# Memory enhancement via mnemonic scaffolds

by

**Felicitas E. Kluger**

A thesis submitted in partial fulfillment of the requirements for the degree of

**Doctor of Philosophy**

Department of Psychology
University of Alberta

# Abstract

Mnemonic techniques can enable learners to memorize vast amounts of information at first encounter. The most effective mnemonic techniques harness mnemonic scaffolds, memory structures of prior knowledge, to which new information is associated. As impressively demonstrated by memory athletes, training in mnemonic scaffolds can lead to superior memory performance and greatly exceed untrained levels of human memory. Importantly, memory athletes neither have extraordinary brain anatomy nor innate memory capacity— their superior memory is attributed to skilled use of mnemonic scaffolds.

Here, we investigate mnemonic scaffolds in novice learners. Our studies include four novel mnemonic scaffolds consisting of different types of prior knowledge and the renowned Method of Loci. Our participants generated their own scaffolds and used them to study word lists in serial order. In addition to the serial recall task, they performed a scaffold-cued recall task, in which they were shown parts of their scaffold as cues to verify whether they used their scaffolds as instructed.

In Chapter 2, we introduce the Autobiographical Story Scaffold. In Chapter 3, we compare autobiographical to fictional stories as mnemonic scaffolds for lists of fifteen words.

In Chapter 4, we compare the Body Scaffold, Autobiographical Story Scaffold, and Routine Activity Scaffold to the Method of Loci and ask whether individual differences in visual imagery and body responsiveness contribute to their effectiveness. We also ask whether different levels of engagement of the body predict the success of the Body Scaffold.

Unlike previous studies, including a formal scaffold-generation phase and a scaffold-cued recall task in all our experiments ensures that success with the strategy can be attributed to participants actually implementing the strategy as instructed. In addition, ours is the

first within-experiment comparison that has revealed differences between different kinds of scaffolds, suggesting that the scaffolds, themselves, could be optimized.

Across all our experiments we have found that not all scaffolds are equally effective, and that the Body Scaffold is on par with the Method of Loci. The ability to form individual scaffold-word associations is a driving factor behind the successful use of mnemonic scaffolds. Embodied cognition, imagined navigation, and visual imagery aptitude may not contribute to their effectiveness. With further fine-tuning of the scaffolds and the way they are instructed, mnemonic scaffolds can be used to greatly boost learning performance.

# Preface

This thesis is an original work by Felicitas Kluger. All research projects contributing to this work received ethics approval from a University of Alberta Research Ethics Board, project names ”Structure of Human Memory” (Pro00105383) and ”Organisation and Retrieval Time-course of Human Memory across Lifespan” (Pro00014801). The study presented in Chapter 4 is an Accepted Manuscript of an article published by Taylor & Francis in Memory on 29 Mar 2022, available online: https://www.tandfonline.com/doi/10.1080/09658211.2022.2052322. I conceived and designed this research with Debby M. Oladimeji, Yuwei Tan, Norman R. Brown, and Jeremy B. Caplan. I also conducted the data analysis, interpreted the results, prepared the figures, wrote the manuscript, and revised it. The co-authors also contributed to data analysis, interpretation of results and edited the manuscript along with Jeremy B. Caplan. Debby M. Oladimeji, Norman R. Brown, and Jeremy B. Caplan contributed to the conceptualization and data interpretation of the experiments in the Chapters 2 and 3. Jeremy B. Caplan edited these chapters.

# Acknowledgements

The work of my dissertation would not have been possible without the tremendous support of many people. First and foremost, I would like to thank my advisor, Dr. Jeremy Caplan, for giving me the opportunity to do this extremely interesting research and for giving me enormous support, great mentorship, and continuous encouragement throughout every step of this thesis. I would also like to thank Dr. Norman Brown for our interesting discussions and valuable advice. Many thanks to my supervisory committee, Dr. Peggy St. Jaques and Dr. Elena Nicholadis for their mindful advice, and encouragement and to my external examiner Dr. Myra Fernandes for the insightful feedback. Special thanks to Debby Oladimeji for her tremendous help and inspiring ideas on the Autobiographical Mnemonic Scaffold and for our vibrant discussions on adapting it for older adults. Special thanks also to Yuwei Tan for her help and valuable discussions on the Body Scaffold. I also thank Dr. Yang Liu and my fellow lab members, Dr. Sucheta Chakravarty, Jeremy Thomas, Tamari Shalamberidze, Ryan Moukhaiber, and many undergraduate research assistants for their support and friendships. I would also like to acknowledge the funding supports that made my research possible, the Natural Sciences and Engineering Research Council of Canada (NSERC), and Department of Psychology of the University of Alberta. I'm also thankful to Dr. Boris Nikolai Konrad and Dr. James Patterson for sharing their insights as memory athletes from their perspectives as a researcher and teacher, respectively. Thanks to the high school students in Munich and the Chiemgau for trying out the Autobiographical Mnemonic Scaffold in the real world and their fascinating insights from their experience. Additionally, I'd like to thank my friends including my amazing Edmonton roommates Darren, Kenny, Ross, Paul, and Garth for supporting and encouraging me throughout my journey. Finally, I would like to thank my parents for supporting me in so many ways.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1

# Mnemonic scaffolds: Anchoring new information onto prior knowledge

The core idea motivating this work is that mnemonic scaffolds, memory structures consisting of ordered prior knowledge, can enable learners to memorize vast amounts of information at first encounter. In fact, mnemonic techniques that rely on ordered prior knowledge have been shown to drastically improve memory to levels that exceed untrained memory performance by a factor greater than ten (Ericsson et al., 1980; Staszewski, 1990). This has been demonstrated in empirical studies (Ericsson et al., 1980; Staszewski, 1990) and at the World Memory Championships (e.g., Foer, 2011). Counterintuitively, extraordinary memory skills do not appear to result from superior cognitive aptitudes or extraordinary brain anatomy (Chase & Ericsson, 1981; Ericsson & Kintsch, 1995; Ericsson & Staszewski, 1989; Maguire, Valentine, Wilding, & Kapur, 2003; Wilding & Valentine, 2006). Rather, they are attributed to training in mnemonic techniques (for a review, see Ericsson, 2003).

Despite their proven benefit in boosting memory performance, mnemonic scaffolds are underexploited in educational settings and for memory-impaired individuals (e.g., J. Richardson, 1995; Twomey & Kroneisen, 2021; Worthen & Hunt, 2011). The underlying cognitive mechanisms of mnemonic scaffolds are unknown. In this dissertation, we ask under which conditions mnemonic scaffolds can enhance memory for novice learners and what driving factors behind their effectiveness are, and whether their effectiveness depends on individual differences in learner skills or affinities.

In Chapters 2 and 3 we focus on novel, story-based mnemonic scaffolds and in Chapter 4, we compare four mnemonic scaffolds to a control condition.

This introductory chapter is structured as follows. We start with a definition of mnemonic scaffolds and provide a brief overview of the general task of our experiments. Then, theories and previous research on the role of prior knowledge for learning new information will be synthesized. This is followed by an overview of how different types of prior knowledge that are investigated in this dissertation may serve as effective mnemonic scaffolds. Lastly, the role of individual differences in learner aptitude for the effectiveness of mnemonic scaffolds will be examined, before the goals, research questions, and following chapters will be outlined.

## 1.1 Mnemonic scaffolds: terminology and general task

The current literature on mnemonic techniques lacks a clear terminological distinction of different techniques, partly due to the fact that the underlying cognitive mechanisms are not well understood. In Skilled Memory Theory, Chase and Ericsson (1981, 1982) argue that structures of existing memory representations providing retrieval cues are central to the effectiveness of mnemonic techniques (see also Roediger, 1980; Wenger & Payne, 1995; Bellezza, 1981; Ericsson & Staszewski, 1989). Different terms have been used to describe these structures of existing memory representations. We use the term "mnemonic scaffolds" to refer to predetermined prior knowledge that is used to remember new information. During study, a mnemonic scaffold provides a system of anchors to which new information is associated. During recall, the scaffold provides those anchors as a set of ordered retrieval cues. The cognitive and neural processes underlying the effectiveness of mnemonic scaffolds are unclear. In Skilled Memory Theory, Chase and Ericsson (1981, 1982), use the related term retrieval structures to refer to hierarchically organized systems of retrieval cues. The concept is very similar to what we define as mnemonic scaffold in that a structured system of anchors is used for associations during study, which serve as retrieval cues during recall. This idea is also captured by the terms "systems of retrieval cues" (Roediger, 1980), "cognitive cuing structures" (Bellezza & Bower, 1981; Bellezza & Buck, 1988), or "scaffolding strategies" (Bouffard, Stokes, Kramer, & Ekstrom, 2017). The difference between what we define as mnemonic scaffolds and what Chase and Ericsson (1981, 1982) define as retrieval structures is that mnemonic scaffolds are a) selected by the learner before studying new items, and b) have a linear rather than a hierarchical order.

The general design used in our experiments tested mnemonic scaffolds in novice learners. We give a brief general overview of the testing sessions below. For a detailed description of the experiments and justification of this approach, refer to the following chapters. Our target task comprises four phases.

1. Participants type their mnemonic scaffold, which consists of ordered prior knowledge. The type of prior knowledge used depends on the specific instructions and experimental design.

2. Participants are shown their scaffold together with a study item following the inherent order of the scaffold and are asked to associate one anchor with one study item at a time.

3. Participants are asked to recall the words in the same order they were presented. In this serial recall phase, the anchors from the scaffolds are to be used as retrieval cues, as participants are asked to go through their scaffold in order to recall the words.

4. Participants are shown the anchors from their scaffold in random order and asked for the word they associated with it. This scaffold-cued phase serves as a verification of whether participants have used the scaffold strategies as instructed.

## 1.2 Anchoring novel information to prior knowledge

The idea that memory for novel information can be enhanced by anchoring it to prior knowledge has been proposed within different frameworks. In what later became the foundation of Schema Theory, Bartlett (1932) proposed that prior knowledge affects the processing of new stimuli and that the activation of specific schemata biases memory retrieval towards such schemata. Schemata are referred to as abstract knowledge structures stored in semantic memory (Alba & Hasher, 1983). Later proponents of Schema Theory argued that such structures provide a framework to organize incoming information and encourage selective attention (Rumelhart, 1980; Schank & Abelson, 1977). This allows anchoring of congruent incoming information to pre-existing knowledge stores, which enhances encoding and subsequent retrieval (Anderson et al., 1976; Alba & Hasher, 1983). Conversely, if novel information is inconsistent with the activated structures, encoding and retrieval become less effective and more effortful (Arbuckle, Vanderleck, Harsany, & Lapidus, 1990). At the neural level, schemata have been related to strongly interconnected, "semanticized" neocortical representations (Van Kesteren, Ruiter, Fernández, & Henson, 2012). Activation of such semantic neural networks affects processing of novel related information so that it can be easily integrated into the associative networks (Coutanche, Thompson-Schill, Sharon, Moscovitch, & Gilboa, 2014; Lewis & Durrant, 2011; Sharon, Moscovitch, & Gilboa, 2011).

The notion that prior knowledge affects learning can also be identified within the Levels of Processing framework (Craik & Lockhart, 1972). According to this view, episodic encoding is facilitated when new information is processed in a meaningful way that engages the learner on a deeper cognitive level (Craik & Lockhart, 1972). Thus, relevant prior knowledge might enable meaningful encoding on such a "deeper" level (Craik & Lockhart, 1972). While Levels of Processing has had profound influence on modern memory research, it should be noted that the link between deep or elaborative processing and retention may be correlational rather than causal (Nairne, 2002).

A prominent example of superior memory for newly learned information that is anchored to prior knowledge is expert memory (Ericsson & Staszewski, 1989). Experts have superior memory for novel information within their area of expertise (e.g., Long & Prat, 2002; Brandt, Cooper, & Dewhurst, 2005; Van Kesteren et al., 2012; Lane & Chang, 2018; Ericsson & Staszewski, 1989; Bruett, Fang, Kamaraj, Haley, & Coutanche, 2018). Ericsson and Kintsch (1995) proposed that prior knowledge improves memory for new information by enabling more effective organizational processing, which allows new information to be associated with retrieval cues and to be integrated into the existing associative network (see also Long & Prat, 2002).

More recent cognitive theories of the prior knowledge effect refer to the idea of semanticization or systems consolidation. Even though episodic (knowledge embedded in its spatiotemporal context) and semantic (knowledge about information independent of its context) memories are dissociable entities that rely on different sets of brain regions (Tulving, 2002), both memory systems closely interact as newly encoded information is affected by preexisting knowledge (McKenzie & Eichenbaum, 2011). Cognitive theories suggest that semantic memories are initially episodic, and over time and with repeated retrieval, they are semanticized, i.e., essential features are extracted and the memories become de-contextualized (Raaijmakers, 1993; Carr et al., 1994). On the neural level, semanticization or systems consolidation has been hypothesized to account for the large-scale reorganization of episodic memories over time (Meeter & Murre, 2004; Winocur & Moscovitch, 2011). According to the Trace Transformation Hypothesis (Winocur & Moscovitch, 2011), episodic memories transform into semantic versions that are represented in extra-hippocampal structures. Insofar as episodic memories are retained, they will continue to rely on the hippocampus, but for retrieval of semantic memories, the hippocampus is not needed (McClelland, McNaughton, & O'Reilly, 1995; Meeter & Murre, 2004; O'Reilly, Bhattacharyya, Howard, & Ketz, 2014). Similarly, the Complementary Learning Systems framework proposes that repetitive reactivation of the memory trace results in a transformation from hippocampal to neocortical retrieval (McClelland et al., 1995; O'Reilly et al., 2014). Novel related information can be rapidly integrated into such semanticized knowledge, leading to superior memory. In a series of studies by Tse et al. (2007, 2011), rodents were trained over several weeks to become experts in locating food in an event arena by repetitively learning 6 odor–place "paired associates" within a grid of 7 × 7 locations. The overlearning of these paired associates rendered memory for them hippocampus-independent (Tse et al., 2007). Novel odor–place associates within the same grid of locations could then successfully be learned with only a single trial, reflecting a prior-knowledge effect. Neocortical regions, in particular, the medial prefrontal cortex and retrosplenial cortex, supported the integration of novel “paired associates” into the existing associative structure. Importantly, prior knowledge also accelerated the transfer of hippocampal memories to neocortex, as recall of the novel integrated odor–place associations was already hippocampus-independent after 24 hours (Tse et al., 2007). Sommer (2017) built on these studies in a longitudinal functional magnetic resonance imaging study in humans. For ten associative structures, each consisting of an irregular layout of 20 locations, participants first overlearned 12 object–location associations. This phase was associated with a shift from hippocampal to ventrolateral prefrontal cortex-mediated retrieval, consistent with semantization theories (Meeter & Murre, 2004; Winocur & Moscovitch, 2011). When participants encoded novel object–location associations, their encoding was associated with ventromedial prefrontal activity (Sommer, 2017). A study found that amnesic patient H.M. could acquire new semantic knowledge when he could relate it to memories established pre-operatively (Skotko et al., 2004), implying the neural substrate for this learning must have been outside the hippocampus. Kan, Alexander, and Verfaellie (2009) found that the potential for eight amnesics patients to benefit from prior knowledge depends on the integrity of their knowledge scaffolding, as inferred from lesion extent. Patients with intact prior knowledge experienced a benefit when learning novel related information, whereas patients with weakened prior knowledge did not, even though these groups did not differ on two baseline measures of episodic learning (Kan et al., 2009). Taken together, these findings suggest that prior knowledge structures can support acquisition of new episodic information by providing extra-hippocampal frameworks into which new information can be incorporated. From this view, mnemonic scaffolds can be seen as extra-hippocampal frameworks that in addition to providing retrieval cues during recall, enable the integration of new information by anchoring it onto existing semanticized knowledge during study.

According to Coutanche et al. (2014) a learning procedure known as “fast mapping,” in which linguistic labels are mapped to referents in the world with only minimal exposure also illustrates the importance of prior knowledge and extra-hippocampal processes for learning new information. In a fast mapping study with amnesic patients, Sharon et al. (2011) found that rapid acquisition of declarative memory can be accomplished independently of the hippocampus, and fast mapping may be mediated by extra-hippocampal neocortical structures that are implicated in semantic memory, such as the lateral temporal lobe and the anterior temporal lobe, as well as the inferior prefrontal cortex (Smith, Urgolites, Hopkins, & Squire, 2014). Coutanche et al. (2014) found that presenting an already-known item during learning was crucial for rapid integration through fast mapping, and conclude that the retrieval of a related known concept is critical for this effect. They also found that object–word associations learned via fast-mapping had semantic-like behavioural characteristics, adding support for the idea that fast mapping could indeed be a quick route to semanticization.

Taken together, in the light of recent research on the neural and cognitive mechanisms underlying fast mapping and the prior knowledge effect, it is a compelling idea that extra-hippocampal learning processes involving prior knowledge may contribute to the effectiveness of mnemonic scaffolds. This leads to our general hypothesis, that all scaffolds generated from prior knowledge should support serial recall far above a no-scaffold control condition (see Chapter 4).

## 1.3 The Method of Loci: Geospatial information in mnemonic scaffolds

By far the most investigated mnemonic scaffolds consist of navigational routes, as used in the famous Method of Loci, also called Memory Palace or Mind Palace Technique (e.g., Foer, 2011; Spence, 1984). This associative encoding technique dates back to classical Greek and Roman rhetorical traditions (Yates, 1966) and is the most common mnemonic technique used by modern memory athletes (Foer, 2011). There are marked individual differences in navigational skills in both real-world and virtual environments (Weisberg, Schinazi, Newcombe, Shipley, & Epstein, 2014; Blajenkova, Motes, & Kozhevnikov, 2005; Chai & Jacobs, 2010; Ishikawa & Montello, 2006; Wen, Ishikawa, & Sato, 2011). While this assumption has not been tested, it is plausible that people with low navigational skills are disadvantaged when using the Method of Loci, which emphasizes the relevance of investigating alternatives to navigational mnemonic scaffolds. Some researchers have argued that visuospatial navigation and the engagement of the medial temporal lobe system are a determining factor in the memory benefit provided by this method, due to the dual role of this network in navigation and episodic memory (e.g., Rolls, 2017; Moser, Rowland, & Moser, 2015). In a combined EEG and fMRI study that compared the Method of Loci to a non-numerical peg-list method consisting of 20 pegs, Fellner et al. (2016) found that the more effective Method of Loci induced a pronounced theta power decrease source-localized to the left medial temporal lobe compared with the nonspatial peg method, mirrored by activity increases in the medial temporal lobe. Theta oscillations have been implicated in spatial processing and navigation (e.g., Ekstrom et al., 2005). The authors suggested this indicates that the Method of Loci may be a particularly efficient encoding strategy by driving the neural processes related to spatial processing and episodic memory. Other findings cast doubt on this, suggesting that navigational cognition may be epiphenomenal, or at least not necessary to excel with the technique (Bouffard et al., 2017; Bower, 1970; Caplan, Legge, Cheng, & Madan, 2019). Instead, these researchers have suggested the effectiveness of the Method of Loci might derive from engaging the learner with the study material in much the same way as other mnemonic scaffolds or peg systems. Bower (1970) noted that numerical peg systems, which use a predefined system of images and numbers as a basis to associate to-be-remembered items, can produce equivalent performance to the Method of Loci, corroborated by Roediger (1980). Bouffard et al. (2017) found that participants' recall performance showed memory boosts comparable to the Method of Loci for strategies based on a sequence of events, suggesting that spatial and non-spatial representations and a series of events can be used to enhance memory performance in a similar way. Caplan et al. (2019) trained naïve participants on virtual environments with particular navigation-based characteristics (conceptual familiarity, boundary, and lines of sight), and found that memory success was not largely determined by navigation-relevant features of the environment, nor knowledge of the environment.

To shed some light on this unresolved controversy around the role of locations for the memory success of mnemonic scaffolds, we ask whether the number of locations mentioned in story-based mnemonic scaffolds predict its memory benefit in Chapter 2. In Chapter 4, we directly compare the Method of Loci to non-navigational mnemonic scaffolds.

## 1.4 The Body Scaffold: Embodiment in mnemonic scaffolds

While the Method Loci is by far the most widely used mnemonic scaffold (e.g., Foer, 2011), there are historical and contemporary anecdotal accounts on a mnemonic scaffold based on the human body (Hunter, 1956), which is also used and recommended by some memory athletes (e.g., Foer, 2011; Konrad, 2013). Despite this, there is, to our knowledge, almost no research on using the human body as a memory aid, and the question of whether embodied cognition may play a role is unanswered. Embodied cognition research suggests that physical actions and sensory experiences shape our mental experience (Barsalou, 1999; Lakoff & Johnson, 1999; Niedenthal, 2007). Embodiment, specifically, refers to the notion that cognition depends on the sensorimotor capacities of the human body and that sensory and motor processes are inseparable in cognition (Varela, Thompson, & Rosch, 1991). Embodied cognition researchers emphasize the notion of "simulation" (Gallese, 2008; Zwaan, 2004), referring to internally representing the meaning of a verbal stimulus. Accordingly, the simulation process involves the same sensorimotor neural correlates that are active during the action execution or interaction with the actual object or entity itself (Zwaan, 2004). Providing explanations for embodied simulation, behavioral and neural evidence has shown that the process of language comprehension elicits activation within primary and secondary motor areas (Pulvermüller, 2005; Barsalou, 2008; Fischer & Zwaan, 2008; Toni, de Lange, Noordzij, & Hagoort, 2008; Handy, Grafton, Shroff, Ketay, & Gazzaniga, 2003). This is exemplified by Sakreida et al.'s (2013) work on the influence of reading on the sensorimotor neural network. Participants were shown concrete nouns (e.g., bucket), non-graspable nouns (e.g., hope), motor verbs (e.g., kick), and non-motor verbs (e.g., marvel) and were instructed to press a button while fMRI was recorded, if the sentence referred to an action performed by the foot or leg (e.g., to kick the ball; Sakreida et al., 2013). The sensorimotor neural network was active during mental simulation as well as when executing an action (Sakreida et al., 2013), showing that body-specific activity is present both during action execution and verbal processing. In the context of memory, D. C. Richardson, Spivey, Cheung, and Cheung (2001) have shown that the representation of a visual stimulus retrieved from memory can activate potential motor interactions, and that memory representations derived from linguistic descriptions can also activate motor affordances. Zimmer and Cohen (2001) argue that sensorimotor details lead to better memory performance due to better encoding elaboration enabling association with preexisting memory representations.

In the only study, to our knowledge, in which body parts were used as memory cues, Bellezza (1984) presented participants with nouns and asked them to come up with a body part or a personal experience that they deem a fitting memory cue for the respective study item. In a free recall task, participants then recalled both the study items and the body part or personal experience they had chosen as a memory cue. No difference in recall using the two types of cues was found (Bellezza, 1984). As the study did not include a control condition, it is unclear whether these memory cues facilitated recall. In contrast to our study, in Bellezza's (1984) study, the human body was not used as a mnemonic scaffold, but individual body parts were selected as cues after viewing the study item. Thus, the advantage of providing a sequence of retrieval cues in a fixed order was dismissed. Therefore, it remains unknown whether prior knowledge in the form of body parts provides a mnemonic benefit as the Body Scaffold and the role of embodied cognition in mnemonic techniques has not been investigated.

There are, however, historical and contemporary anecdotal accounts of a mnemonic scaffold based on the human body. In his memory training book *Plutosofia* Gesualdo (1592) describes how to remember information by associating it with parts of the human body. Some memory athletes describe using their own body to remember information by associating study items with body parts (e.g., Foer, 2011; Konrad, 2013). To our knowledge, the experiments reported in Chapter 4 are the first empirical studies of memory performance using a body-based mnemonic technique. Although we had no *a priori* reason to expect the Body Scaffold to excel to the same level as the Method of Loci, we had wondered if emphasizing embodiment, or individual differences in tendency toward embodiment, might drive the success of the Body Scaffold (Chapter 4).

## 1.5 Story-based scaffolds: autobiographical and fictional stories as mnemonic scaffolds

In story-based mnemonic scaffolds, fictional or autobiographical stories function as prior knowledge to which the study items are associated. Participants first recall a story and are then instructed to incorporate study items into their sentences, imagining that the story actually happened this way. To our knowledge, this type of mnemonic scaffold has never been investigated.

The story-based mnemonic scaffolds of the present study are not to be confused with the story mnemonic described in reviews by Bower and Clark (1969); Bellezza (1983, 1986) and Worthen and Hunt (2008, 2011). The authors use the term "story mnemonic" to refer to a narrative chaining method similar to the Link Method (Bower & Gilligan, 1979), where learners construct a meaningful story woven around the study items. Bellezza and Bower (1981) classify this technique as a "chain-type" mnemonic where recall is based on cues which themselves are part of the studied list. This means that one word is used as the cue to recall the next word. This is referred to as intrinsic cueing because the study items dictate the content of the story and the cues belong to the studied list (Bellezza & Bower, 1981). In our story-based mnemonic scaffolds, in contrast, the stories are produced before participants view the study items and thus provide external anchors to which the study items are associated. In one of the very few studies that directly compare different mnemonic strategies, Roediger (1980) showed that mnemonic scaffolds providing extrinsic cues are more effective than the Link Method, a chain-type mnemonic which uses intrinsic cueing. While little is known about the role of autobiographical memories in mnemonic techniques, research and theory on the self-reference effect provide important theoretical arguments that autobiographical memories may serve as effective mnemonic scaffolds. In a series of studies, Rogers, Kuiper, and Kirker (1977) extended Levels of Processing Theory (Craik & Lockhart, 1972) to the realm of the self. Depth of processing research had already demonstrated that semantic-encoding tasks ("Does the word mean the same as x?") resulted in higher recall compared with phonemic ("Does the word rhyme with x?") or graphemic ("Does the word have capital letters?") tasks (Craik & Tulving, 1975). Rogers et al. (1977) found that self-referential tasks (e.g., "Does the word describe you?") produced even better memory than semantic orienting tasks. The authors concluded that the self acts as a "superordinate schema" (p. 686) that facilitates encoding and retrieval (Rogers et al., 1977). Subsequent studies using a variety of encoding tasks including autobiographical memory and imagery (Bower & Gilligan, 1979; Brown, Collins, & Duguid, 1989) confirmed this notion. A meta-analysis by Symons and Johnson (1997) highlights the importance of the self-reference effect in memory, emphasizing that self-referent encoding tasks yield superior memory in free recall, cued recall and recognition tasks relative to both semantic and other-referent encoding tasks. Symons and Johnson (1997) conclude that this is because the self is a well-developed and often-used construct that promotes elaboration and organization of encoded information. In addition, since autobiographical memories are highly self-relevant and rich in detail they possibly invoke extra-hippocampal structures, supplementing the function of the hippocampus (e.g., Cabeza & St Jacques, 2007). Thus, due to their highly self-relevant nature, we expect that autobiographical events can serve as effective mnemonic scaffolds.

In a behavioral study of hippocampal function (to our knowledge, the only study on mnemonic strategies that includes autobiographical memories), Bouffard et al. (2017) compared an autobiographical, so-called "temporal" scaffold, consisting of a timeline of autobiographical events to the Method of Loci. Participants were instructed to create a chronological timeline using ten of their most memorable memories (Bouffard et al., 2017). Their participants' final recall performance showed a similar memory increase for the Method of Loci and autobiographical timelines, suggesting that spatial and temporal representations can be used to enhance memory performance in a similar way (Bouffard et al., 2017). It is important to note, however, the focus of this study was on time-based scaffolds, not autobiographical memory, itself. In our experiments, we developed a novel autobiographical technique, in which single autobiographical events *per se* comprise the mnemonic scaffold, more in line with how participants might spontaneously remember events from their lives.

To test whether the self-relevance effect translates to autobiographical mnemonic scaffolds and renders them more effective than non-autobiographical scaffolds, we compare autobiographical to fictional stories as mnemonic scaffolds in Chapter 3.

A potential disadvantage of using autobiographical stories as mnemonic scaffolds might be that recall of autobiographical narratives is not stable and based on the narrative and other factors (e.g., Greenberg & Rubin, 2003; Habermas, 2018; Hirst & Echterhoff, 2011; McAdams & McLean, 2013), including individual abilities (Rubin, 2020, 2021). To investigate whether stories that are recalled more reliably than stories recalled with lower accuracy function as more effective mnemonic scaffolds, we tested whether recall accuracy of the story scaffolds themselves has an effect on recall accuracy of the study items in Chapter 3.

## 1.6 The Routine Activity Scaffold: Routine activities in mnemonic scaffolds

As a third alternative to the Method of Loci, we hypothesize that routine activities may provide effective mnemonic scaffolds. Routine activities are central to Script Theory, the assumption that part of our knowledge and cognitive processes is organized around hundreds of stereotypical situations (Schank & Abelson, 1977). Script Theory, built on Schema Theory (Bartlett, 1932), was proposed by Schank and Abelson (1977) to unify central notions of memory and knowledge acquisition in developmental, clinical, social, and cognitive psychology (Abelson, 1981). A dynamic version of schemata comprising activities, scripts are defined as organized knowledge stores which consist of routine activities and serve as a base for elaborations surrounding a topic (Bower, 1970). In other words, Schank and Abelson (1977) refer to scripts as cultural stereotypes that people acquire through direct or vicarious experiences along with idiosyncratic variations that provide preexisting knowledge structures involving event sequences. The classic example of a script is going to a restaurant, which contains a standard sequence of events characterizing typical activities in a restaurant (Schank & Abelson, 1977). Because scripts incorporate causal and practical constraints on order, it seemed plausible that routine activities might be excellent scaffolds for serial recall.

Considering the large body of literature on routine activities and knowledge acquisition via prior knowledge in the form of schemata and scripts, we were surprised that we could not find any studies that used routine activities for mnemonic purposes, with the exception of Bouffard et al. (2017). Their study, however, only included one routine activity, the steps to making a sandwich, to investigate whether sequences with easily accessible temporal features provide similar memory boosts as the Method of Loci and timelines of autobiographical events (as described above). The authors argued that the steps of making a sandwich are fixed in time, well-rehearsed, and familiar to most people and may therefore have a stronger temporal connection than autobiographical memories have with each other (Bouffard et al., 2017). The steps of making a sandwich resulted in similar memory performance as autobiographical timelines or the Method of Loci (Bouffard et al., 2017). Based on these findings and the notion that routine activities are well rehearsed and highly familiar, we expected routine activities to be effective mnemonic scaffolds, particularly due to their natural constraints on order (Chapter 4).

## 1.7 Individual differences in learner aptitude and the usefulness of mnemonic scaffolds

Viewing mnemonic strategies as skills (Ericsson et al., 1980), one might expect individual differences, in particular, cognitive abilities or affinities to particular kinds of information, might determine how well a participant can excel with a particular scaffold. Despite its importance for the application of mnemonic strategies in educational and cognitive rehabilitation settings, the role of individual differences in the usefulness of mnemonic scaffolds has received almost no scientific attention. One notable exception, Sanchez (2019), found evidence suggestive that effective usage of the Method of Loci was dependent on participants' visuospatial ability, measured with the Cube Comparisons Task (CCT) and Paper Folding

Task (PFT), both drawn from French, Ekstrom, and Price (1963). In fact, those lower in visuospatial aptitudes may actually have been disadvantaged by using navigational scaffolds for serial recall (Sanchez, 2019). The precise interpretation of the result is more subtle, and we revisit this study in Chapter 4. Aside from Sanchez' (2019) study that only examined the Method of Loci and visuospatial ability, the experiments in Chapter 4 are the first study investigating different effects of individual differences in the usefulness of the Method of Loci in addition to non-navigational mnemonic scaffolds. In an attempt to replicate Sanchez' (2019) findings, we used the PFT to measure visuospatial aptitude (Chapter 4). In addition, we used the Vividness of Visual Imagery Questionnaire (VVIQ, Marks, 1973) to assess the self-reported vividness of participants' imagery. In Chapter 4, we explain why we used those self-report tools.

There is some anecdotal evidence that visual imagery skill or vividness might determine the effectiveness of some mnemonic scaffolds. Many memory athletes contend that "thinking in images," i.e., vivid visual imagery, is key to successful application of mnemonic strategies (e.g., Foer, 2011; Konrad, 2013). This introspection of world-class mnemonic strategy users, however, has not been confirmed by research, and at least two studies that addressed this question found no relationship between vividness of visual imagery and success with the Method of Loci (Kliegl, Smith, & Baltes, 1990 and McKellar, Marks and Barron reported by Marks, 1972a, p. 96).

Notably, visual imagery capacity appears to vary greatly between individuals, with aphantasics reporting a lack of the ability to create visual images (Keogh & Pearson, 2018). It therefore stands to reason that the relationship between individual differences in visual imagery and usefulness of visual-based mnemonic strategies might have important practical and theoretical implications. Interestingly, De Beni and Cornoldi (1985) found that congenitally blind participants can perform well with the Method of Loci (De Beni & Cornoldi, 1985). This suggests that visual imagery may not be the basic reason why imagery-based strategies are effective. In addition to the VVIQ and PFT, we used the Body Responsiveness Questionnaire (BRQ, Daubenmier, 2005) to assess participants' awareness of internal body sensations expecting it to relate to success with the Body Scaffold (Chapter 4).

## 1.8 Recall accuracy of the scaffolds themselves and order in mnemonic scaffolds

Previous research on mnemonic scaffolds has shown that the strength of mnemonic scaffolds lies in facilitating memory for ordered information (Bouffard et al., 2017; Foer, 2011; Ericsson et al., 1980; Roediger, 1980; Yates, 1966). In serial recall, memory for order is commonly investigated with a strict position scoring criterion under which items are only scored as correct when recalled in the same position they were presented. Memory for items is scored with a lenient position scoring criterion under which an item is scored as correct if it was recalled in any position of the list. Bouffard et al. (2017), Roediger (1980) and our experiments described in the following chapters also show that the mnemonic advantage of mnemonic scaffolds over the control condition is higher when scored under the strict scoring criterion than under the lenient scoring criterion. This raises the question of whether mnemonic scaffolds that have a stable internal order and the recall accuracy of which is consequently higher are more effective than mnemonic scaffolds where the order is not easily retrieved. While this question, to our knowledge, has never been tested directly, findings that the Method of Loci and the Body Scaffold, the order of which we consider as stable (in the Body Scaffold the order of anchors is prescribed by the invariable order of body parts, and in the Method of Loci the order can be retrieved by following a fixed route without backtracking) outperform scaffolds consisting of autobiographical stories, the order of which is variable (autobiographical events are not reliably retrieved in a chronological order, e.g., E. F. Loftus & Fathi, 1985) suggest that order influences the effectiveness of mnemonic scaffolds (Chapter 4). To investigate whether memory for the stories predicts the memory success of story-based mnemonic scaffolds, participants in Chapter 3 completed a recall attempt of their scaffolds. This allows us to test whether recall accuracy of the scaffold itself predicts recall accuracy of the study items.

## 1.9 Goals, research questions, and chapter overview

The overarching goal of this dissertation is to test conditions under which mnemonic scaffolds can enhance memory in novice learners and to investigate driving factors behind their effectiveness. We ask four main research questions:

1. Under which conditions can story-based mnemonic scaffolds enhance memory for lists of words?
2. What are alternatives to the Method of Loci?
3. Do individual differences in visual ability and body responsiveness contribute to the effectiveness of mnemonic scaffolds?
4. Does engagement of the body contribute the effectiveness of the Body Scaffold?

Here, we provide an overview of how the following chapters answer these research questions. Chapters 2 and 3 focus on the first main research question. In Chapter 2, we introduce the story-based mnemonic scaffold and we compare four different presentation modes for mnemonic scaffolds consisting of autobiographical stories and ask whether spatial locations mentioned within those stories increase their effectiveness in facilitating serial recall. We identified the presentation mode in which one word and one part of the scaffold are shown on the same screen at a time as the most effective one. This is because compliance rates were the highest compared to the other presentation modes and participants in this group outperformed Control and the other groups in recency positions. Consequently, we used this presentation mode of showing one word and one part of the scaffold at a time for all following experiments reported in this dissertation. The experiment of Chapter 2 also showed that the number of locations mentioned in the stories that were used with the most effective presentation mode predicted memory success. Additionally, scaffold-word associations showed a strong relationship with serial recall accuracy in most groups, suggesting that scaffold-word associations are important for the effectiveness of mnemonic scaffolds. We followed up on these effects in Chapter 3.

Since the advantage of the Autobiographical Story Scaffold was observed in recency positions, we increased list length from ten to fifteen in Chapter 3. Using the most effective presentation mode identified in Chapter 2, we compared autobiographical to fictional scaffolds and to a control condition that was instructed to say the words out loud. In addition to testing whether the number of locations mentioned in the autobiographical and fictional stories predicts the success with which participants use them for serial recall, we also tested whether recall accuracy of the story scaffolds themselves predicts serial recall accuracy of the study items. Since we found no advantage of neither the autobiographical nor the fictional story scaffold neither over baseline memory (i.e., participants recall accuracy before they were instructed to use mnemonic scaffolds) nor the control condition, the conditions to test our hypotheses were not met. We concluded that list length may be a boundary condition for the effectiveness of story-based mnemonic scaffolds in novice learners. We found that scaffold-word associations predicted serial recall accuracy of the study items, suggesting that some participants may have relied upon scaffold-word associations when recalling the study items.

Chapter 4 provides answers to research questions 2 to 4 and includes two more experiments. In Experiment 1 of Chapter 4, we compare mnemonic scaffolds consisting of a) parts of the human body, b) autobiographical stories, and c) steps of routine activities to the Method of Loci and to a control condition for lists of ten words, and test whether visual ability and body responsiveness contribute to their effectiveness. We found that the Body Scaffold is equally effective as the Method of Loci and that the Autobiographical Story Scaffold provides a mnemonic benefit over baseline memory, while the Routine Activity Scaffold does not improve memory. We found no relationship between our individual differences measures and effectiveness of the scaffolds. As in Chapter 2, we found evidence that individual scaffold-word associations are important for the effectiveness of mnemonic scaffolds. Surprised by the high level of success participants had with the Body Scaffold, we asked if this is based on engagement of the body in Chapter 4. We found that varying levels of embodiment do not have an effect on the effectiveness of the Body Scaffold.

In Chapter 5, we summarize the findings from the previous Chapters. Given the richness of the data from the scaffold-cued recall task, we discuss in how far individual scaffold-word associations are a driving factor behind the effectiveness of mnemonic scaffolds. We then further discuss the implications of our findings on advancing theories of memory enhancement via mnemonic scaffolds.

# Chapter 2

# The Autobiographical Story Scaffold

# Abstract

Mnemonic scaffolds, structured memories to which study items are associated, enable memory champions to remember vast amounts of information in order and at first encounter. Previous research is largely restricted to the familiar-route scaffold (Method of Loci). However, there is evidence that this method is not suitable for all learners, and alternatives to the Method of Loci are needed. Here, we describe and test such an alternative, a novel mnemonic scaffold consisting of autobiographical stories. We compared four presentation modes of the scaffolds and study items for lists of ten words and investigated whether mentioning locations in the autobiographical stories increases their effectiveness in facilitating serial recall. The most effective presentation mode, where participants were shown a sentence from their autobiographical stories and a study item at a time, outperformed Control in recency positions, had the highest compliance rates and benefited from spatial locations incorporated into those stories. In short, we showed that Autobiographical Story Scaffolds can enhance memory under certain conditions, and with further development, they could be an alternative for people who struggle with the Method of Loci.

## 2.1 Introduction

An effective way to enhance memory is by the use of mnemonic techniques that provide a scaffold of existing memories to which new information is associated. As demonstrated by world memory champions, mnemonic techniques that use structures of prior knowledge, after sufficient training, enable learners to memorize vast amounts of information at first encounter (e.g., Foer, 2011; Ericsson, 2003). Crucially, superior memory is based on training in mnemonic techniques and does not require superior cognitive aptitudes or extraordinary brain anatomy (Chase & Ericsson, 1981; Ericsson & Staszewski, 1989; Ericsson & Kintsch, 1995; Wilding & Valentine, 2006; Ericsson, 2003; Maguire et al., 2003). Ericsson et al. (1980)

and Staszewski (1990) have shown that training with mnemonic techniques can improve human memory by a factor greater than ten.

Despite this, factors underlying the effectiveness of these techniques are largely under-investigated. As explained in detail in Chapter 1, there is converging evidence that anchoring to-be-remembered information to a scaffold of existing memories in systematic order is key to exceeding normal, i.e., untrained recall. In Skilled Memory Theory, Chase and Ericsson (1981, 1982) propose that the anchors within the scaffold function as retrieval cues during recall and thus facilitate memory (see also Ericsson & Staszewski, 1989; Bellezza, 1981; Roediger, 1980; Wenger & Payne, 1995; Bellezza, 1981). Different terminology has been used to describe these structures of existing memory representations, including retrieval structures (Chase & Ericsson, 1981, 1982), systems of retrieval cues (Roediger, 1980), cognitive cuing structures (Bellezza & Bower, 1981; Bellezza & Buck, 1988), or scaffolding strategies (Bouffard et al., 2017). For simplicity and to capture the idea that an internal order of anchors may be important for a mnemonic scaffold to be effective, we use the term "mnemonic scaffold" (see also Chapters 3 and 4).

Research on mnemonic scaffolds is largely restricted to a particular type of mnemonic scaffold consisting of a route through a familiar environment. This mnemonic scaffold is called the Method of Loci or Mind Palace Technique (e.g., Foer, 2011; Yates, 1966). It uses a route through a familiar environment as a mnemonic scaffold, whereby specific locations (Loci) are associated with study items. For retrieval, the learner imagines navigating along the same route reporting study items along the way (e.g., Foer, 2011). The Method of Loci has been used since the ancient Greeks to facilitate accurate recollection of information in order (Yates, 1966). Even after modest training (e.g., Roediger, 1980) this method produces superlative memory performance and is used for personal memory enhancement (Foer, 2011; Raz et al., 2009) and in memory championships (e.g., Dresler et al., 2017; Maguire et al., 2003). Research has shown that older adults (e.g., Anschutz, Camp, Markley, & Kramer, 1987; Baltes & Kliegl, 1992; Gross & Rebok, 2011; Kliegl et al., 1990; Karbach & Verhaeghen, 2014; Yesavage, Lapp, & Sheikh, 1989), and possibly younger adults with low visuospatial aptitude (Sanchez, 2019) struggle with the Method of Loci. This highlights the need to find alternatives to the Method of Loci (see also Chapter 1 and 4).

One such alternative might be autobiographical stories used as mnemonic scaffolds. My research on the Autobiographical Story Scaffold is motivated by my real-world teaching experience. High school students were instructed to integrate study material such as Latin vocabulary and Grade 10 Biology material into autobiographical stories. Those students reported better memory for their study material when using this strategy. This led to the idea that autobiographical stories might function as an effective alternative to the Method of Loci, and potentially work better for participants who have trouble with the Method of Loci. In Chapter 4 we explain in detail why autobiographical stories may be particularly well suited for seniors.

As reported in Chapter 4, autobiographical stories can provide a mnemonic benefit over uninstructed baseline memory. Research and theory on the self-reference effect suggest that Autobiographical Story Scaffolds, in theory, may have an advantage over non-autobiographical ones (e.g., Bower & Gilligan, 1979; Brown et al., 1989; Rogers et al., 1977; Symons & Johnson, 1997). We describe this in detail in Chapter 4 and also in Chapter 1.

The story-based mnemonic scaffolds described here are not to be confused with the story mnemonic described in reviews by Bellezza (1983, 1986) and Worthen and Hunt (2008, 2011). The authors use the term "story mnemonic" to refer to a narrative chaining method similar to the Link Method, (Bower & Gilligan, 1979), where learners construct a meaningful story woven around the study items. In other words, the study items are linked to one another so that they form a story. When recalling the story, one study item serves as a cue for the next. Bellezza and Bower (1981) classify this technique as a "chain-type" mnemonic where recall is based on cues which themselves are part of the studied list. This is referred to as intrinsic cueing because the study items dictate the content of the story and because the cues belong to the studied list (Bellezza & Bower, 1981). In our story-based mnemonic scaffolds, in contrast, the stories are produced before participants view the study items and thus provide external anchors to which the study items are associated.

### 2.1.1 Presentation modes for Autobiographical Story Scaffolds

It is important to test how different ways of presenting scaffolds and study items may affect the effectiveness of mnemonic scaffolds because this might influence how learners excel with the strategy in several ways. First, having the whole story available during study may be important because not having to remember the story while studying the words might make it easier to use the strategy. In this case, having the whole story on the screen when studying the lists words should outperform the other groups. Second, it might be important that participants' attention is directed to individual sentences, in which case re-typing the story or viewing one word and one sentence at a time should be most effective. Third, it might be that participants' attention needs to be focused on the study items alone without potentially distracting information of the scaffold on the screen. In this case, only showing words during study should be most effective. Fourth, forcing a clear-cut integration of the word into the sentence might be important, in which case participants who retype the words with the sentence integrated should see the highest levels of success with the strategy.

To our knowledge, it has never been tested whether different ways of presenting mnemonic scaffolds and list words during study have an effect on the success with which participants use these scaffolds for serial recall. Previous related studies on mnemonic scaffolds (e.g., Bouffard et al., 2017; Roediger & Crowder, 1976; Dresler et al., 2017) asked participants to study the list words without showing the scaffold. To test how different ways of presenting study items and scaffolds affect memory, we compared those four different presentation modes to each other and to an uninstructed Control.

### 2.1.2 Spatial locations in mnemonic scaffolds

As a learned environment forms the basis for the encoding process during use of the Method of Loci, this mnemonic scaffold is thought to rely on an imagined spatial/navigational memory substrate for study items that are not spatial themselves (Fellner et al., 2016; Müller et al., 2018; Moser et al., 2015; Kondo et al., 2005; Maguire et al., 2003; Nyberg et al., 2003; Mallow, Bernarding, Luchtmann, Bethmann, & Brechmann, 2015).

On the one hand, some researchers argue that the effectiveness of the Method of Loci is based on engaging the learner with the study material in much the same way as other non-spatial strategies (Chapter 4, Bower, 1970; Caplan et al., 2019; Wang & Thomas, 2000). Bower (1970) noted that numerical peg systems, which use a pre-defined system of images and numbers as a basis to associate to-be-remembered items, can produce equivalent performance to the Method of Loci. Bouffard et al. (2017) found that participants' recall performance showed memory boosts comparable to the Method of Loci for both spatial and sequentially ordered strategies, suggesting that spatial and and sequentially ordered representations can be used to enhance memory performance in a similar way. Caplan et al. (2019) trained naïve participants on virtual environments with particular navigation-based characteristics (conceptual familiarity, boundary, and lines of sight), and found that memory success was not determined by navigation-relevant features of the environment nor individual differences in knowledge of the environments or experience with first-person video games. The experiments described in Chapter 4 showed that a mnemonic scaffold based on the human body is equally effective as the Method of Loci, providing further evidence that the effectiveness of the Method of Loci may not be due to imagined navigation but due to characteristics shared with non-navigational mnemonic scaffolds.

On the other hand, some researchers argue for the primacy of spatial cognition for the effectiveness of mnemonic strategies. Rolls (2017) suggested that the Method of Loci may be effective because it activates the neural navigation system which is also specialized for episodic memory. Theoretical interpretations of studies on spatial navigation and memory suggest that spatial context is a prominent defining feature of episodic memory (Hebscher, Levine, & Gilboa, 2018; Robin, 2018; Robin, Buchsbaum, & Moscovitch, 2018; Robin, Wynn, & Moscovitch, 2016). It has been noted, for example, that spatial cues lead to quicker and more detailed memories (Hebscher et al., 2018; Horner, Bisby, Wang, Bogus, & Burgess, 2016; Robin et al., 2016). In this context, the Method of Loci has been associated with neurons in the hippocampus and neighbouring regions (e.g., place and grid cells; Becchetti, 2010), which are also active during mental navigation (Bellmund, Deuker, Schröder, & Doeller, 2016) and when information is encountered at imagined locations (Moser et al., 2015). Evidence that those place cells are also involved in episodic memory encoding and retrieval (Miller et al., 2013; Monaco, Rao, Roth, & Knierim, 2014) suggests a synergy between navigational processes and episodic memory, which might contribute to the effectiveness of the Method of Loci. However, neurons representing temporal information have also been described within the hippocampus (Eichenbaum, 2014; Fellner et al., 2016; Itskov, Curto, Pastalkova, & Buzsáki, 2011; Mankin et al., 2012, 2012) which challenges the primacy of spatial-based functions to hippocampal processing. These findings challenge the notion that imagined navigation contributes to the effectiveness of the Method of Loci, even though several neuroimaging studies have reported navigation-like brain activity during the application of the Method of Loci (Fellner et al., 2016; Müller et al., 2018; Moser et al., 2015; Kondo et al., 2005; Maguire et al., 2003; Nyberg et al., 2003; Mallow et al., 2015).

If spatial locations explain supremacy of the Method of Loci, then the Autobiographical Story Scaffold might be optimized when the story contains lots of spatial features. Alternatively, if not, then spatial features might be less determining of serial-recall success with the story scaffold. To test this, we asked whether the number of locations mentioned in Autobiographical Story Scaffolds predicts memory success. If there is such a relationship, the instructions to use autobiographical stories as mnemonic scaffolds could be improved by asking the users to include locations in their stories.

### 2.1.3 Scaffold-word associations and verification of strategy use

Instructions to apply a mnemonic strategy do not guarantee that participants use the strategy to study and recall list items (Bellezza, 1981). Previous studies have assumed compliance of participants (e.g., Roediger, 1980; Bouffard et al., 2017). This is problematic because self-reported compliance rates of using the instructed strategy cannot be expected to be particularly high (Sahadevan, Chen, & Caplan, 2021). For example, the self-reported compliance rate in a study of the Method of Loci was only 40 and 58% in the two strategy groups respectively (Legge, Madan, Ng, & Caplan, 2012). Thus, in addition to concerns with the validity of subjective report of instruction compliance, we were concerned that our comparison of the effectiveness of our four mnemonic scaffolds could be confounded by including participants who do not apply the strategy as instructed. We therefore included the scaffold–cued recall task, where participants were tested directly for memory for scaffold–word associations (Sahadevan et al., 2021). This gives us the unique opportunity to check whether participants were actually using the scaffold strategies as instructed, and actually forming scaffold–word associations, and secondly, whether success in scaffold–word memory, itself, might largely explain differences in serial-recall success across the presentation modes and story types. We expect that when participants can recall the study items when cued with sentences they were asked to associate it with, they may rely on the association during serial recall. In other words, we expect scaffold-cued recall to predict serial recall. Alternatively, it is conceivable that participants may not be able to recall the words during serial recall despite being able to recall them when cued with the anchor (Sahadevan et al., 2021). In this case, the story scaffold would only show a benefit for scaffold-cued recall but no benefit for serial recall.

### 2.1.4 Goals and hypotheses of the current study

Our specific goals were to a) describe a novel mnemonic scaffold based on autobiographical stories, b) compare different presentation modes of autobiographical stories and scaffolds (and participants' compliance) to an uninstructed Control, and c) test whether spatial locations mentioned in Autobiographical Story Scaffolds contribute to their effectiveness. Since this is, to our knowledge, the first study that compares different presentation modes of mnemonic scaffolds, we did not have an a-priori hypothesis as to how the presentation modes differ in their effectiveness. We hypothesized that locations mentioned in the autobiographical stories might increase their mnemonic benefit because locations might engage the learner in spatial cognition, which, according to some researchers (e.g., Fellner et al., 2016), has a memory-enhancing effect as described above.

## 2.2 Methods

The experiment was programmed and run in MATLAB with PsychToolBox (Psychophysics Toolbox Version 3) experiment programming extensions (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997), and the CogToolBox Library (Fraundorf et al., 2014).

### 2.2.1 Participants

Participants (N = 115) were recruited from the introductory psychology research participation pool in partial fulfillment of course requirements. Data from two participants were excluded from the analyses because they used animal tales rather than autobiographical stories. All participants were required to have English as their first language and/or have learned English before the age of six and were all older than 17 years old. Written informed consent was obtained prior to the experiment in accordance with a University of Alberta ethical review board.

### 2.2.2 Materials and overview of the procedure

Study lists were random sets of ten 4–8 letter nouns of high and low imagery (e.g., MEADOW, DOUBLE, EFFORT, TIMBER) drawn from the Toronto Word Pool with frequency ratings by Kucera and Francis,[1] also used by Bouffard et al. (2017). Words were drawn at random, without replacement, to construct the complete set of serial lists, each comprising ten words. The experiment was presented in individual closed testing cubicles. First, participants in all five groups typed up a short autobiographical story comprised of ten sentences. Next, participants in the four experimental groups were told that they were going to be trained

[1] `http://memory.psych.upenn.edu/Word_Pools` accessed 12/01/2021

on an effective memory technique that they were to use to study lists of words. After generating an autobiographical story comprised of ten sentences, participants learned lists of ten words by embedding the words into sentences within their story convincing themselves that it actually happened this way. The four experimental groups varied by presentation mode of the words and participants' typed autobiographical reports. As illustrated in Figure 2.1, the Type Group typed the sentences from their story with the study words integrated; the Word-Only Group, saw only words; the Word-plus-Story Group saw one word together with their whole story on the same screen, and the Word-plus-Sentence Group saw one word and one sentence from their autobiographical story on the same screen. The Control Group received no instruction on how to study the words. Due to additional time required for re-typing the sentences with the words integrated, the Type Group learned only seven lists. Experimental groups Word-Only, Word-plus-Story, Word-plus-Sentence and the Control Group learned nine word lists. After the study phase, all participants completed serial recall of the most recently studied word list, asked to type as many words in order of presentation as they could. Immediately following serial recall, participants four experimental groups completed a scaffold-cued recall task, to evaluate the extent to which they had successfully learned scaffold-word associations, as requested by the instructions. Lastly, participants were asked three self-report questions pertaining to how useful they found the strategy and how much they used it. The self-report questions were added partway through data collection, and are thus available for 72 of the 115 participants. The phases of the experiment are described in more detail below.

### 2.2.3 Story generation

Participants were asked to think of a personal event they remember well, like a vacation, or their first day of school, and break the sequence of events into 10 sentences. An example story is shown in Figure 2.2. Participants were instructed to report the story as if they were telling it to a friend, to include details, and to maintain chronological order. Once their stories were completely typed up, they were given the option to edit sentences or alter the order of sentences within their story until they were satisfied. The story phase was only completed once by each participant, and thus the same autobiographical story was used for all word lists. As shown in previous studies on mnemonic scaffolds, proactive interference is not to be expected (Chapter 4, Massen & Vaterrodt-Plünnecke, 2006; Bass & Oswald, 2014).

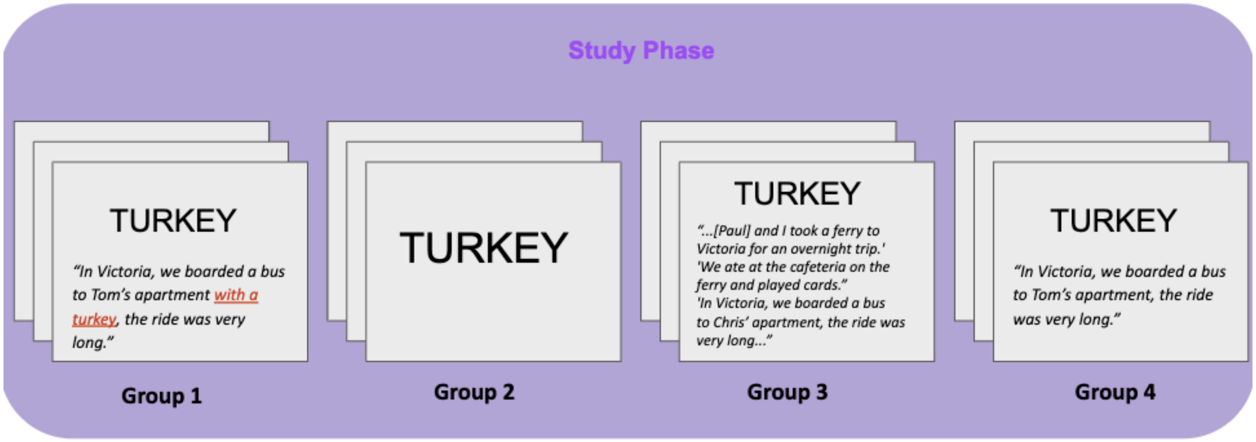



Figure 2.1: Illustration of the four different presentation modes used four experimental groups. Group 1 retyped the words with the study items integrated; Group 2 only saw the words; Group 3 saw one study item together with the whole story; Group 4 saw one study item together with one sentence.

### 2.2.4 Serial recall and scaffold-cued recall

Participants in the experimental groups were instructed to associate one word with each sentence of their story. When they believed they had mentally incorporated one study word into one sentence of their story by imagining that it actually happened this way, they advanced to the next word list and sentence in their story. To recall the words, participants were asked to repeat this mental incorporation, envisioning each word as a part of a sentence in their autobiographical story. The entire task was self-paced, and participants were asked to press ENTER when they were ready to get to the next screen. As shown in Figure 2.1 and explained earlier, the presentation modes of the words and stories varied across the experimental groups. After studying a list, participants were asked to recall them in order of presentation and type them on the keyboard. Response lines were visible in a vertical configuration from the start of recall. Participants were asked to type PASS if they could not recall a word. The study and serial recall and scaffold-cued recall task was repeated until all lists of 10 words each were studied. After the recall phase, participants in all groups with exception of the Control were shown their sentences, one at a time, in a random order, and asked to type the word they associated with the sentences. We used this scaffold-cued recall task because instructions to apply a mnemonic strategy do not guarantee that participants actually use the strategy (Bellezza, 1981). As pointed out earlier, most previous studies have assumed compliance of participants (e.g., Roediger, 1980; Bouffard et al., 2017) even

List words (1st half )
UNIT – DEVICE – TURKEY – COPY – GENIOUS

**Original Story (1st half)**

`[Paul] and I took a ferry to Victoria for an overnight trip.'

'We ate at the cafeteria on the ferry and played cards.'

'In Victoria, we boarded a bus to Toms' apartment, the ride was very long.'

'Afterward we stopped at his apartment to drop off my belongings before heading out.'

'Our first task was to find Brescia Bakery and buy poppy seed cake.'

`Because [Paul] is an absolute **UNIT**, he invited me to stay at this apartment in Victoria; we boarded a ferry there.'

`We ate at the cafeteria on the ferry and played cares, no **DEVICES** around, just people living in the moment.'

`In Victoria we boarded a bus to Toms' apartment, there was a **TURKEY** on the bus, which was strange and unexpected but that's just how it is sometimes.'

'Afterward we stopped at his apartment to drop off my belongings and because he had to **COPY** some documents before heading out again.'

`Our first task in Victoria was to find Brescia Bakery because their poppy seed cake is **GENIOUS** and I wanted some.'

Figure 2.2: Anonymized autobiographical story of a participant the Type Group with and without the study words integrated

though self-reported compliance rates of using the instructed strategy cannot be expected to be particularly high (Sahadevan et al., 2021).

### 2.2.5 Scoring of locations mentioned in the stories

The autobiographical stories typed by each participant were later scored by two researchers to count the number of locations according to agreed criteria regarding what constitutes a location. If there was any disagreement between the number of locations counted by each scorer, the average of both scores of locations was taken. A spatial reference mentioned autobiographical stories was counted as a location if a) it referred to a place where an action was performed or observed, b) it was the destination of an action, c) it was part of a journey. Various places within a larger area (e.g., cities within a country, or countries within a continent, or stores within a mall) were counted separately including the name of the area (e.g., *Europe, West Edmonton Mall*). A spatial reference was not counted as a location if a) it was used as an attribute to a person or object (e.g., *a friend from Austria*), b) the location was used as a time reference, (e.g., *when I was in high school*), c) it was a mode of transportation (e.g., *car, on the plain*). A location was only counted once if it was mentioned multiple times same sentence. An example of the first half of an anonymized autobiographical story of a participant in the Type Group with and without the study words integrated is shown in Figure 2.2.

### 2.2.6 Data analyses

Serial recall accuracy was scored in two ways: a) strict scoring, in which a word was scored as correct if it was recalled position it was presented, sensitive to order-errors, and b) lenient scoring, in which a word was scored as correct if it came from the current list, regardless of order. Statistical analyses were conducted in JASP (JASP Team, 2019) using simple linear regressions or analyses of variance (ANOVA) to test main and interaction effects of categorical variables. In the analyses, we call the main factor "Group" when the Control is included, and "Presentation Mode" when the four presentation modes are compared. The Greenhouse-Geisser correction was applied where sphericity was violated. When conducting post-hoc tests on significant group effects, post-hoc Tukey's Honestly Significant Difference tests were used. We also conducted Bayesian linear regressions, which produce a Bayes Factor for the linear regression because we needed to determine if the null effect Control Group is stable. The advantage is that Bayesian model comparison techniques provide support for one model over another in contrast to classical hypothesis testing, which looks for evidence against only one model (the null hypothesis). The Bayes Factor is the ratio of Bayesian probabilities for the alternative and null hypotheses; $BF = p(H1)/p(H0)$. For smaller values of this ratio, the null is more strongly supported by the data under consideration than the alternate hypothesis, whereas a greater value of this ratio would indicate otherwise. By convention (Raftery & Kass, 1995) there is "some" evidence for the null when $BF = < 0.3$, and correspondingly, "some" evidence for the alternate hypothesis when $BF > 3$. "strong" evidence is inferred when $BF < 0.1$ or $> 10$.

## 2.3 Results

To answer our first question (Which presentation mode of Autobiographical Story Scaffolds is most effective in facilitating recall?), we analyzed recall data from the scaffold-cued and serial recall tasks for strict and lenient scoring and asked whether self-reported compliance varies by presentation mode. To answer our second question (Do locations mentioned in Autobiographical Story Scaffolds contribute to their effectiveness in facilitating serial recall?) we conducted frequentist and Bayesian linear regressions on the number of locations mentioned in the autobiographical stories and recall accuracy for strict and lenient scoring. We provide detailed analyses below.

### 2.3.1 Effect of the presentation modes on scaffold-cued recall

Before asking how the presentation modes compare in facilitating serial recall accuracy, we were interested in whether participants may have relied on scaffold-cued recall when recalling the words in serial order. Therefore, we conducted linear regressions to predict serial recall accuracy (strict scoring) based on scaffold-cued recall accuracy. Scaffold-cued recall did indeed predict serial recall accuracy in the Type ($p < 0.001$, $\beta = 0.90$, $R^2 = 0.80$, $BF_{inclusion} > 100$), Word-plus-Sentence ($p < 0.001$, $\beta = 0.90$, $R^2 = 0.80$, $BF_{inclusion} > 100$), and Word-only Group ($p < 0.001$, $\beta = 0.80$, $R^2 = 0.63$, $BF_{inclusion} > 100$). The strong relationship between scaffold-cued recall and serial recall in these groups suggests that scaffold-cued recall is a good indicator of strategy use. In the Word-plus-Story Group, the relationship was non-significant ($p < 0.05$), suggesting that in this group, participants struggled to use the instructed strategy.

Next, we were interested in whether the four presentation modes differed success with which participants formed word-sentence associations as instructed. To test this, we analyzed the scaffold-cued recall data on their own in a 4 (Type, Word-Only, Word-plus-Story, Word-plus-Sentence) × 10 (Serial Position 1-10) mixed ANOVA on the average scaffold-cued recall accuracy for each serial position (Figure 2.3). This produced a significant main effect of Presentation Mode ($F(3, 101) = 5.13$, $p = 0.002$, $\eta_p^2 = 0.13$, $BF_{inclusion} > 100$). Tukey's post-hoc tests revealed that scaffold-cued recall accuracy was significantly higher Type Group than Word-Only Group (Mean Difference=0.22, $SE = 0.06$, $d = 0.34$, $p = 0.004$) and Word-plus-Story Group (Mean Difference=0.17, $SE = 0.06$, $d = 0.26$, $p = 0.042$). Participants in the Word-plus-Sentence Group formed almost significantly more word-sentence associations than participants in the Word-Only Group (Mean Difference=0.17, $SE = 0.06$, $d = 0.25$, $p = 0.054$), indicating that participants in the Type Group and the Word-plus-Sentence Group were more successful in associating study items with sentences as instructed than the other two groups.

The main effect was qualified by a significant interaction of Presentation Mode × Serial Position ($F(22.05, 742.46) = 3.30$, $p < 0.001$, $\eta_p^2 = 0.09$, $BF_{inclusion} > 100$). To follow up, we conducted one-way ANOVAs on Presentation Mode at each Serial Position. This revealed that Type > Word-Only ($p < 0.05$) in positions 5 and 6; Type > Word-Only ($p = 0.001$), Type > Word-plus-Story ($p < 0.05$), Type > Word-Only ($p < 0.05$), and Word-plus-Sentence > Word-Only ($p < 0.05$) in position 7; Type > Word-Only ($p < 0.05$) in position 8; Type > Word-Only ($p < 0.001$), Type > Word-plus-Story ($p < 0.05$), and Word-plus-Sentence > Word-Only ($p < 0.05$) in position 9; Type > Word-Only ($p < 0.001$), Type > Word-plus-Story ($p < 0.001$), and Word-plus-Sentence > Word-Only ($p < 0.05$) in position 10. This indicates that the significant interaction of Presentation Mode × Serial Position for scaffold-cued recall accuracy is modulated by the advantage of the Type Group and over both the Word-Only Group and the Word-plus-Story Group, and the advantage of the Word-plus-Sentence Group over the Word-Only Group in middle and recency positions of the word list (Figure 2.3). Taken together, the scaffold-cued recall data show that participants in the Type Group and Word-plus-Sentence Group were more successful in forming word-sentence associations than participants other groups.

Next, we were interested in whether serial recall may have relied upon scaffold-word associations. Therefore, we collapsed across serial position and included scaffold-cued recall accuracy as a covariate in a one-way ANCOVA to determine significant effects of Presentation Mode on serial recall accuracy (strict scoring) controlling for scaffold-cued recall accuracy. If scaffold-word associations are the driving force behind serial recall using the scaffolds, the addition of scaffold-cued recall as a covariate would be expected to render the main effect of Presentation Mode non-significant. The effect of the covariate scaffold-cued recall accuracy itself was significant, strongly supported by the Bayes Factor ($F(1, 100) = 132.43$, $p < 0.001$, $\eta_p^2 = 0.57$, $BF_{inclusion} > 100$). The main effect of Group after controlling for scaffold-cued recall accuracy was non-significant, with the Bayes Factor providing some evidence for a null-effect ($F(3, 100) = 1.90$, $p = 0.13$, $\eta_p^2 = 0.05$, $BF_{inclusion} = 0.30$). In light of the strong relationship between serial recall accuracy and scaffold-cued recall accuracy this suggests that scaffold-word associations may, indeed, have been relied upon the scaffolds during serial recall, itself.

### 2.3.2 Self-reported compliance

We were interested in whether self-reported compliance varied by Presentation Mode, and whether this corroborates our findings from the scaffold-cued recall data that the Type Group and Word-plus-Sentence have the highest compliance rates.

The groups differed significantly in self-reported compliance ($\chi^2(9) = 17.63$, $p < 0.05$; Table 2.1 and Figure 2.4). This appeared largely due to Word-plus-Sentence Group responses being higher (more compliant) than the other groups; when Word-plus-Sentence Group was removed, the chi-square became non-significant ($\chi^2(6) = 9.72$, $p = 0.14$).

To assess the validity of self-reported compliance on serial recall accuracy, we plotted scaffold-cued recall as a function of serial position (Figure 2.4). A mixed ANOVA of self-reported compliance (Always, Mostly, Sometimes, Never) on scaffold-cued recall accuracy for each serial position (1-10) revealed a significant main effect of self-reported compliance ($F(3, 55) = 11.30$, $p < 0.001$, $\eta_p^2 = 0.38$, $BF_{inclusion} > 100$). The interaction of self-reported compliance × Serial Position was non-significant. Tukey's post-hoc tests indicated that participants who answered “always” had significantly ($p < 0.001$) higher scaffold-cued recall accuracy than those who answered “sometimes” or “never” and participants who answered “mostly” had significantly ($p < 0.05$) higher scaffold-cued recall accuracy than participants who answered “sometimes” and “never”. “Sometimes” and “never” scaffold-cued recall ac-

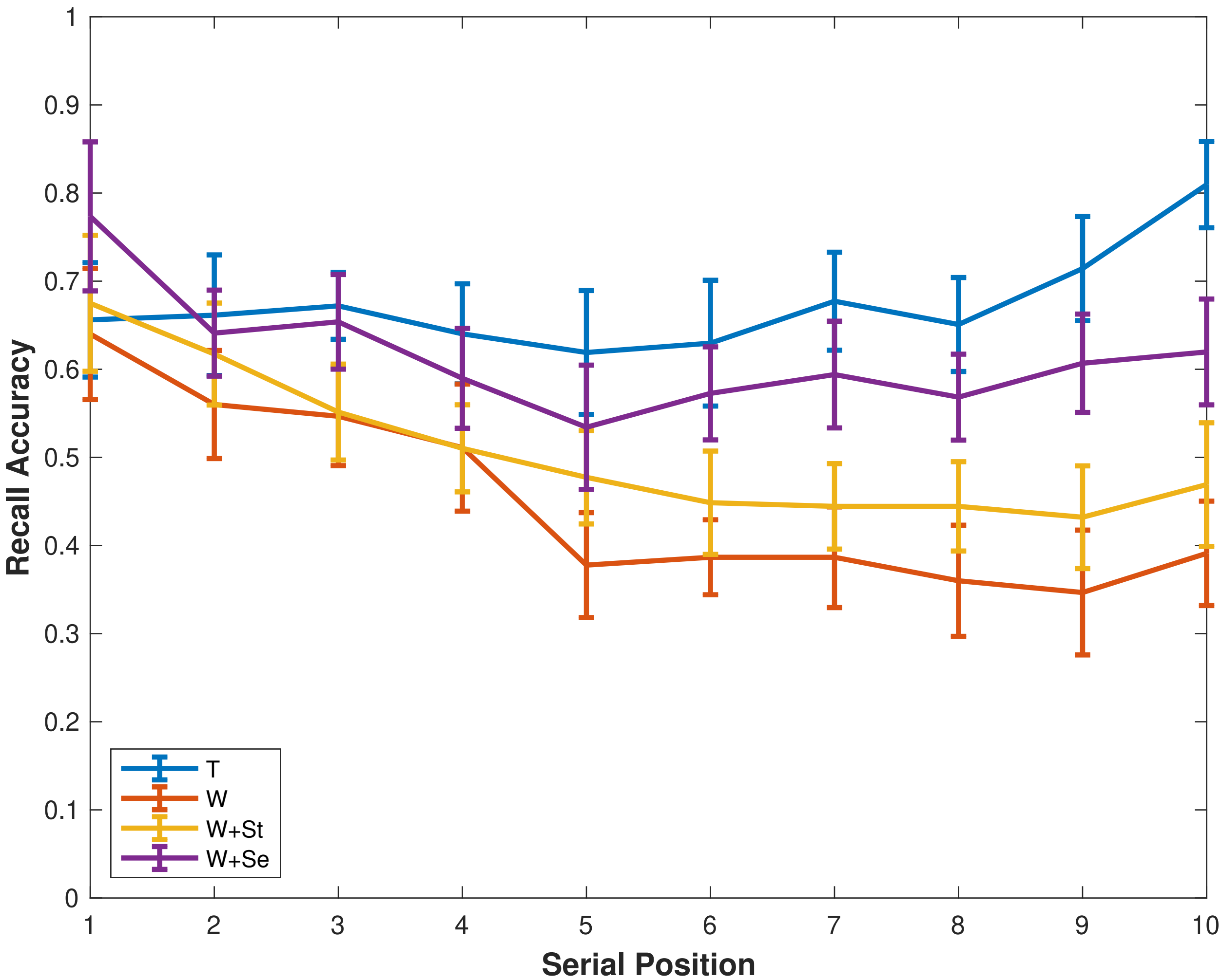



Figure 2.3: Serial position curves of the scaffold-cued recall task. T stands for Type Group, which retyped the sentences from from the autobiographical stories with the words integrated, W stands for Word-only Group, which only saw the list words during study, W+St stands for Word-plus-Story Group, which saw the whole story and one list word at a time, W+Se stands for Word-plus-Sentence Group which saw one word and one sentence at a time, and C stands for Control. Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994).

| Group | Always | Mostly | Sometimes | Never | Total |
|---|---|---|---|---|---|
| 1 | 0 | 8 | 4 | 1 | 13 |
| 2 | 0 | 5 | 8 | 3 | 16 |
| 3 | 2 | 5 | 8 | 0 | 15 |
| 4 | 4 | 8 | 3 | 0 | 15 |

Table 2.1: Rates of self-reported compliance with the instructed strategy, as a function of group. Values correspond to numbers of participants responding to the question "Did you integrate the words into your story?" with "Never," "Sometimes," "Mostly," or "Always."

curacies did not differ from one another ($p = 0.662$), and neither did "always" from "mostly," ($p = 0.117$). This indicates that self-reported compliance corresponds with objectively measured compliance. To test whether the effect of self-reported compliance observed in scaffold-cued recall accuracy translates to serial recall accuracy under the strict scoring criterion, we turned to the serial recall data with an ANOVA with the same design, as shown in Figure 2.5. This revealed a significant main effect of self-reported compliance ($F(3, 55) = 6.06$, $p = 0.001$, $\eta_p^2 = 0.25$, $BF_{inclusion} > 100$). The interaction of self-reported compliance × Serial Position was significant ($F(14.94, 273.91) = 2.97$, $p < 0.001$, $\eta_p^2 = 0.14$, $BF_{inclusion} > 100$). Tukey's post-hoc tests revealed that participants who answered "always" had significantly higher serial-recall accuracy than participants who answered "sometimes" ($p = 0.005$) and participants who answered "never" ($p = 0.012$). Participants who answered "mostly" recalled the words almost significantly better than participants who answered "sometimes"($p = 0.054$). Following up on the significant interaction of Self-reported Compliance × Serial Position with one-way ANOVAs at each Serial Position, we found that "always"> "sometimes ($p = 0.063$) at position 3, "always"> "sometimes ($p < 0.05$) and "always"> "never"($p = 0.064$) at position 6, "always"> "sometimes" ($p < 0.05$) and "always"> "sometimes" ($p < 0.05$) at position 7, "always"> "sometimes" ($p < 0.05$) and "always"> "never" ($p < 0.05$), and "mostly"> "sometimes" ($p = 0.060$) at position 8, "always"> "sometimes" ($p < 0.05$) and "always"> "sometimes" ($p < 0.05$), and "mostly"> "never" ($p < 0.05$) at position 9, and "always"> "sometimes" ($p < 0.05$) and "always"> "never" ($p < 0.001$), and "mostly"> "sometimes" ($p < 0.001$) and "mostly"> "never" ($p < 0.001$) at position 10. This suggests that participants with high self-reported compliance rates have not only higher objectively measured compliance but also higher serial recall for the whole list, and particularly for middle and recency positions.

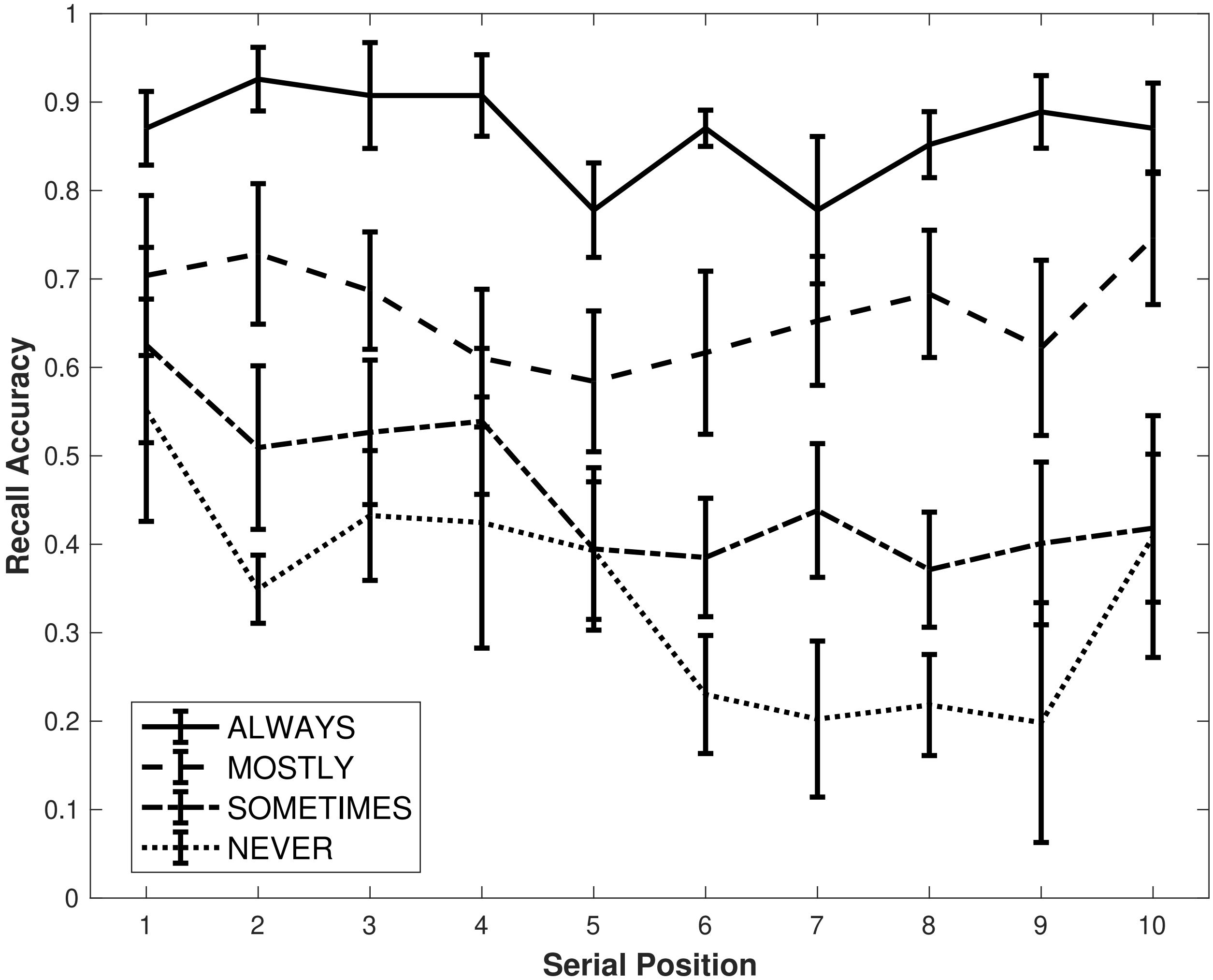



Figure 2.4: Scaffold-cued recall as a function of serial position, separated by participants' answers to the question "Did you envision the words of part of your story when studying them?"

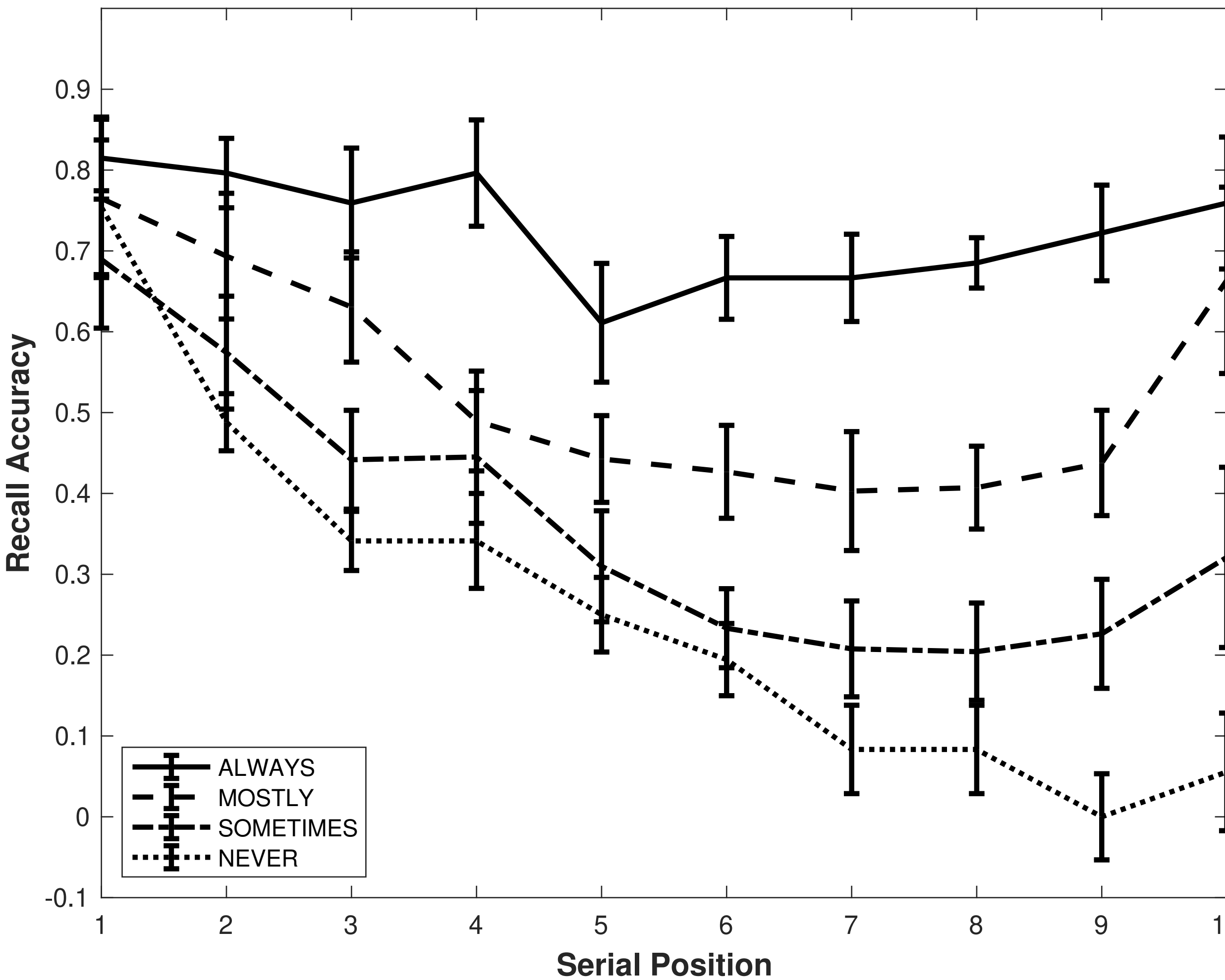



Figure 2.5: Serial recall accuracy (strict scoring) as a function of serial position, separated by participants' answers to the question "Did you envision the words of part of your story when studying them?"

### 2.3.3 Effect of the presentation modes on serial recall

The scaffold-cued recall responses assess how successfully participants complied with the instructions to associate the study words with sentences from their autobiographical stories. This does not tell us the degree to which participants were able to use these word-sentence associations to perform serial recall of the list. In addition to testing whether participants in the experimental groups recalled more words than uninstructed Control participants (Figures 2.6 and 2.7), we asked whether the advantage of retyping the story with the words integrated (Type Group) and viewing one word together with one sentence at a time (Story-plus-Sentence Group) over the other two presentation modes observed scaffold-cued recall task is also observable in serial recall using a strict (Figure 2.6) and lenient (Figure 2.7) scoring criterion.

#### Strict serial recall accuracy

Comparing the effect of Group (Type, Word-Only, Word-plus-Story, Word-plus-Sentence, Control) on serial recall accuracy using the strict scoring criterion (Figure 2.6), a mixed ANOVA revealed a non-significant effect of Group with a contradicting Bayes Factor strongly supporting an effect of Group ($F(5, 128) = 1.32$, $p = 0.267$, $\eta_p^2 = 0.04$, $BF_{inclusion} > 100$), and a significant interaction of Group × Serial Position ($F(21.99, 703.64) = 7.67$, $p < 0.001$, $\eta_p^2 = 0.19$, $BF_{inclusion} > 100$).

Following up on the contradicting Bayes Factor and the significant interaction with one-way ANOVAs at each serial position, Tukey's post-hoc tests revealed that Control > Type ($p < 0.05$) in position 1. No significant effects were found from positions 2 to 8. In position 9, Tukey's post-hoc tests revealed that Type > Word-Only and Type > Control ($p < 0.05$). In position 10, Type > Word-Only, Type > Word-plus-Story, Type > Control, Word-plus-Sentence > Word-Only, Word-plus-Sentence > Control ($p < 0.05$), and Word-plus-Sentence > Word-plus-Story approached significance ($p = 0.57$). This indicates that the significant interaction of Group × Serial Position is mainly characterized by the advantage over the Type Group and the Word-plus-Sentence Group over the other presentation modes and the Control in recency positions. In light of this interaction, the Bayes Factor contradicting the non-significant $p$-value of the main effect of Group does not necessarily mean that there is no null effect because the JASP algorithm requires the significant interaction to be part of

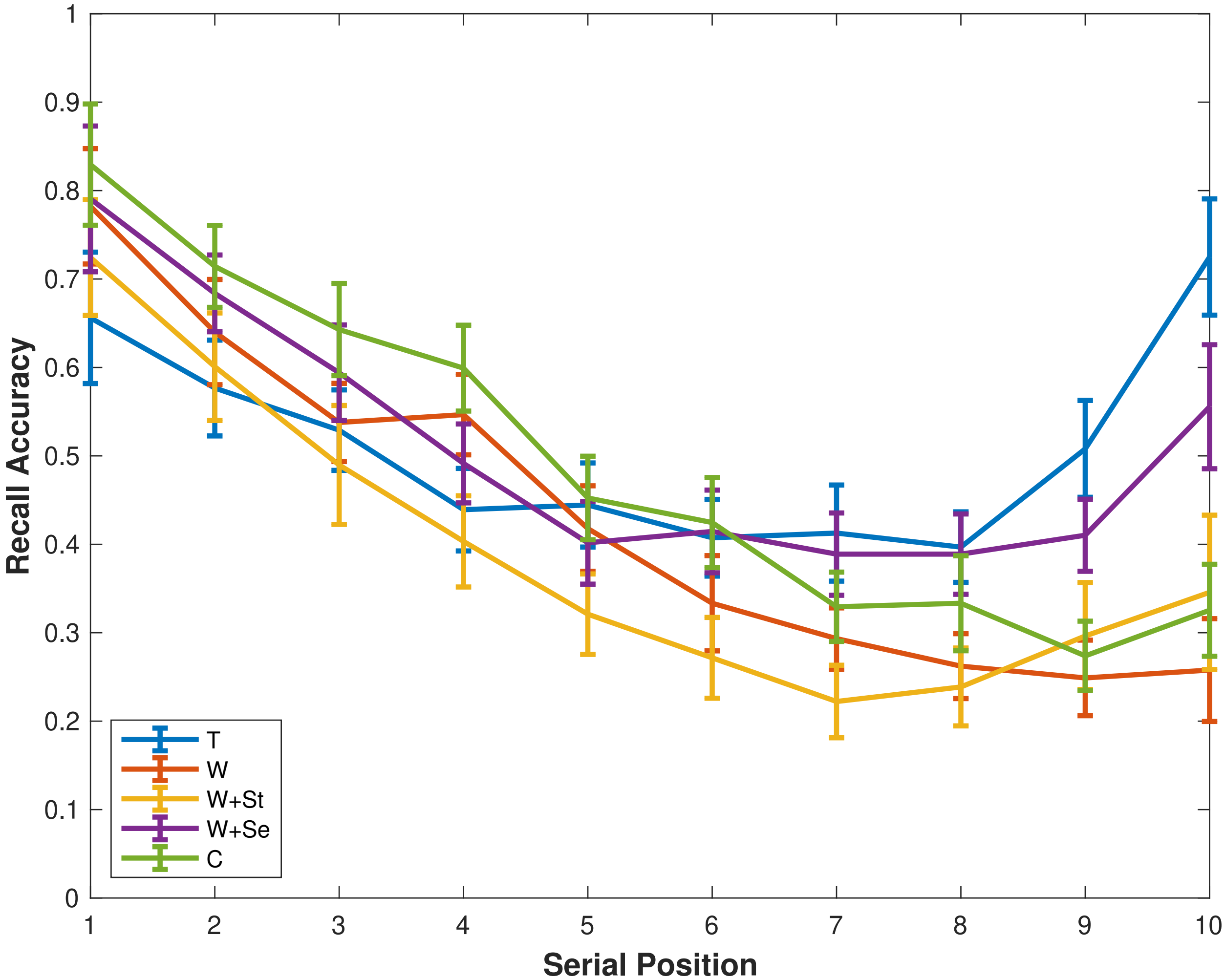



Figure 2.6: Serial position curves of serial recall (strict scoring). Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994).

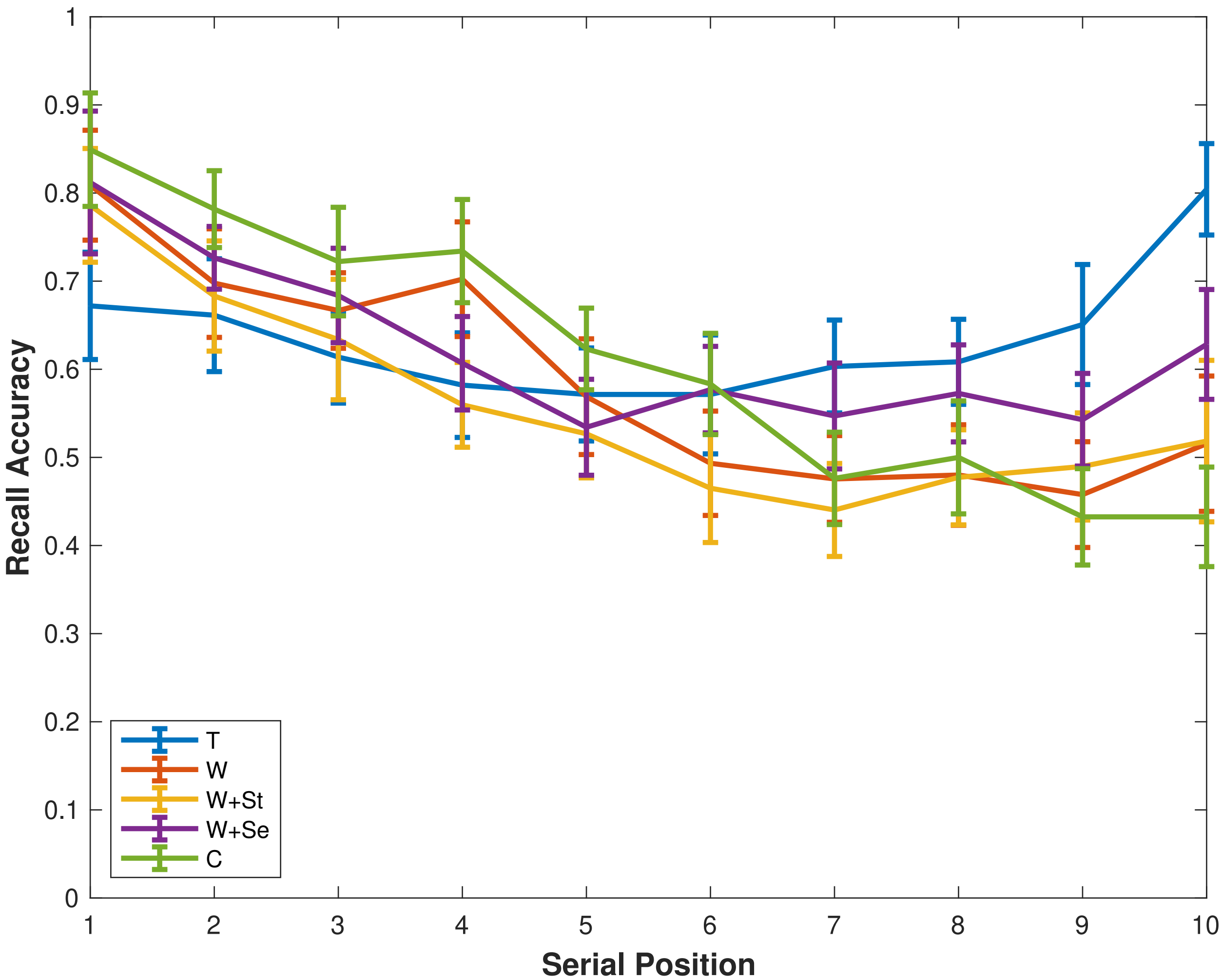



Figure 2.7: Serial position curves of serial recall (lenient scoring). T stands for Type Group, which retyped the sentences from from the autobiographical stories with the words integrated, W stands for Word-only Group, which only saw the list words during study, W+St stands for Word-plus-Story Group, which saw the whole story and one list word at a time, W+Se stands for Word-plus-Sentence Group which saw one word and one sentence at a time, and C stands for Control. Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994).

the Bayesian mixed ANOVA.

Together, this suggests that the instructions of the Type Group and the Word-plus-Sentence Group have a memory-enhancing effect in recency positions of the list when scored under the strict criterion.

### Lenient serial recall accuracy

We were interested in whether lenient scoring, which reflects memory for items regardless of their order (Figures 2.7) yields different results than strict scoring. Comparing the effect of Group (Type, Word-Only, Word-plus-Story, Word-plus-Sentence, Control) on serial recall accuracy (lenient scoring (Figures 2.7), a mixed ANOVA revealed a non-significant effect of Group with a contradicting Bayes Factor ($F(2, 128) = 0.60$, $p = 0.666$, $\eta_p^2 = 0.02$, $BF_{inclusion} > 100$) and a significant interaction of Group × Serial Position ($F(26.98, 863.23) = 5.12$, $p < 0.001$, $\eta_p^2 = 0.14$, $BF_{inclusion} > 100$). Following up on the contradicting Bayes Factor and the significant interaction with one-way ANOVAs at each serial position Tukey's post-hoc tests revealed that Control > Type ($p < 0.05$) in position 1 and Type > Control in position 9 and Type > Word-Only, Type > Word-plus-Story, Type > Control, and Word-plus-Sentence > Control in position 10, indicating a similar pattern as with strict scoring, in which participants in the Type Group and Word-plus-Sentence Group outperformed the other groups in recency positions.

### Spatial Locations and Recall Accuracy

To investigate the relationship between recall accuracy and number of locations used autobiographical stories, simple linear regressions were calculated for all groups. The dependent variable was recall accuracy, and the explanatory variable was number of locations mentioned autobiographical stories. First, we conducted a linear regression for all participants four experimental groups as shown in Figure 2.8a. The regression found a significant relationship between number of locations and recall accuracy ($p = 0.015$, $\beta = 0.25$, $R^2 = 0.063$, $BF_{10} = 0.038$), showing that a higher number of spatial locations was related with higher recall accuracy.

To rule out the possibility that participants who mention many spatial locations in their autobiographical stories had higher recall accuracy, regardless of whether they use the autobiographical mnemonic strategy or not, we conducted a linear regression for the Control

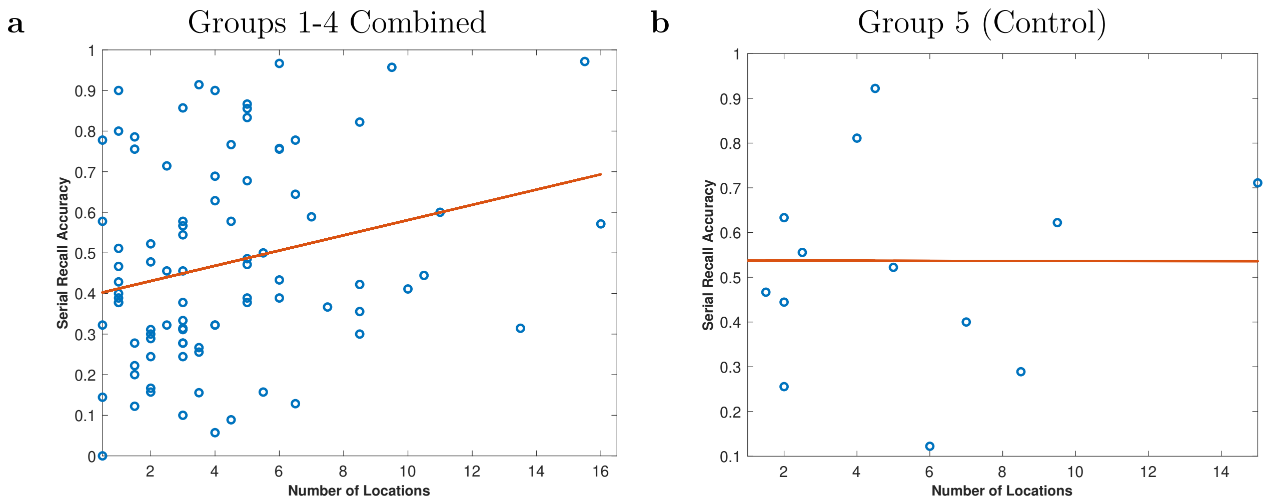



Figure 2.8: Linear regressions (strict scoring) of recall accuracy against number of locations mentioned autobiographical stories for all participants in the four experimental groups and for participants in the Control. Each point plots one participant, and the red line plots the best-fitting linear function.

Group, shown in Figure 2.8b. In contrast to the experimental groups, a supported null effect was found Control Group ($p = 0.95$, $BF_{10} = 0.29$).

To investigate the relationship between number of locations with experimental groups, we conducted linear regressions for each experimental group separately. The regression was only significant in the Word-plus-Sentence Group, with the Bayes Factor providing strong evidence for an effect ($F(1, 21) = 11.77, p = 0.003$, $\beta = 0.60$ $R^2 = 0.33$, $BF = 14.72$) as shown in Figure 2.9d.

For the remaining experimental groups, we found no significant relationship ($p > 0.05$) between recall accuracy and number of locations with Bayes Factor either providing evidence for the null or inconclusive range (Figure 2.9a-c). Taken together, this indicates that the number of locations mentioned autobiographical stories predicted serial recall accuracy when study words and sentences from the autobiographical story were shown on the same screen.

## 2.4 Discussion

The main goal of this study was to test whether autobiographical stories can serve as effective mnemonic scaffolds and to determine the most effective presentation mode. In order for a mnemonic scaffold to be effective, participants must be able to use it to form scaffold-word associations as instructed and receive a memory benefit compared to not using the

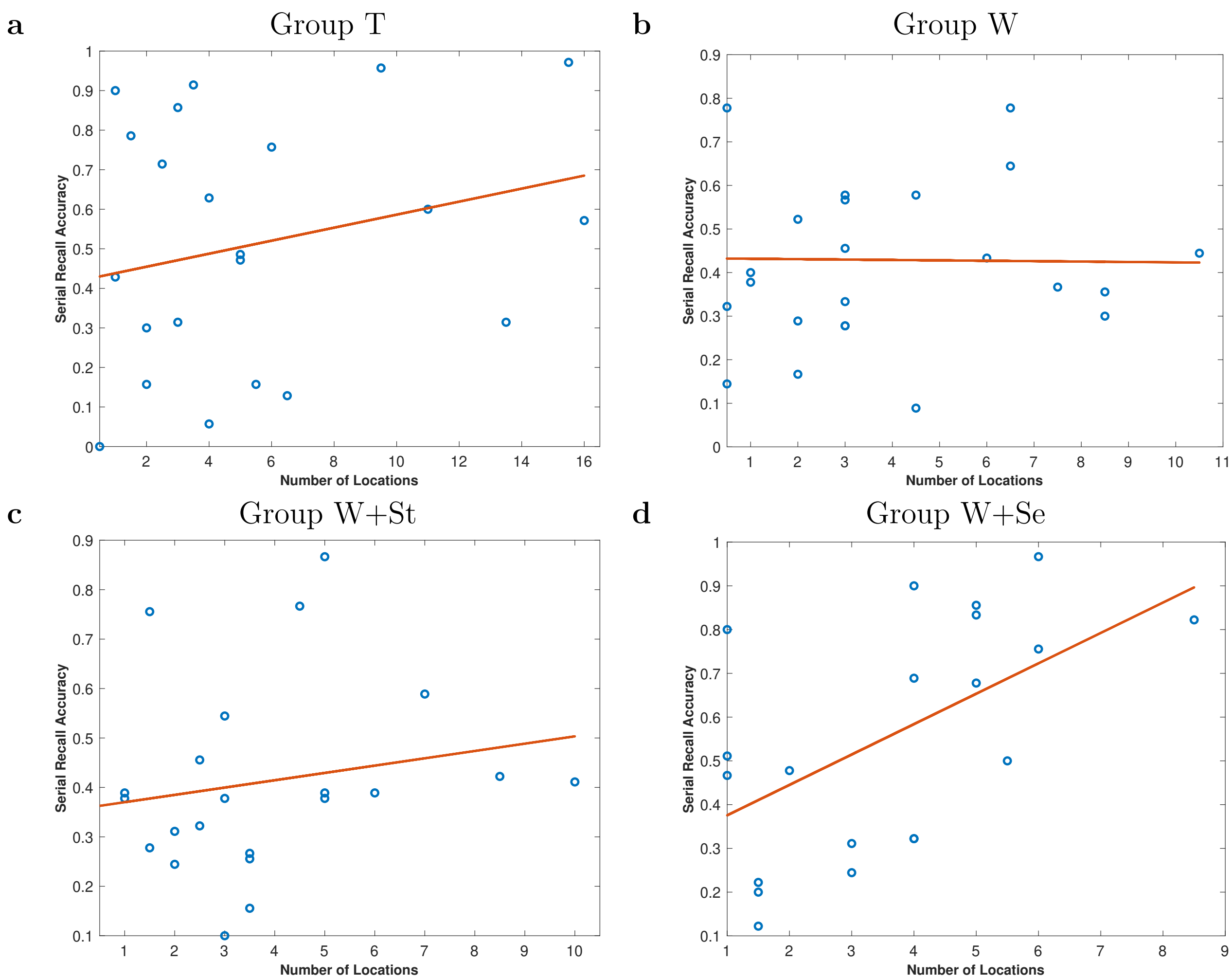



Figure 2.9: Linear Regression of recall accuracy (strict scoring) and number of locations mentioned autobiographical story experimental groups (a-d, respectively).

scaffold. Our results show that autobiographical stories can serve as effective mnemonic scaffolds, i.e. participants can use them to form scaffold-word associations and benefit from those during serial recall for words in recency positions if certain conditions are met. Some participants followed the instructions as verified by the scaffold-cued recall data. This is important because previous studies have assumed compliance of participants (e.g., Roediger, 1980; Bouffard et al., 2017), which is problematic because self-reported compliance rates of using the instructed strategy cannot be expected to be particularly high (Sahadevan et al., 2021). Participants who either typed the sentences with the words integrated or viewed one word and one sentence at a time had relatively high self-reported and objectively verified compliance rates and outperformed control in recency positions. Since viewing words and sentences is more time-efficient than retyping the sentences, this was determined as the preferred presentation mode. We also found some evidence that spatial locations mentioned in Autobiographical Story Scaffolds contribute to their effectiveness when using this presentation mode. Below, we discuss the role of different presentation modes and of spatial locations in Autobiographical Story Scaffolds in detail.

### 2.4.1 Autobiographical Story Scaffolds and procedural variants

Although the four presentation modes were not different from Control overall, our more fine-grained analyses lead to the observation that viewing one study item and one autobiographical sentence at a time is the most effective presentation mode for Autobiographical Story Scaffolds. First, this is because self-reported compliance rates were the highest for this condition, and participants in the Word-plus-Sentence and Type Groups were more successful in forming word-sentence associations than other two groups. Importantly, recall accuracy in the Type Group might confound the effectiveness of the strategy with a production effect of typing the words. This also slows down study time, which in turn might increase memory independently of the strategy. Whether or not the production effect enhances memory or not is not clear and dependent on various task conditions (for an extensive review, see MacLeod & Bodner, 2017). Second, the significant interaction between group and serial position is important. While serial-recall accuracy of the Control Group and the Word-Only and Word-plus-Story Groups dropped over the course of the list and does not show a recency effect, recall accuracy of the Type and Word-plus-Sentence Groups only dropped first half of the list. We speculated that the rise of the advantage toward later serial positions hints that these presentation modes might prove even more superior as lists lengthen (Chapter 3).

Regarding the lack of a mnemonic benefit in the Word-only Group that studied the words without viewing their story, it is possible that those participants found it difficult to remember the sentences from their autobiographical story. Recalling the autobiographical story while trying to encode the study items may introduce a high working memory load, increasing the difficulty of the task as compared to the other presentation modes. This likely also explains why scores from the Word-plus-Story Group scaffold-cued recall task were the lowest.

The relatively low strategy compliance in the Word-plus-Story Group might be due to visual overload and distraction caused by seeing the full story and one study item on the screen. In addition, it is important to note that participants received the instructions to use their autobiographical story as a memory scaffold only after they typed them. More training with the strategy or simply informing participants of the purpose of their stories might have an effect on both strategy compliance and effectiveness.

Overall, the presentation mode of the Word-plus-Sentence Group emerges as the preferred way to instruct the strategy for multiple reasons: a) participants' compliance is high, b) there is a mnemoinc benefit in recency positions, showing potential for successful encoding of longer lists, c) success of the strategy is not confounded by a production effect of typing the words and d) the number of spatial locations used in the Autobiographical Story Scaffolds correlates positively with recall accuracy, suggesting that their when using this presentation mode could be greatly improved by explicitly instructing the learners to include spatial information in their mnemonic scaffolds, as discussed further below.

### 2.4.2 The role of spatial information in successful application of the autobiographical scaffold

We asked whether mentioning spatial locations in Autobiographical Story Scaffolds increases their effectiveness. We found a significant correlation between serial recall accuracy and the number of spatial locations used in the Autobiographical Story Scaffolds. This held for the experimental groups altogether and was most prominent in the presentation mode that we identified to be the most successful one (Word-plus-Sentence Group). The notion that spatial cognition plays a role in the effectiveness of an autobiographical scaffolding approach is further supported by the null result in the Control Group. This shows that participants who mention more locations in their autobiographical stories they were asked to type without any mention of using these stories to study the words, did not recall more words as shown in Figure 2.8b. Thus, the higher recall success of participants in the Word-plus-Sentence Group can be attributed to the higher number of spatial locations mentioned in those stories. However, as shown in Figure 2.9a there were still some participants in the Type Group who had high recall accuracy despite incorporating very few spatial locations in their story. While the role of spatial information for memory success is still unclear and research from various perspectives on memory is ongoing, generalizing the role of spatial cognition for the success of mnemonic strategies is very likely the wrong approach. Instead, individual differences in learner aptitude might be the key to understanding the success of mnemonic scaffolding strategies involving spatial thinking. Sanchez (2019) found that individual differences in visuospatial ability determined whether the Method of Loci was helpful or, in fact, harmful to serial recall. From our results of this study, in addition to identifying the most effective presentation mode for the autobiographical strategy, we inferred a way to make the strategy more effective by instructing participants to include spatial locations in their autobiographical scaffolds. The existence of participants who excelled with the strategy but without relying on spatial locations suggests that spatially heavy stories may not be the only way that the story-based strategy could succeed.

### 2.4.3 Individual scaffold-word associations may drive serial recall accuracy

We found that scaffold-cued recall is a strong predictor of serial recall accuracy in all experimental groups except for the Word-plus-Story Group. As pointed out above, the lack of a mnemonic benefit with this presentation mode might be due to visual overload and distraction caused by seeing the full story and one study item on the screen. The strong positive relationship between scaffold-cued recall and serial recall suggests that participants have relied upon word-sentence associations when recalling lists of ten words. Given our findings from the scaffold-cued, and serial recall data, we speculate that scaffold-word associations may be the driving force behind serial recall using the scaffolds. This is important considering our alternative hypothesis that if participants would only benefit from the scaffolds in scaffold-cued recall but not in serial recall. Failing to apply the scaffold-word associations would mean that there would only be very limited real-world usefulness of the story scaffolds.

This is because learners would not be able to recall study items without being shown their personal stories, which in most learning and testing scenarios is not feasible. Consequently, our finding that participants who apply the story scaffolds as instructed do see a benefit from scaffold-word associations during serial recall demonstrates that this novel mnemonic technique can be used in real-world learning situations.

## 2.5 Conclusion

In sum, we found that the Autobiographical Story Scaffold is effective for recency positions. It is most effectively applied when one word and one sentence from the autobiographical story are presented together during study phase. We found a positive relationship between number of locations and recall accuracy with this presentation mode. Moreover, participants who followed the instructions to associate study items with sentences from their autobiographical stories were able to use those associations for serial recall.

These findings offer important potential for the strategy to be improved upon: instructing learners to incorporate spatial locations in their autobiographical scaffolds to improve the effectiveness of the strategy. In future real-world and experimental applications of the Autobiographical Story Scaffold, we suggest that there be more participant training on how to use the strategy which includes the explicit recommendation to use spatial locations if desired autobiographical stories. To further develop the autobiographical mnemonic strategy to facilitate best possible learning outcomes for individual learners, future studies should incorporate ways to measure differential learner aptitudes and give the learners the opportunity to become acquainted with story-based mnemonic scaffolds in a multi-session training program. To test whether the the effectiveness of Autobiographical Story Scaffolds is based on the stories being autobiographical, future experiments should directly compare autobiographical to non-autobiographical scaffolds as we will attempt in Chapter 3.

# Chapter 3

# Story scaffolds: self-relevance and boundary conditions

# Abstract

Chapter 2 showed that autobiographical stories can be used as effective mnemonic scaffolds. They offer a recall advantage in lists of ten words. Spatial locations mentioned in Autobiographical Mnemonic Scaffolds can increase their effectiveness in facilitating serial recall.

Here we a) test whether this effect can be replicated with lists of fifteen words and b) compare autobiographical scaffolds to fictional ones to test whether self-reference has an effect. We neither found an advantage of neither scaffold type over an initial assessment of serial-recall accuracy prior to using mnemonic scaffolds nor Control, suggesting that list length may be a boundary condition for story-based mnemonic scaffolds. More-fine grained analyses revealed that participants in both experimental groups who formed scaffold-word associations as instructed benefited from their scaffolds. This suggests that participants who associate words and individual sentences as instructed can use those for serial recall.

## 3.1 Introduction

As shown in Chapter 2 (see also Chapter 4), mnemonic scaffolds based on autobiographical stories can enhance memory for lists of ten words in recency positions when participants retype their stories with the words integrated and when being shown the sentences from their autobiographical stories together with the study items. As the benefit of the autobiographical scaffolds was present in middle and recency positions, we speculated that this effect would be more pronounced in longer word lists and stories, respectively. To test this, we increased list length to fifteen.

One might expect that the mnemonic benefit of Autobiographical Story Scaffolds is due to the stories being autobiographical as previous experiments did not include a comparison with non-autobiographical stories. To test whether the effectiveness of the Autobiographical Story Scaffolds is due to autobiographical content of the stories, we directly compare autobiographical scaffolds to fictional ones and to a control condition using the most effective presentation mode determined in Chapter 2. We also tested whether the mnemonic benefit of locations mentioned in autobiographical stories described in Chapter 2 also applies to fictional stories. These insights would not only allow us to refine the instructions of the strategy for better memory outcomes in novice learners, but also provide more information on the disputed role of spatial cognition in the effectiveness of the Method of Loci (Chapter 2). Finally, we tested whether recall accuracy of the stories themselves predicts recall accuracy of the study items. Again, insights from this would allow us to refine the instructions of the strategy for better memory outcomes in novice learners and provide information on whether memory for the scaffolds themselves is important for participants to use the strategy effectively or whether alternative factors may play a role.

### 3.1.1 The self-reference effect in story-based mnemonic scaffolds

As mentioned in Chapter 1, research and theory on the self-reference effect suggest that Autobiographical Mnemonic Scaffolds, in theory, may outperform fictional ones (e.g., Bower & Gilligan, 1979; Brown et al., 1989; Rogers et al., 1977; Symons & Johnson, 1997).

On the other hand, a potential disadvantage of using autobiographical stories as mnemonic scaffolds might be that recall of autobiographical narratives is not stable and based on the narrative and other factors (e.g., Greenberg & Rubin, 2003; Habermas, 2018; McAdams & McLean, 2013; Hirst & Echterhoff, 2011), including individual abilities (Rubin, 2020, 2021). To adjudicate between these two predictions we tested whether recall accuracy of the story scaffolds themselves has an effect on recall accuracy of the study items.

### 3.1.2 Recall accuracy of the scaffolds themselves and order in mnemonic scaffolds

It stands to reason that, if one cannot remember one's scaffold, it will be a poor scaffold for serial recall. This is why in addition to investigating the role of locations in mnemonic techniques (see also Chapter 2) we were interested in whether a reliable order of sentences in story-based scaffolds contributes to their memory success.

Previous research on mnemonic scaffolds has shown that the strength of mnemonic scaffolds lies in facilitating memory for ordered information (Foer, 2011; Yates, 1966; Bouffard et al., 2017; Ericsson et al., 1980; Roediger, 1980). In serial recall, memory for order is commonly investigated with a strict position scoring criterion under which items are only scored as correct when recalled in the same position they were presented. Memory for items is scored with a lenient position scoring criterion under which an item is scored as correct if it was recalled in any position of the list. Bouffard et al. (2017); Roediger (1980) and the experiments described in Chapter 4 have shown that the mnemonic advantage of mnemonic scaffolds over non-scaffold strategies is higher when scored under the strict scoring criterion than under the lenient scoring criterion. This raises the question of whether mnemonic scaffolds that have a stable internal order and are easier to remember are more effective than mnemonic scaffolds where the order is not easily retrieved. While this question, to our knowledge, has never been tested directly, findings that the Method of Loci and Body Scaffold, the order of which is presumably stable (the order of anchors is prescribed by the fixed order of body parts, or can be retrieved by following a fixed route without backtracking) outperform Autobiographical Story Scaffolds, the order of which is variable (autobiographical events are not reliably retrieved in a chronological order (Loftus & Fathi, 1985)) suggest that order influences the effectiveness of mnemonic scaffolds. We further explain this in Chapter 4. To investigate whether reliability of order predicts the memory success of story-based mnemonic scaffolds, participants completed a recall attempt of their scaffolds. This allows us to test whether recall accuracy of the scaffold itself predicts recall accuracy of the study items. It also allows us to check whether it might be promising to improve participants' memory of their story as a way to improve their performance with story-based mnemoic scaffolds.

### 3.1.3 Scaffold-word associations and verification of strategy use

As described in Chapter 2, instructions to apply a mnemonic strategy are no guarantee that participants actually use the strategy to study and recall list items (Bellezza, 1981). Previous studies have simply assumed compliance of participants without attempting to verify it objectively (e.g., Roediger, 1980; Bouffard et al., 2017). As in Chapter 2, we included the scaffold–cued recall task, in which we tested whether participants formed scaffold-word associations.

We expect to replicate the finding from Chapter 2 that when participants can recall the study item when cued with sentences they were asked to associate it with, they may rely on the association during serial recall. In other words, we expect scaffold-cued recall to predict serial recall. Alternatively, it is conceivable that with longer lists and stories, participants may not be able to recall the words during serial recall despite being able to recall them when cued with the anchor (Sahadevan et al., 2021). In this case, the story scaffolds for list length fifteen would only show a benefit for scaffold-cued recall but no benefit for serial recall.

### 3.1.4 Goals and hypotheses of the current study

Our specific goals were to a) test whether the effect of Autobiographical Mnemonic Scaffolds and locations mentioned within those stories on serial recall accuracy for lists of ten words transfers to lists of fifteen words, b) compare autobiographical to fictional mnemonic scaffolds and to a Control, and c) test whether recall accuracy of the story scaffolds themselves contributes to their effectiveness.

The experiment first started as an in-person experiment but because of the pandemic, we moved our data collection online (for more details on data collection, refer to the Methods section). We hypothesize that participants who use Autobiographical Mnemonic Scaffolds outperform participants who use fictional ones given the memory-enhancing self-relevance effect described above. We also hypothesize that recall accuracy of the stories themselves contributes to the effectiveness of both autobiographical and fictional story scaffolds because a reliable order may facilitate self-cueing of anchor points and thus enhance serial recall. We expect participants in the Control Group who use a reading loud strategy to perform at the lowest levels because they do not use a mnemonic scaffold. As explained above, we also expect that scaffold-cued recall predicts serial recall accuracy.

## 3.2 Methods

The experiment was pre-registered at `osf.io/t7zpq`. It was programmed and run in MATLAB with Psychophysics Toolbox experiment programming extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997), and the CogToolbox (a set of functions for MATLAB and Psychophysics Toolbox 3 for cognitive psychology experiments). Library (Fraundorf et al., 2014). It was later ported to PsychoPy3 (Peirce et al., 2019), translated to JavaScript for the `Pavlovia.org` platform. We used the most effective presentation mode determined in Chapter 2, i.e., presenting one study item together with one sentence at a time in the study phase. We did not use the presentation mode that requires participants to type their stories because is it less efficient in that typing takes more time and thus fewer list can be studied.

The overall procedure of the experiment described in Chapter 2 and of the present experiment is similar, but we added two phases to the experiment described below. Participants completed a pre-instruction, baseline memory task, where they recalled lists of words in serial order prior to receiving instructions on mnemonic scaffolds. This allows for a within-subject comparison of recall accuracy with and without the technique. In addition, participants in the experiment described here completed a recall attempt of their stories, which allows us to test whether recall accuracy of the story scaffold predicts recall accuracy of the study items.

**Participants**

In our first attempt to collect data, we recruited 76 participants from the University of Alberta introductory psychology research participation pool in partial fulfillment of course requirements. Those participants were required to have English as their first language and/or have learned English before the age of six and were all older than 17 years old. Written informed consent was obtained prior to the experiment in accordance with a University of Alberta ethical review board. Due to the pandemic, we were unable to reach our pre-registered target sample size of $> 30$ participants per group. This is why we moved the experiment online to collect a complete data set. For the main data set presented here, participants (N = 128) were recruited through the online participant recruitment service Prolific (`https://prolific.co/`). We consulted the in-person data to check for the possibility of differences due to the online versus in-person samples, and report the online data fully since this data set was complete and met our pre-registered sample size.

The experiment was run on the online data collection platform `Pavlovia.org`. As with the participants from the experiment of Chapter 2 and the participants from the incomplete in-person data set, all online participants were required to have English as their first language and/or have learned English before the age of six and were all older than 17 years old. Written informed consent was obtained prior to the experiment in accordance with a University of Alberta ethical review board. Participants whose recall accuracy was at ceiling (1 word or less incorrect) or at floor (1 word or less correct) in either the pre-instruction phase, the post-instruction phase or both, or whose stories had more than one missing or more than one incomprehensible sentence per story were excluded from the final sample. In addition, participants who wrote an autobiographical story despite being in the Fictional Group, participants who wrote a review of a book or movie, the lyrics of a song, or a prayer were excluded. After those exclusions, the final sample included 99 participants. Participants were required to be fluent in and or have English as their first language, and an approval rating[1] greater than 70%.

### Materials

As in Chapter 2, study items were 4-8 letter nouns drawn at random (without replacement) from the Toronto Word Pool with frequency ratings by Kucera and Francis,[2] also used by Bouffard et al. (2017). Each word list comprised fifteen words. Participants completed the experiment remotely on their own computers.

### Overview of the procedure

After agreeing to the consent form, participants studied and recalled 3 lists without instruction in a mnemonic strategy, which served as our measure of uninstructed baseline memory. After that, participants were informed that they were going to learn an effective memory technique that they were to use to study more word lists.

Next, participants in the Autobiographical Group were asked to type an event from their own life they remember well, such as their first day of school or a holiday trip. Participants in the Fictional Group were asked to type a passage from their favourite book or movie, split up into fifteen sentences. Immediately after typing the scaffold, participants were asked to recall the sentences from their story in order. Participants then used their self-generated scaffold to study more word lists. Immediately following serial recall of each list, participants were shown the words from the just-recalled list in randomized order and asked for the study items they associated with each sentence. Subsequently, participants answered three self-report questions on strategy use and usefulness. The phases of the experiment and justifications for their use are described in detail below.

**Pre-instruction serial recall** The first phase measured serial recall ability prior to instruction and was thus identical for the three groups. Participants studied three lists of

[1]representing the percentage of studies which a participants has completed by following the instructions
[2]`http://memory.psych.upenn.edu/Word_Pools` accessed 12/01/2021

fifteen words and were tested with serial recall. Study items were presented individually, centrally on the screen, and remained on the screen until participants pressed ENTER. After viewing the whole list, participants were asked to type the words in the order they were presented. Each response was entered on a separate response line. The response lines were not numbered. They were visible in a vertical configuration from the start of recall. Participants were not given the option to edit their responses apart from using the backspace key, before pressing ENTER to submit the response. Typed words remained on the screen until all responses were entered. Participants were instructed to type PASS if they could not remember a word.

**Story generation and recall** As in Chapter 2, participants in the Autobiographical Group were asked to think of a personal event they remember well, such as their first day of school or a holiday trip in first-person perspective, and type the story split up into individual sentences. Participants in the Fictional Group typed up a passage from their favourite book or movie. Examples given were a scene from Harry Potter or the ending of a favourite book. The length of the story was increased to fifteen sentences in this experiment because lists of fifteen words were studied. This is because we wanted to test whether the mnemonic benefit is higher for longer lists compared to Chapter 2, where the difference was highest in recency positions. Participants were given the option to edit their sentences or change the order within their story. As shown in previous studies on mnemonic scaffolds, proactive interference is not to be expected (e.g., Massen & Vaterrodt-Plünnecke, 2006; Bass & Oswald, 2014; Legge et al., 2012; Caplan et al., 2019). We also explain this in Chapters 2 and 4.

Next, participants did a recall attempt of their story by typing the story in the same order they typed it earlier. This allowed us to measure recall accuracy of the story and to test whether that predicts recall accuracy of the study items. Participants did not have the option to edit the story during nor after the recall attempt.

**Post-instruction study, serial, and scaffold-cued recall** The post-instruction study and recall phase of the present experiment followed the same scheme as in Chapter 2. After generating their stories, participants received instructions on how to use those to study six more word lists. For the two experimental groups, this entailed associating one word with one sentence from their story. Scaffold-word pairs were presented in serial order with the list word in uppercase in the center of the screen and the respective sentence below. Participants saw one scaffold-word pair on each screen and pressed ENTER after they were satisfied that they had associated the word with the sentence to get to the next one. Participants in the Control Group received the filler instruction to read the words aloud to (supposedly) make remembering easier (Bodner & Taikh, 2012) (but see Chapter 4 for the controversy about the production effect). After participants in all groups had studied a list, they were asked to recall the list in serial order. The sentences from the stories were not shown during the recall phase. Following serial recall of each list, participants in the experimental groups were tested directly for memory for scaffold-word associations (regardless of whether or not the story had just supported serial recall of the study items) with the scaffold-cued recall task.. As the Control Group did not use their stories for serial recall, the scaffold-cued recall task only applied to the experimental groups. After scaffold-cued recall of each list, participants in the experimental groups were informed how many lists remained to be studied; participants in the Control Group, who did not have a scaffold-cued recall task, received this information after serial recall of each list.

**Self-report questions** Participants in the two experimental groups answered the following self-report questions at the end of the experiment: 1.) Did you envision the list words as part of your story when studying them? 2). Did integrating the words into the story make remembering the words easier? 3.) Did you use a different strategy to study the word lists? Briefly describe what you did. The response options for question 1.) were Always/Mostly/Sometimes/Never for question 2.) Yes/No/I don't know.

### Scoring of locations

The autobiographical and fictional stories typed by each participant were later scored by two researchers to count the number of locations. The criteria for what constitutes a location were the same as in Chapter 2. If there was any disagreement between the number of locations counted by each scorer, the scorers scored the individual sentences together and resolved the disagreement for each case.

#### Scoring of recall accuracy of the scaffold

The recall attempt of the stories was scored by two researchers using a scoring program written in MATLAB that displayed the original story (with numbered sentences) participants wrote together with each sentence from the recall attempt individually on the same screen. The scorers entered the number of the sentence from the recall attempt in the original story, and the scoring program collected the numbers of the sentence. If there was any disagreement regarding the sentence number in the original story, the scorers scored those respective sentences together and resolved the disagreement for each case. Recall accuracy of story recall was scored with the strict scoring criterion (a recalled sentence was only scored as correct if recalled in its original position) and the lenient scoring criterion (a recalled sentences was scored as correct if recalled in any position) because we wanted to test whether sentences recalled both in order and regardless of order predict recall accuracy of the study items.

### 3.2.1 Data analyses

We used the same data analyses approach as in Chapter 2 and followed our pre-registered analyses plan. We call the main factor "Group" when the Control is included and "Scaffold Type", when the Autobiographical and Fictional Group are compared.

## 3.3 Results

We asked whether autobiographical or fictional story scaffolds have a mnemonic benefit over Control for lists of fifteen words. We also asked whether locations and recall accuracy of the story scaffolds increases their effectiveness. Foreshadowing our results, we found supported nulls for the effect of scaffold. Consequently, the conditions to test some of our hypotheses were not met. We report more detailed analyses below.

### 3.3.1 Preliminary results from the in-person data set

Even though we were not able to complete our data-set with in-person participants from the University of Alberta introductory psychology research participation pool due to the pandemic, we report the findings here as preliminary results. Since the number of participants in each group was unequal with 33 participants in the Autobiographical Group, 27 participants in the Fictional Group, and only 16 participants in the Control, our planned comparison of the the Groups in an ANOVA is statistically under-powered. This is why we conducted frequentist and Bayesian paired samples t-test comparing pre- to post recall accuracy in both experimental groups. In both the Autobiographical ($t(26) = -0.74$, $p = 0.464$, $BF_{10} = 0.24$), and the Fictional Group ($t(26) = -0.42$, $p = 0.671$, $BF_{10} = 0.22$), we found no difference between pre- and post-instruction accuracy (strict scoring). This suggests that the instructions to use autobiographical or fictional stories had no effect on recall accuracy.

### 3.3.2 Results from the main data set

#### Validity of the sample

To ensure the validity of our main sample, we tested whether there was a subject sampling bias across groups and a learning-to-learn effect in the Control, as described below.

**Absence of a subject sampling bias** To verify the absence of a subject sampling bias across groups, we conducted a 3 (Autobiographical, Fictional, Control) $\times$ 15 (Serial Position 1-15) mixed ANOVA on the average serial recall accuracy for each serial position averaged across the three lists that were studied prior to receiving any information on mnemonic techniques (Figure 3.4 dotted lines, Table 3.1).

This produced neither a significant main effect of Group (strict: $F(2, 96) = 0.06$, $p = 0.940$, $\eta_p^2 < 0.01$, $BF_{inclusion} = 0.07$, lenient: $F(2, 96) = 0.39$, $p = 0.679$, $\eta_p^2 = 0.01$, $BF_{inclusion} = 0.05$), nor a significant interaction (strict: $F(15.31, 734.69) = 0.78$, $p = 0.706$, $\eta_p^2 = 0.02$, $BF_{inclusion} < 0.01$, lenient: $F(13.70, 657.65) = 1.12$, $p = 0.342$, $\eta_p^2 = 0.02$, $BF_{inclusion} = 0.03$), with Bayes factors providing strong evidence for the null, confirming the absence of a subject sampling bias.

**Absence of a learning-to-learn effect in the Control** To check whether participants in the Control received a memory benefit from the filler instruction (reading the words out loud) we conducted frequentist and Bayesian paired-samples $t$-tests comparing pre- with post-instruction recall accuracy in the Control (Figure 3.4c). The $t$-tests were non-significant ($p < 0.05$) with the Bayes Factor providing evidence ($BF_{10} < 0.3$) for the null or with only anecdotal evidence for an effect, suggesting the absence of a measurable learning-to-learn effect in the Control (Table 3.1).

| | Autobiographical | | Fictional | | Control | |
|---|---|---|---|---|---|---|
| | Pre | Post | Pre | Post | Pre | Post |
| **Strict** | 0.39(0.19) | 0.43(0.23) | 0.37(0.24) | 0.46(0.27) | 0.37(0.24) | 0.42(0.30) |
| | $t(40) = 1.13$ | | $t(30) = 1.60$ | | $t(27) = 1.60$ | |
| | $p = 0.267$ | | $p = 0.120$ | | $p = 0.122$ | |
| | $BF_{10} = 0.30^*$ | | $BF_{10} = 0.61$ | | $BF_{10} = 0.62$ | |
| **Lenient** | 0.64(0.16) | 0.66(0.20) | 0.66(0.21) | 0.66(0.20) | 0.66(0.21) | 0.64(0.12) |
| | $t(40) = 0.83$ | | $t(30) < 0.01$ | | $t(27) = -0.57$ | |
| | $p = 0.414$ | | $p = 1.000$ | | $p = 0.573$ | |
| | $BF_{10} = 0.23^*$ | | $BF_{10} = 0.19^*$ | | $BF_{10} = 0.23^*$ | |

Table 3.1: Comparison of pre- and post-instruction recall accuracy in all groups, including the Control, where the null effect confirms the absence of a learning-to-learn effect. No significant effects were found, and the Bayes Factors either provide evidence for a null effect or are in the inconclusive range. The mean and standard deviation for pre- and post-instruction recall accuracy are reported as M(SD); asterisks denote significant p-values conclusive Bayes Factors $BF_{10} > 3$ $BF_{10} < 0.3$.
.

| | Strict | Lenient | Cued |
|---|---|---|---|
| **Main Effect** | $F(2, 96) = 0.13$ | $F(2, 96) = 0.10$ | $F(1, 96) = 0.44$ |
| | $\eta_p^2 < 0.01$ | $\eta_p^2 < 0.01$ | $\eta_p^2 < 0.01$ |
| | $p = 0.880$ | $p = 0.907$ | $p = 0.508$ |
| | $BF_i = 0.05^*$ | $BF_i = 1.82$ | $BF_i = 0.19^*$ |
| **Interaction** | $F(12.03, 577.08) = 1.06$ | $F(11.28, 541.46) = 2.36$ | $F(10.05, 693.14) = 0.94$ |
| | $\eta_p^2 = 0.02$ | $\eta_p^2 = 0.05$ | $\eta_p^2 = 0.01$ |
| | $p = 0.391$ | $p = 0.007^*$ | $p = 0.495$ |
| | $BF_i = 0.03^*$ | $BF_i = 9.72^*$ | $BF_i < 0.01^*$ |

Table 3.2: Comparison of the mnemonic scaffolds and Control. The median and standard deviation for pre- and post-instruction recall accuracy is reported as M(SD); $BF_i$ is short for $BF_{inclusion}$; significant $p$-values ($p < 0.01$) and conclusive Bayes Factors $BF_i > 3$ $BF_i < 0.3$ or are marked with an asterisk.

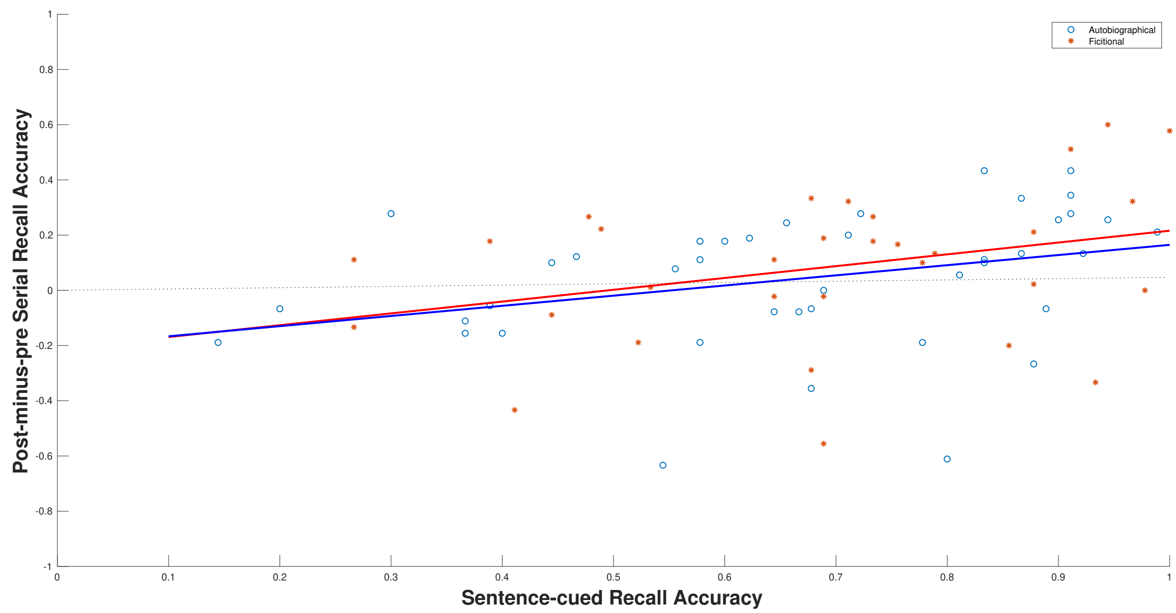



Figure 3.1: Linear regressions of serial recall accuracy (strict scoring) against scaffold-cued recall accuracy. Each point plots one participant, and the red and blue lines plot the best-fitting linear function.

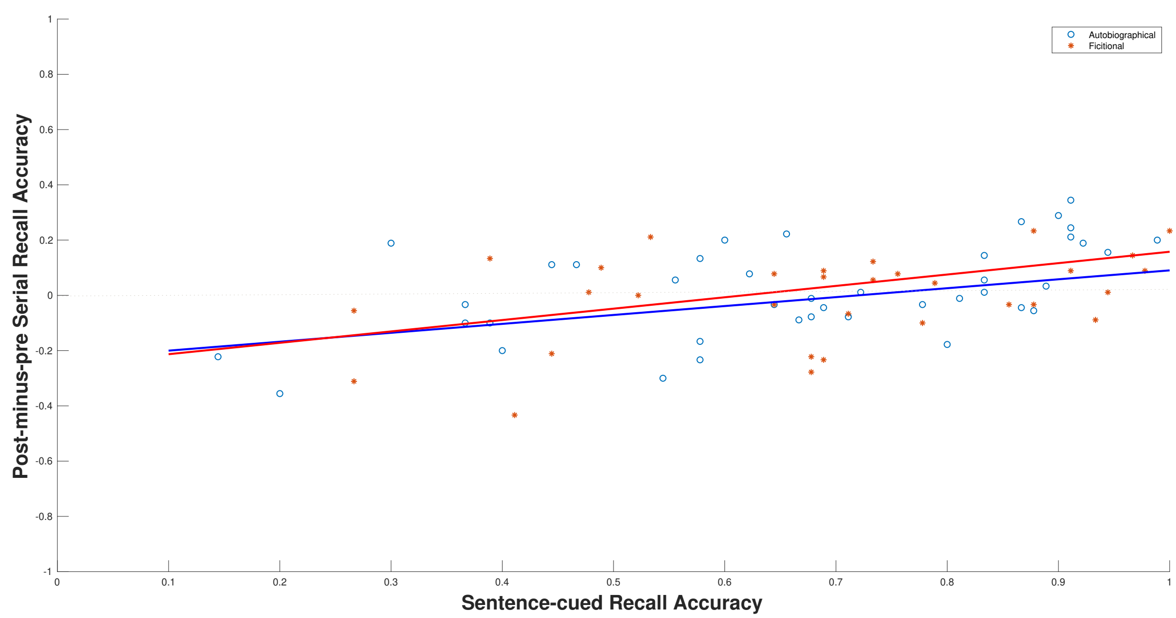



Figure 3.2: Linear regressions of serial recall (lenient scoring) accuracy against scaffold-cued recall accuracy. Each point plots one participant, and the red and blue lines plot the best-fitting linear function.

**Absence of an effect of list number** We were interested in whether participants' serial recall accuracy of each word list changed over the course of the experiment. An increase in serial recall accuracy over the course of the experiment could indicate a learning-to-learn effect, while a drop serial recall accuracy in later lists could indicate fatigue. To test this, we conducted a 3 (Autobiographical, Fictional, Control) $\times$ 6 (list 1-6) frequentist and Bayesian mixed ANOVA on the average serial recall accuracy (strict scoring) for each serial list (Figure 3.3). The ANOVA neither found a significant main ($F(2, 98) = 0.43$, $p = 0.958$, $\eta_p^2 < 0.01$, $BF_{inclusion} = 0.09$) nor interaction effect ($F(10, 490) = 0.54$, $p = 0.863$, $\eta_p^2 = 0.01$, $BF_{inclusion} < 0.01$) with Bayes factors confirming the absence of an effect of list number. This suggests that list number did not have an effect on serial recall accuracy. This replicates previous studies that have found no evidence of proactive interference or fatigue over successive lists using the method of loci or similar strategies (e.g., Bass & Oswald, 2014; Caplan et al., 2019; Legge et al., 2012; Massen & Vaterrodt-Plünnecke, 2006)

**Relationship between pre-instruction and scaffold-cued recall accuracy** To test whether there was a pre-existing relationship between pre-instruction, baseline memory and scaffold-cued recall accuracy, we conducted linear regressions of strict and lenient pre-instruction recall accuracy based on scaffold-cued recall accuracy combining both experimental groups. For both strict ($p < 0.002$, $\beta = 0.36$, $R^2 = 0.12$, $BF_{10} = 17.30$) and lenient ($p < 0.001$, $\beta = 0.57$, $R^2 = 0.32$, $BF_{10} > 100$) scoring, scaffold-cued recall accuracy predicted pre-instruction memory. Due to this confounding, pre-existing relationship, we subtracted pre from post-instruction recall accuracy and used this measure of memory increase for our regressions of scaffold-cued accuracy.

### Effect of scaffold type

To test our hypothesis that Autobiographical Story Scaffolds have an advantage over fictional ones, we conducted within-subject and between-subject comparisons of scaffold-cued, strict and lenient recall accuracy reported below.

**Effect of the story types on scaffold-cued recall** Before asking how the story types compare in facilitating serial recall accuracy, we were interested in whether participants may have relied on scaffold-cued recall when recalling the words in serial order. Due to the

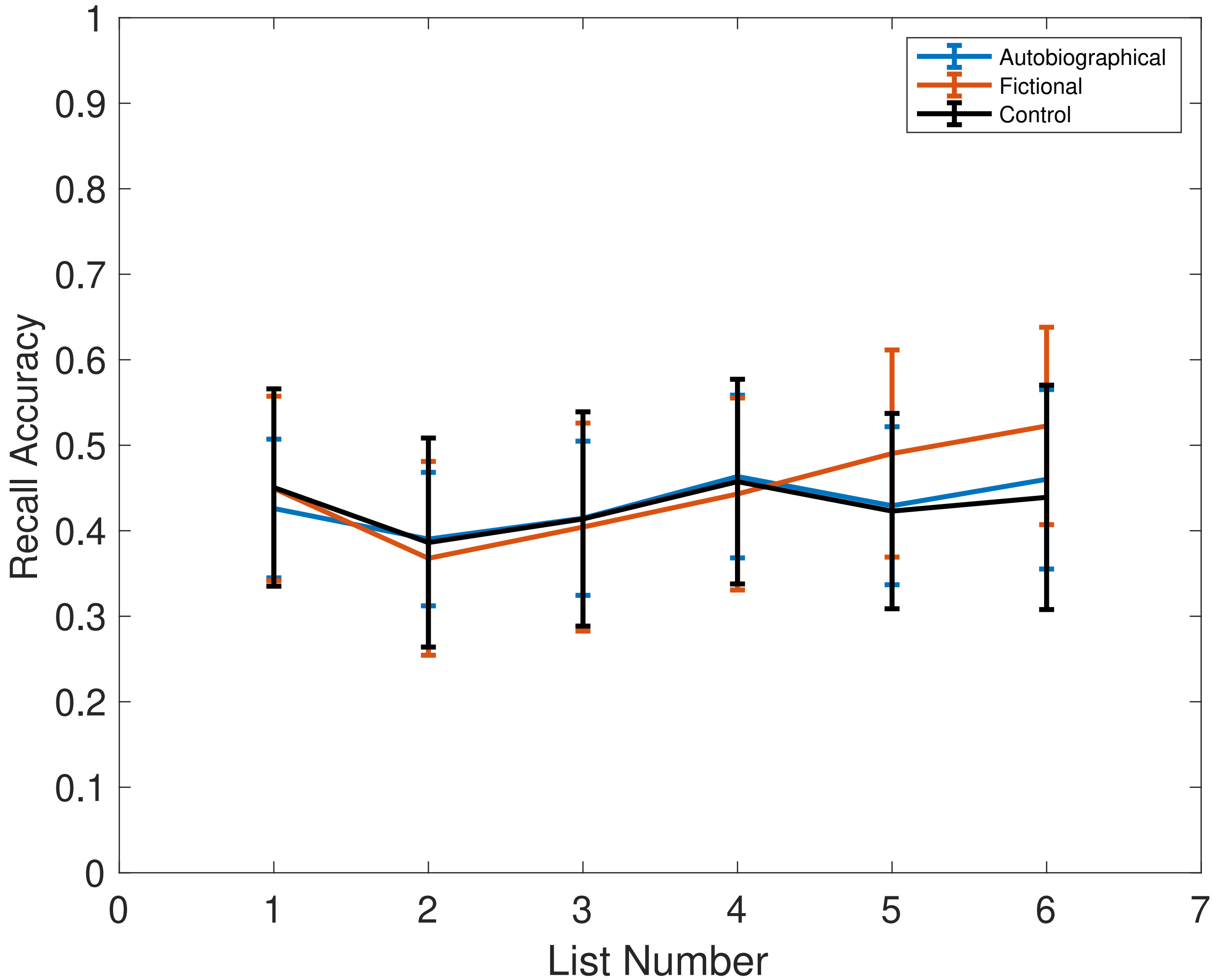



Figure 3.3: Serial Recall accuracy for each list averaged across serial position for all Groups. Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994)

confounding relationship between baseline memory and scaffold-cued recall reported above, we subtracted pre- from post-instruction recall accuracy (strict scoring) and tested whether scaffold-cued recall accuracy predicts this measure of memory-increase. This was the case in both Scaffold Types for both strict (Autobiographical: $p = 0.037$, $\beta = 0.33$, $R^2 = 0.08$, $BF = 1.85$; Fictional: $p = 0.045$, $\beta = 0.37$, $R^2 = 0.11$, $BF_{10} = 1.71$; both experimental groups combined: $p = 0.005$, $\beta = 0.33$, $R^2 = 0.09$, $BF_{10} = 7.63$) and lenient scoring (Autobiographical: $p < 0.001$, $\beta = 0.53$, $R^2 = 0.23$, $BF_{10} = 78.39$; Fictional: $p = 0.022$, $\beta = 0.41$, $R^2 = 0.14$, $BF_{10} = 2.80$; both experimental groups combined: $p < 0.001$, $\beta = 0.48$, $R^2 = 0.23$, $BF_{10} > 100$). This suggests that participants who formed scaffold-word associations may have relied on those during serial recall.

Next, we were interested in whether the two story types differed in the success with which participants formed scaffold-word associations as instructed. To test this, we analyzed the scaffold-cued recall data from the scaffold-cued recall task on their own in a 2 (Autobiographical, Fictional) × 15 (Serial Position 1-15) mixed ANOVA on the average scaffold-cued recall accuracy for each serial position. This revealed a non-significant main ($F(1, 70) = 1.78$, $p = 0.705$, $\eta_p^2 = 0.01$, $BF_{inclusion} = 0.18$) and interaction ($F(10, 693.14) = 0.94$, $p = 0.495$, $\eta_p^2 = 0.01$, $BF_{inclusion} < 0.01$) effect with Bayes Factors supporting the null, suggesting no difference between scaffold-cued recall accuracy in the two experimental groups.

**Within-subject pre-post comparison** To test whether the instruction to use autobiographical or fictional story scaffolds has an effect on recall accuracy within-subjects, we conducted frequentist and Bayesian paired samples t-tests comparing post- and pre-instruction recall accuracy for both scoring methods in all groups (Figure 3.4, Table 3.1). Those were non-significant with Bayes Factors either in the inconclusive range or providing strong evidence for a null effect. This suggests that neither the Autobiographical Scaffold nor the Fictional Scaffold, nor Control had an effect on within-subject recall accuracy for neither strict or lenient scoring.

**Between-subject comparison** To test whether Group had an effect on recall accuracy, we conducted 3 (Autobiographical, Fictional, Control) and 2 (Autobiographical, Fictional) × 15 (Serial Position 1-15) mixed ANOVAs on the average serial recall accuracy for each serial position averaged across the six lists from the post-instruction phase for strict, lenient,

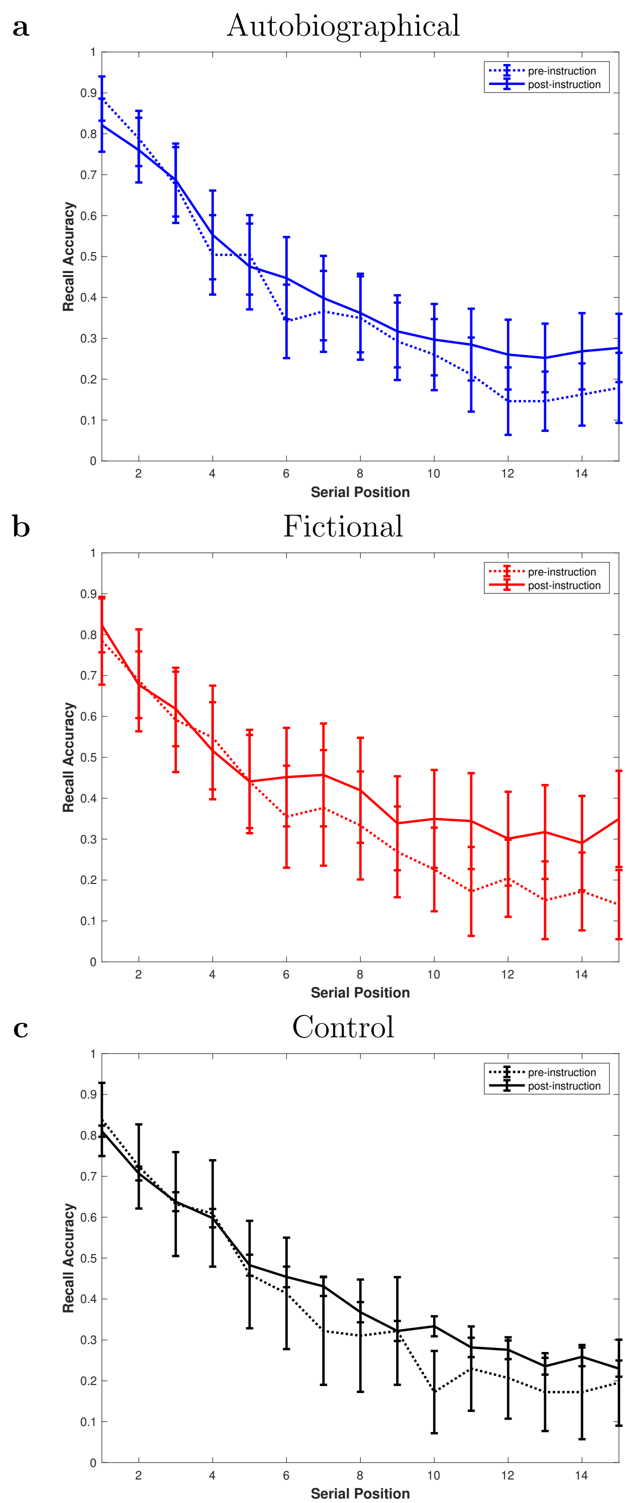



Figure 3.4: Serial Position Curves for individual groups comparing pre- and post-instruction serial recall accuracy (strict scoring). Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994).

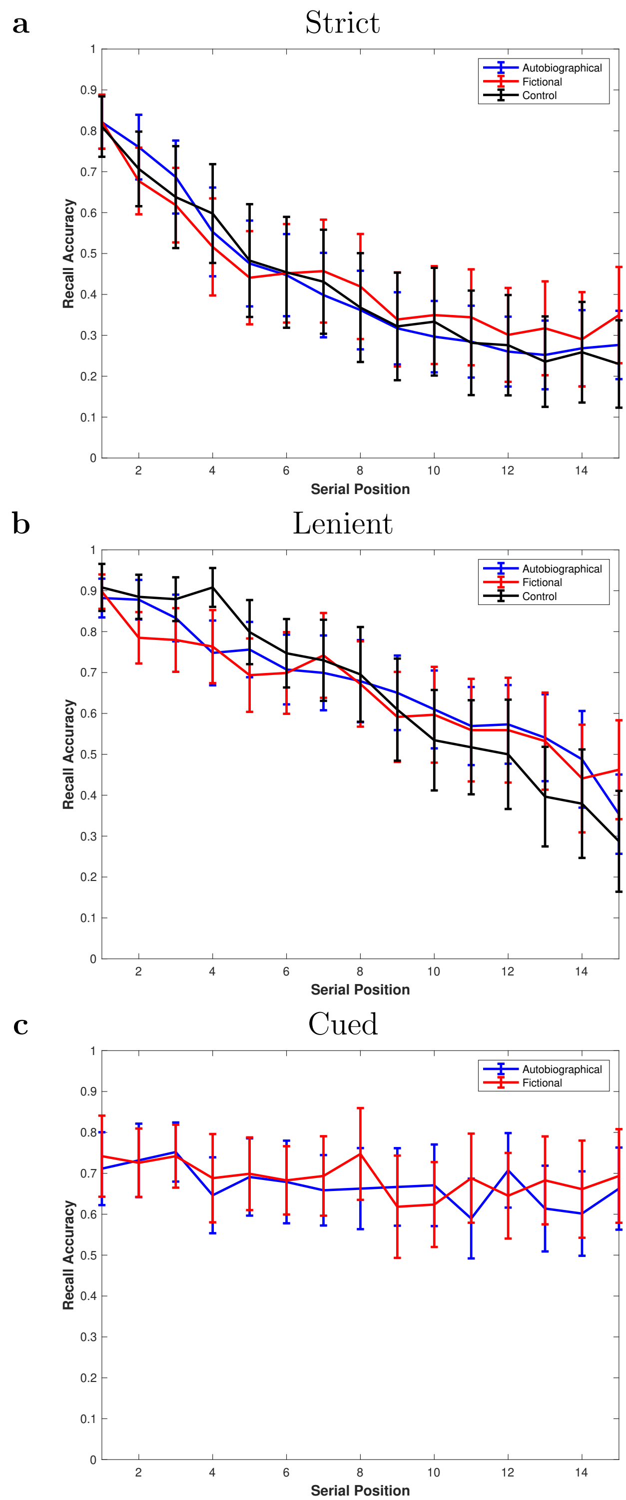



Figure 3.5: Post-instruction accuracy as a function of serial position, as a function of group. Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994).

and scaffold-cued recall accuracy (Figure 3.5, Table 3.2).

For strict and cued post-instruction recall (Figure 3.5a,c), the ANOVAs produced non-significant main and interaction effects with Bayes Factors providing strong evidence for a null effect of Group. For lenient scoring (Figure 3.5b) the interaction was significant with the Bayes Factor providing evidence for an interaction. To follow up on this significant interaction, we conducted one-way ANOVAs at each serial position. These ANOVAs were significant at position 2, where the Control and Autobiographical Group outperformed the Fictional Group ($p < 0.05$), at position 4, where Control outperformed the Autobiographical Group ($p < 0.05$), and at position 15, where the Fictional Group outperformed Control ($p < 0.05$). Together, these findings suggest that neither autobiographical nor fictional scaffolds improve memory when studying lists of fifteen words across the whole list, but there were differences at individual positions.

**Self-reported Compliance** Next, we were interested in whether answers to the self-report questions on strategy use differed between the story experimental groups. Therefore, we conducted a chi-square test on those two questions. Each participant responded with the same number keys for both questions (nobody responded with "Never" in question 2), and the chi-square test was non-significant for ($\chi^2(2) = 0.412$), suggesting that the answers were distributed evenly across the two groups.

### Spatial Locations

To investigate the relationship between recall accuracy and number of locations used in the story scaffolds, we conducted simple linear regressions for the Autobiographical and Fictional Group. The dependent variable was the respective type recall accuracy (post-instruction strict, post-instruction lenient, and scaffold-cued recall respectively), and the explanatory variable was number of locations mentioned in the stories. In neither the Autobiographical nor in the Fictional Group, number of locations predicted neither strict post-instruction accuracy, lenient post-instruction accuracy, nor scaffold-cued accuracy ($p > 0.05$). Bayes Factors either provided evidence for the null or were in the inconclusive range. This suggests no relationship between number of locations mentioned in the story scaffolds and recall success.

### Story Recall Accuracy

To investigate the relationship between recall accuracy of the story scaffolds themselves and increase in serial recall accuracy (post-instruction minus pre-instruction recall) we conducted simple and linear regressions for all groups. The dependent variable was memory increase (strict and lenient scoring), and the explanatory variable was story recall accuracy (strict and lenient scoring respectively). For strict scoring memory increase with strict scoring story recall accuracy ($R^2 = -0.02$, $\beta = 0.12$, $p = 0.536$, $BF_{10} = 0.40$) and for lenient storing memory increase with strict scoring story recall accuracy ($R^2 = -0.03$, $\beta = 0.06$, $p = 0.738$, $BF_{10} = 0.36$), the relationship was non-significant with an inconclusive Bayes Factor. This was also true for strict scoring memory increase with lenient scoring story recall accuracy ($R^2 < -0.01$, $\beta = -0.17$, $p = 0.360$, $BF_{10} = 0.47$), and for lenient scoring memory increase with lenient scoring story recall accuracy ($R^2 = -0.04$, $\beta = -0.02$, $p = 0.899$, $BF_{10} = 0.35$). This suggests that recall accuracy of the story scaffolds themselves do not affect memory success.

## 3.4 Discussion

The goals of the experiment described in the present chapter were to a) test whether the effect of Autobiographical Mnemonic Scaffolds and locations mentioned within those stories on serial recall accuracy for lists of ten words generalizes to lists of fifteen words, b) compare autobiographical to fictional mnemonic scaffolds and to a Control condition, and c) test whether recall accuracy of the story scaffolds themselves contribute to their effectiveness.

We summarize our main findings below, before we discuss them in more detail in the following paragraphs. The effect of Autobiographical Mnemonic Scaffolds on recall accuracy for lists of ten words described in Chapter 2 was not replicated with list length fifteen in the present experiment. We found a benefit of the instructions in neither the Autobiographical Group, nor the Fictional Group over neither uninstructed baseline memory nor the Control condition. The overlapping null-effect between pre-and post instruction recall accuracy in the Autobiographical and Fictional Group in the in-person data set and the online data set suggest that this finding is not confounded by the data collection modality.

The positive relationship between spatial locations mentioned in the autobiographical stories observed in Chapter 2 was also not replicated in the present experiment. Neither did recall accuracy of the story scaffolds themselves correlate with serial recall accuracy. Interestingly, however, we found a significant positive relationship of scaffold-cued recall accuracy and memory increase in both experimental groups, suggesting that participants who complied with the instructions to form scaffold-word associations did benefit from those during serial recall. We discuss these findings in more detail below.

### 3.4.1 List-length may be a boundary condition in story-based scaffolds in novice users

Given the self-reference effect in memory (e.g., Rogers et al., 1977; Bower & Gilligan, 1979; Brown et al., 1989; Symons & Johnson, 1997), we hypothesized that autobiographical scaffolds may have an advantage over fictional ones. However, in the present experiment with lists length fifteen, we failed to find a mnemonic benefit in both scaffold conditions. This precludes us from answering the question of whether spatial locations or recall accuracy of the scaffold themselves contribute to their effectiveness.

When it comes to interpreting the null effects of our analyses, it is important to note that, even though we ruled out fatigue, there are other factors such as perceived task difficulty, cognitive load, motivation, etc. that we did not control for.

The reason why we increased list length to fifteen in the present experiment was the advantage of the most effective presentation mode determined in Chapter 2 over Control in recency positions leading us to speculate that this advantage would be even more pronounced for longer lists. On the contrary, however, the mnemonic benefit across positions was not found for list length fifteen. Notably, in the experiment described in Chapter 4, we were able to replicate the mnemonic benefit of the Autobiographical Story Scaffold over uninstructed recall for list length ten. We therefore presume that list length might be a boundary condition for story-based mnemonic scaffolds for novice users.

We did not observe any sign of fatigue, since participants' recall accuracy did not drop over the course of the experiment, which may have affected how the groups may have differed from one another. Even though we can only speculate, it seems plausible that it is considerably more challenging to come up with stories that are five sentences longer than stories consisting of ten words. If the scaffold generation phase was already perceived as challenging, this might have decreased participants' motivation and effort for the study and recall phases. Additionally, the quality of the stories might decrease as participants might struggle to come up with additional sentences even if they might think that their story is already complete. In other words, if some sentences lack detail because participants inserted filler sentences as a result of having to write fifteen sentences to proceed in the experiment, they might find it hard to associate list words with these sentences.

An important factor when it comes to using story-based mnemonic scaffolds effectively is that, as mentioned earlier, participants were not aware that they were going to use their stories as mnemonic scaffolds. Instead, they were asked to associate words with the sentences from their stories immediately after they generated them. With more time to construct the scaffolds, and the information that those will be used to study words, some participants might have come up with more effective stories, especially if they had have the opportunity to try out the strategy with different stories. In two case studies of intensive uninstructed memory training, two individuals were able to increase their memory for digits by a factor greater than ten based on the insights they gained from trial and error of constructing and using mnemonic scaffolds consisting of prior knowledge (Ericsson et al., 1980; Ericsson, Delaney, Weaver, & Mahadevan, 2004). Thus is plausible that "trying out" the strategy in multiple sessions may lead to learners gaining valuable insights into which kind of stories, both autobiographical or fictional, are more effective than others.

### 3.4.2 Scaffold-word associations may be more important than recall accuracy of the scaffold

There was one important effect observed in Chapter 2 that we replicated in the present experiment: Scaffold-cued recall predicted serial recall. This suggests that participants who formed scaffold-word associations as instructed may have relied upon those for serial recall, and that individual scaffold-word associations may be the driving factor behind the effectiveness of mnemonic scaffolds. The effect was weaker in the present experiment than in the experiment described in Chapter 2, which could be because the task is more challenging for longer lists for various reasons as described above. However, it is important to mention that we only had post-instruction data in the Chapter 2 because the experiment did not include a pre-instruction recall phase. This is why the the same correlations may have been inflated because they would have included overall subject-variability effects. Overall, this finding has promising implications for the strategy to become more effective, even with longer word lists, when participants are more motivated to engage in this challenging task or receive more training or the opportunity to try out the strategy themselves.

Interestingly, recall accuracy of the story scaffolds themselves did not have an effect on serial recall accuracy of the study items. We had hypothesized that recall accuracy of the stories themselves contributes to the effectiveness of both autobiographical and fictional story scaffolds because a reliable order may facilitate self-cueing of anchor points and thus enhance serial recall. This assumption was based on findings that the strength of mnemonic scaffolds lies in facilitating memory for ordered information (Foer, 2011; Yates, 1966; Bouffard et al., 2017; Ericsson et al., 1980; Roediger, 1980). The lack of a significant positive relationship between of recall accuracy of the story scaffolds on memory increase suggests that individual scaffold-word associations may be more important than memory for the scaffolds themselves. It is important to note that participants were shown the sentences of their stories during study and were also shown their sentences in the scaffold-cued recall task. The repeated exposure to the stories may have reduced a possible effect recall accuracy of the story in the recall attempt. This is why, to answer the question of whether memory of the story scaffolds themselves contribute to their effectiveness, more research is needed.

In any case, the finding has important implications for the story scaffold technique to be improved upon: during study, it may be beneficial to instruct novice users of story-based mnemonic scaffolds to focus on integrating the study items into each individual sentence rather than thinking about their whole story. Similarly, during recall, it may be beneficial to instruct learners to focus on individual sentences as cues rather than focusing on recalling the whole story. Our results suggest that when developing a multi-step training program for story-based mnemonic scaffolds, the initial training stages should be focused on individual scaffold-word associations rather than fine-tuning of the stories.

## 3.5 Conclusion

Together, our findings suggest that for novice users, who only receive brief instructions on story-based mnemonic scaffolds, list length may pose a limit on the effectiveness of story-based mnemonic scaffolds. While there is evidence that the Autobiographical Mnemonic Scaffold enhances memory for lists of ten words (see Chapter 2 and 4), we found a supported null effect for lists of fifteen words for both autobiographical and fictional scaffolds. While number of locations mentioned in the stories and recall accuracy of the scaffolds themselves did not predict serial recall, and even though there was no overall effect of the scaffolds on serial recall accuracy, scaffold-cued recall predicted serial recall. This means that participants with more scaffold-word associations benefited more from both autobiographical and fictional scaffolds. This suggests novice learners who, despite the challenging task, follow the instructions to integrate the study items into the sentences from their story scaffolds, rely on those during serial recall. While further research on story-based mnemonic scaffolds and the effect of list length is needed, our findings may inform further development of this mnemonic technique: novice learners should be encouraged to focus on forming associations between study items and individual sentences rather than on their whole stories.

# Chapter 4

# Alternative mnemonic scaffolds to the Method of Loci

# Abstract

Memory champions remember vast amounts of information in order and at first encounter by associating each study item to an anchor within a scaffold— a pre-learned, structured memory. The scaffold provides direct-access retrieval cues. Dominated by the familiar-route scaffold (Method of Loci), researchers have little insight into what characteristics of scaffolds make them effective, nor whether individual differences might play a role. We compared participant-generated mnemonic scaffolds: a) familiar routes (Loci), b) autobiographical stories (Story), c) parts of the human body (Body), and d) routine activities (Routine Activity). Loci, Body, and Story Scaffolds benefited serial recall over Control (no scaffold). The Body and Loci Scaffold were equally superior to the other scaffolds. Measures of visual imagery aptitude and vividness and body responsiveness did not predict accuracy. A second experiment tested whether embodiment could be responsible for the high level of effectiveness of the Body Scaffold; this was not supported. In short, mnemonic scaffolds are not equally effective, and embodied cognition may not directly contribute to memory success. The Body Scaffold may be a strong alternative to the Method of Loci and may enhance learning for most learners including those who do not find the Method of Loci useful.

## 4.1 Introduction

Arguably the most effective mnemonic techniques are those that leverage previously learned material (Roediger, 1980; Staszewski, 1990). These techniques require learners to form associations between prior knowledge and study items in serial order. We refer to that ordered prior knowledge as a "scaffold." As demonstrated by memory world champions, after sufficient training, such techniques can enable learners to memorize vast amounts of information at first encounter (e.g., Foer, 2011). Superior memory does not require superior cognitive aptitudes or extraordinary brain anatomy (Chase & Ericsson, 1981; Ericsson &

Staszewski, 1989; Ericsson & Kintsch, 1995; Maguire et al., 2003; Hu et al., 2009; Takahashi et al., 2006; Wilding & Valentine, 2006). Through training, participants can improve serial recall by more than ten times untrained memory (Ericsson et al., 1980; Staszewski, 1990). The vast majority of research on mnemonic techniques is restricted to the Method of Loci, and fundamental questions about the cognitive processes underlying the effectiveness of mnemonic techniques, as well as desirable properties of mnemonic scaffolds, are largely unanswered. Only two studies, to our knowledge, have directly compared different scaffold-based mnemonic techniques (Bouffard et al., 2017; Roediger, 1980), and only one study has investigated individual differences in learner aptitudes predicting the usefulness of such techniques (Sanchez, 2019).

Here, we test two general hypotheses: a) that all scaffolds constructed from prior knowledge may provide a mnemonic benefit, b) that individual differences in skills and affinities related to the type of scaffold (specifically, visual imagery, spatial aptitude and body awareness), may partly determine memory success. In Experiment 1, we compare three mnemonic scaffolds to the Method of Loci and a no-scaffold Control. For reasons we explain below, the scaffolds were based on parts of the body (Body Scaffold), autobiographical stories (Story Scaffold), and routine activities (Routine Activity Scaffold). Surprised by the high level of success participants had with the Body Scaffold, in Experiment 2, we test the hypothesis that attention drawn to the human body drives the success of this mnemonic scaffold.

### 4.1.1 Mnemonic scaffolds and the role of prior knowledge

While the underlying cognitive mechanisms of mnemonic techniques are unclear, there is converging evidence to suggest that a large portion of the memory benefit of a scaffold is because prior knowledge can enhance new learning. In their Skilled Memory Theory, Chase and Ericsson (1981, 1982) argue that structures of existing memories providing retrieval cues are central to the effectiveness of mnemonic techniques (see also Roediger, 1980; Wenger & Payne, 1995; Bellezza, 1981; Ericsson & Staszewski, 1989). While studying a list, a mnemonic scaffold provides a system of pegs or anchors[1] to which new information is attached, or associated. During recall, the scaffold provides those anchors as a set of ordered retrieval cues.

Theories of expert memory provide examples of how superior memory for newly learned

[1]Not to be confused with usage of the term "anchor" in the judgement and decision-making literature.

information is connected with prior knowledge (Ericsson & Staszewski, 1989). These theories presume that prior knowledge improves memory by allowing new information to be associated with retrieval cues and to be integrated into the existing associative network (e.g., Brandt et al., 2005; Bruett et al., 2018; Ericsson & Staszewski, 1989; Long & Prat, 2002; Lane & Chang, 2018; Van Kesteren et al., 2012). Neurocognitive theories of the so-called "prior-knowledge effect" suggest that memories are initially episodic, and over time and with repeated retrieval, they are semanticized, or decontextualized (Raaijmakers, 1993; Carr et al., 1994). Interestingly, novel information can sometimes be rapidly integrated into prior, presumably semanticized, knowledge, leading to superior memory (McClelland et al., 1995; O'Reilly et al., 2014). Evidence suggests that in this way, memories that are normally hippocampal-dependent may take a fast route, bypassing the hippocampus to be stored immediately in neocortical areas, and taking on semantic-memory properties (Coutanche et al., 2014; Kan et al., 2009; Meeter & Murre, 2004; Sharon et al., 2011; Skotko et al., 2004; Smith et al., 2014; Sommer, 2017; Tse et al., 2007, 2011; Winocur & Moscovitch, 2011). In other words, whether due to bypassing the hippocampus or providing rich, reliable retrieval cues, or both, these convergent lines of research motivate our general hypothesis that all scaffolds generated from prior knowledge should support serial recall far above a no-scaffold control condition.

### 4.1.2 Mnemonic scaffolds comprising different types of prior knowledge

To our knowledge, there are only two studies that directly compared scaffold-based encoding strategies to one another. Roediger (1980) found that the Method of Loci and the Numerical Peg System— both scaffold-based mnemonic techniques— were more effective in facilitating recall than non-scaffold-based techniques. Bouffard et al. (2017) showed that the Method of Loci was as effective as scaffolds consisting of temporally ordered events. We directly compare the effectiveness of four mnemonic scaffolds, three of which have previously not been investigated, that harness four different types of prior knowledge.

#### The Method of Loci

The Method of Loci (also called Memory Palace or Mind Palace) is the most common mnemonic technique (Foer, 2011; Spence, 1984). The mnemonic scaffold used in the Method of Loci is a familiar route through a known environment. During study, the learner imagines walking the route, “placing” study items along the way by associating them with locations or objects along that route. During recall, learners re-walk the same route in their mind’s eye, “picking up” the study items in the same order they were studied. Some researchers have argued that visuospatial navigation and the engagement of the medial temporal lobe system are determining factors in the memory benefit provided by this method, due to the dual role of this network in navigation and episodic memory (e.g., Fellner et al., 2016; Moser et al., 2015; Rolls, 2017). Other findings cast doubt on this, suggesting that navigational cognition may be epiphenomenal, or at least not necessary to excel with the technique (Bouffard et al., 2017; Bower, 1970; Caplan et al., 2019). Instead, these researchers have suggested the effectiveness of the Method of Loci might derive from engaging the learner with the study material in much the same way as other mnemonic scaffolds or peg systems. If the effectiveness of the Method of Loci is not driven by imagined navigation but by features shared with other non-navigational mnemonic scaffolds, such as harnessing prior knowledge, we should see a similar recall accuracy when harnessing prior knowledge in the form of body parts, autobiographical stories, and routine activities.

### The Body Scaffold

Although there is almost no research on using the human body as a memory aid, there are historical and contemporary anecdotal accounts on a mnemonic scaffold based on the human body (Hunter, 1956). Gesualdo (1592) describes how to remember information by associating it with parts of the human body. Some memory athletes describe using their own body to remember information by associating study items with body parts (e.g., Foer, 2011; Konrad, 2013). Embodiment, specifically, refers to the notion that cognition depends on the sensorimotor capacities of the human body and that sensory and motor processes are inseparable in cognition (Varela et al., 1991). Behavioral and neural evidence has shown that language comprehension elicits activation within primary and secondary motor areas (Barsalou, 2008; Fischer & Zwaan, 2008; Toni et al., 2008; Pulvermüller, 2005; Handy et al., 2003). In the context of memory, D. C. Richardson et al. (2001) showed that the representation of a visual stimulus retrieved from memory can activate potential motor interactions, and that memory representations derived from linguistic descriptions can also activate motor affordances. Zimmer and Cohen (2001) argue that sensorimotor details lead to better memory performance due to better encoding elaboration, enabling association with preexisting memory representations.

In the only study, to our knowledge, in which body parts were used as memory cues, Bellezza (1984) presented participants with nouns and asked participants to come up with a body part or a personal experience that they deem a fitting memory cue for the respective study item. In a free recall task, participants then recalled both the study items and the body part or personal experience they had chosen as a memory cue. No difference in recall using the two types of cues was found (Bellezza, 1984). As the study did not include a control condition, it is unclear whether these memory cues facilitated recall. In contrast to our study, in Bellezza's (1984) study, the human body was not used as a mnemonic scaffold, but individual body parts were selected as cues after viewing the study item. Thus, the advantage of providing a sequence of retrieval cues in a fixed order was missed. It remains unknown whether prior knowledge in the form of body parts provides a mnemonic benefit as the Body Scaffold and the role of embodied cognition in mnemonic techniques has not been investigated.

We wondered if experimentally drawing additional attention to the body, or individual differences in tendency toward embodiment, might drive the success of the Body Scaffold. We incorporated these questions into the design of both experiments.

### The Autobiographical Story Scaffold

Little is known about whether prior knowledge in the form of autobiographical stories can boost memory for new material. Indeed, research and theory on the self-reference effect provide important theoretical arguments that autobiographical memories may serve as effective mnemonic scaffolds. A meta-analysis by Symons and Johnson (1997) highlights the importance of the self-reference effect in memory, emphasizing that self-referential encoding tasks yield superior memory in free recall, cued recall and recognition tasks relative to both semantic and other-referent encoding tasks. Symons and Johnson (1997) conclude that this is because the self is a well-developed and often-used construct that promotes elaboration and organization of encoded information. In addition, since autobiographical memories are highly self-relevant and rich in detail they possibly invoke extra-hippocampal structures, supplementing the function of the hippocampus (e.g., Cabeza & St Jacques, 2007).

In a behavioral study of hippocampal function (to our knowledge, the only study on mnemonic strategies that includes autobiographical memories), Bouffard et al. (2017) compared an autobiographical, so-called "temporal" scaffold, consisting of a timeline of autobiographical events to the Method of Loci. Participants were instructed to create a chronological timeline using ten of their most memorable memories (Bouffard et al., 2017). Their participants' final recall performance showed a similar memory increase for the Method of Loci and autobiographical timelines, suggesting that spatial locations as in the Method of Loci and temporally ordered events can be used to enhance memory performance in a similar way (Bouffard et al., 2017). In our experiment, we developed a novel autobiographical technique, in which single autobiographical events *per se* comprise the mnemonic scaffold, which is more in line with how participants might spontaneously remember events from their lives.[2] Taken together, due to the highly self-relevant nature of autobiographical prior knowledge, we expect that autobiographical stories can serve as effective mnemonic scaffolds.

### The Routine Routine Activity Scaffold

It has been proposed in Script Theory that routine activities facilitate memory, as part of our knowledge is organized around stereotypical situations (Bartlett, 1932; Schank & Abelson, 1977; Abelson, 1981). Dynamic versions of schemata comprising activities, scripts are defined as organized knowledge stores which consist of routine activities and serve as a base for elaborations surrounding a topic (Bower, 1970). Considering the large body of literature on routine activities and knowledge acquisition via prior knowledge in the form of schemata and scripts, we were surprised that we could not find any studies that used routine activities for mnemonic purposes, with the exception of Bouffard et al. (2017). Their free recall task, however, only included one routine activity, the steps to making a sandwich, to investigate whether sequences with easily accessible temporal features provide similar memory boosts as the Method of Loci and timelines of autobiographical events (as described above). The steps of making a sandwich resulted in similar memory performance as autobiographical timelines and the Method of Loci (Bouffard et al., 2017). Based on these findings and the notion that routine activities are well rehearsed and highly familiar, we expected routine activities to be effective mnemonic scaffolds.

[2] Our Story Scaffold Method, where study items are integrated into an autobiographical story from the learners' own life is not to be confused with the story mnemonic described in reviews by Bellezza (1983, 1986) and Worthen and Hunt (2008, 2011), where word lists are studied by combining the words in sentences that make up an ad-hoc story (and typically not autobiographical).

### 4.1.3 Individual differences in learner aptitude and the usefulness of mnemonic scaffolds

Viewing mnemonic strategies as skills (Ericsson et al., 1980), one might expect individual differences might determine how well a participant can excel with a particular scaffold. Despite its importance for the application of mnemonic strategies in educational and cognitive rehabilitation settings, the role of individual differences in the usefulness of mnemonic scaffolds has received almost no scientific attention. One notable exception, Sanchez (2019) found evidence suggestive that effective usage of the Method of Loci was dependent on participants' visuospatial ability, measured with the Cube Comparisons Task and the Paper Folding Task (PFT; French et al., 1963). In fact, those lower in visuospatial aptitudes may actually have been disadvantaged by using navigational scaffolds for serial recall (Sanchez, 2019). In an attempt to replicate Sanchez' (2019) findings, we used the PFT to measure visuospatial aptitude. Aside from Sanchez' (2019) study, this is the first study investigating effects of individual differences in the usefulness of the Method of Loci in addition to non-navigational mnemonic scaffolds.

There is some anecdotal evidence that visual imagery skill or vividness might determine the effectiveness of some mnemonic scaffolds. Many memory athletes contend that "thinking in images," i.e., vivid visual imagery, is key to the successful application of mnemonic strategies (e.g., Foer, 2011; Konrad, 2013). This introspection of world-class mnemonic strategy users, however, has not been confirmed by research, and at least two studies that addressed this question found no relationship between vividness of visual imagery and success with the Method of Loci (Kliegl et al., 1990 and McKellar, Marks and Barron reported by Marks, 1972b). Notably, visual imagery capacity appears to vary greatly between individuals, with aphantasics reporting no ability to create visual images (Keogh & Pearson, 2018). The relationship between individual differences in visual imagery and usefulness of visual-based mnemonic strategies might therefore have important practical and theoretical implications. However, the fact that congenitally blind participants can perform well with the Method of Loci (De Beni & Cornoldi, 1985) suggests that visual imagery may not be the basic reason why imagery-based strategies are effective. We used the Vividness of Visual Imagery Questionnaire (VVIQ, Marks, 1973) to assess the self-reported vividness of participants' imagery.

In addition to the VVIQ and PFT, we used the Body Responsiveness Questionnaire (BRQ, Daubenmier, 2005) to assess participants' awareness of internal body sensations expecting it to relate to success with the Body Scaffold.

### 4.1.4 Goals and hypotheses of the current experiments

With two experiments, we address three main questions: 1) Do mnemonic scaffolds differ in their effectiveness in facilitating serial recall?, 2) Do visuospatial ability, vividness of visual imagery, and body responsiveness affect the usefulness of those mnemonic scaffolds?, and 3) Does bodily engagement contribute to the success of the Body Scaffold?

In Experiment 1, participants were assigned to one of four Mnemonic Scaffold Groups instructed to generate mnemonic scaffolds using either a) the Method of Loci (Loci Scaffold), b) a sequence of body parts (Body Scaffold), c) autobiographical stories (Story Scaffold), or d) routine activities (Routine Activity Scaffold). After generating their own scaffold, participants were instructed to use it to study lists of ten words by making associations between the words and scaffold. Serial recall (recalling the list in order) accuracy was compared to a Control Group that used a read-aloud strategy in place of a mnemonic scaffold. In light of theories and empirical findings on the prior knowledge effect, we hypothesized that the four mnemonic scaffolds should outperform the control condition. We reasoned that serial and scaffold-cued (recalling the corresponding word when cued with its anchor from the scaffold) recall may differ if participants are not applying the respective mnemonic scaffold as instructed. In addition, some participants might remember the scaffold–word associations but be unable to use them to perform serial recall because they might rely on the cues separately from the scaffold to remember the associations, for example.

As for the individual differences measures, we predicted a positive correlation of VVIQ scores and effectiveness of all mnemonic scaffolds relative to Control, given that all scaffolds are visually rich and using vivid mental imagery is often mentioned as practical, anecdotal advice by professional memory athletes. For the PFT, we predicted a relative benefit of the Loci Scaffold for participants with high PFT scores, as we expected to replicate Sanchez' (2019) findings. Finally, we predicted a relative benefit of the Body Scaffold for participants with high BRQ scores, if the Body Scaffold benefits from embodiment.

In Experiment 2, participants were instructed in one of three variants of the Body Scaffold with varying levels of bodily engagement, involving either a) no physical engagement of the body parts using the same instructions as in the Body Scaffold Group of Experiment 1, b) repetitive hand movement, or c) touching the respective body parts during study. Hypothesizing that using one's own body as a scaffold may enable a deeper engagement with the list items during study, in which the coupling between sensorimotor perceptions or actions and the study items might consequently provide a memory benefit, we predicted that the group with the highest level of bodily engagement (c) would perform best.

## 4.2 Experiment 1

In Experiment 1, we compared four mnemonic scaffolds to a non-scaffold Control. We also tested whether success with any scaffold covaried with several individual-difference measures.

### 4.2.1 Method

**Participants**

Participants ($N = 221$) were recruited from the introductory psychology research participation pool in partial fulfilment of course requirements. There were 44, 45, 44, 43, 45 participants in the Body, Loci, Activity, Story, and Control Group, respectively. The mean age of the participants who reported their age (11 omissions) was 19.58 years. All participants were required to have English as their first language and/or have learnt English before the age of six. All participants were older than 17 years. Written informed consent was obtained prior to the experiments in accordance with the University of Alberta ethical review board.

The sample size was selected to be close to related studies on mnemonic scaffolds (Bouffard et al., 2017; Legge et al., 2012; Roediger, 1980) that observed significant effects of the Method of Loci and other scaffold-based strategies on learning. We also used G*power (Faul et al., 2007) as a post-hoc evaluation of the sample size. For a repeated measures, between factors calculation of the required sample size for main effects with an alpha error probability of 0.05, and an estimated medium effect size of 0.25, and power of 0.8 the required sample size is indicated as 110. We exceeded this required sample size. For a post-hoc evaluation of sample size for interaction effects ("ANOVA: Fixed, special, main effects, and interactions" in G*Power with Greenhouse-Geisser corrected degrees of freedom of 5.45), a sample size of 218 participants was indicated. Thus, our interaction effects of Serial Recall Accuracy

× Serial Position are slightly under-powered and must be interpreted with caution. This is why we followed up on interaction effects with single-measures ANOVAs at each serial position. In addition, we relied on Bayes Factors as an indication whether our sample size was sufficient.

### Materials

Study lists were random sets of ten 4–8 letter nouns of high and low imagery (e.g., MEADOW, DOUBLE, EFFORT, TIMBER) drawn from the Toronto Word Pool with frequency ratings by Kucera and Francis,[3] also used by Bouffard et al. (2017). Words were drawn at random, without replacement, to construct the complete set of serial lists, each comprising ten words. The experiment was presented in individual closed testing cubicles, each with a chair, table, and PC desktop computer.

### Procedures

Both Experiment 1 consisted of five phases (Figure 4.1), described in more detail below. Experiment 1 lasted no longer than one hour and fifty minutes and had five Groups, explained in detail below.

**Pre-instruction baseline serial recall** The task design is visualized in Figure 4.1. The first phase measured serial recall ability prior to instruction and was thus identical for all groups. Participants studied two lists of ten words and were tested with serial recall. Words were presented individually, centrally on the screen, self-paced. That is, a word remained on the screen until participants pressed ENTER. Following the study phase, participants were asked to type the words in the order they were presented. Each response was entered on a separate response line. All ten response lines were visible from the start of recall in a vertical configuration and were not numbered. Participants were not allowed to edit their responses apart from using the backspace key, before pressing ENTER to submit the response. Typed words remained on the screen until all ten responses were entered, but no backtracking was allowed. If participants could not remember a word, they were instructed to type PASS.

[3] http://memory.psych.upenn.edu/Word_Pools

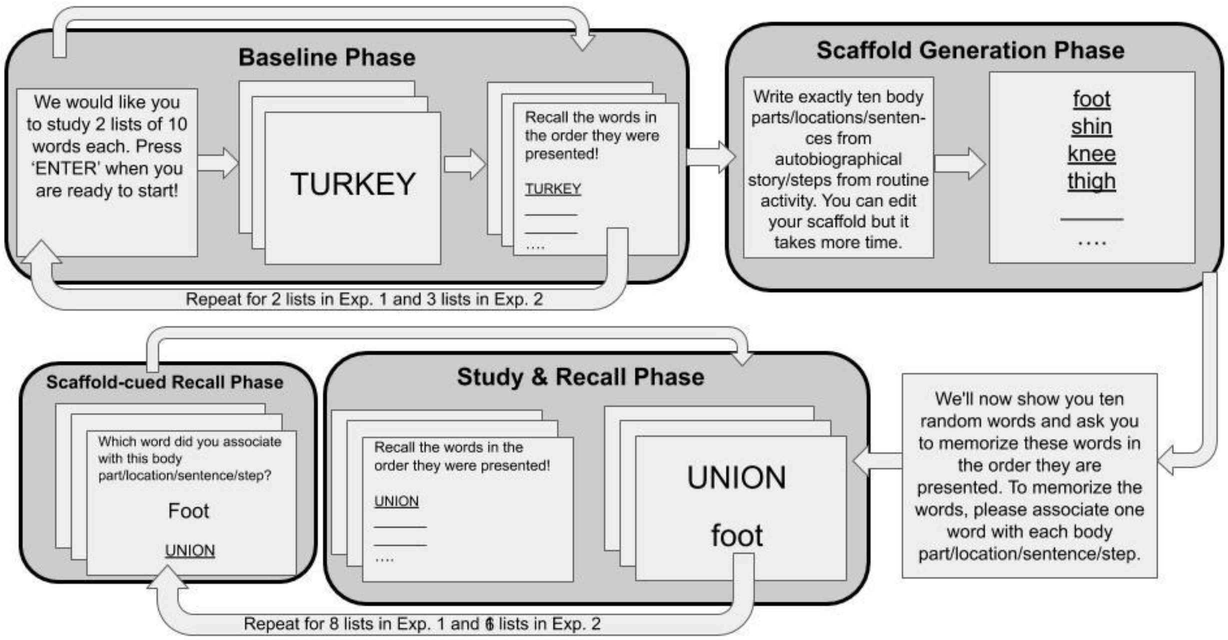



Figure 4.1: Illustration of the experimental design.

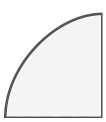


Figure 4.2: Illustration of the experimental groups in both experiments.

**Scaffold-generation phase** After the pre-instruction baseline memory test, participants in all groups, including Control, were informed they will learn a mnemonic technique to make remembering words lists easier. Participants in the four Mnemonic Scaffold Groups read the same instructions with the only variation between groups being the type of scaffold used.

Participants in the Body Group typed ten body parts. They were asked to start at their feet and follow their body upward to their head, to ensure a sequential order of the body parts selected.

Participants in the Loci Group typed ten locations or objects along a familiar, frequently travelled route; the example given was the way from their house to the university. They were informed that they could only type each location once and that it is important to follow the chronological order of locations in which they are encountered on the familiar route.

Participants in the Story Group typed an event from their own life they remembered well, split up into ten sentences; the example given was their first day of school. They were informed that it was important to follow a chronological order of events and to write their story in first-person perspective as if telling it to a friend.

Participants in the Routine Activity Group typed an activity performed on a daily basis, split up into ten steps; the examples given were: brushing your teeth, walking your dog, making a sandwich. They were informed that it was important to follow a chronological order of steps and to type the steps in the imperative tense, as if giving instructions to someone else.

After typing the ten parts of their scaffolds, participants were asked to proofread their scaffolds and were able to edit by repeatedly changing the order of the ten parts or re-writing a part, if they wished. Participants were informed that they will use their scaffolds to study eight more word lists by associating them with their body parts, locations, sentences from their autobiographical stories, or steps from their routine activities, respectively.

Numerous prior studies have failed to find any evidence of proactive interference using similar strategies (e.g., Bass & Oswald, 2014; Caplan et al., 2019; Legge et al., 2012; Massen & Vaterrodt-Plünnecke, 2006), so we did not expect to observe proactive interference here. Given that, as has been done in those previous studies, we asked each participant to memorize numerous lists, to obtain more reliable measures of their performance, and in case we could check for evidence of training effects within the experimental session.

Participants in the Control Group were asked to type a sequence of ten body parts as a filler instruction and did not receive any instructions on how to use it later. None of the participants reported having used their body parts to study the words when asked whether they used a different strategy than reading the words aloud. Participants in the Body, Loci, and Activity Group studied 10 lists in total; two lists in the pre-instruction baseline phase and eight lists using their scaffold or the saying words aloud (Control).

**Post-instruction study and serial recall** After typing their scaffolds, participants received instructions on how to use those to study eight more word lists. For the four Scaffold Groups, the study phase entailed associating one word with one part of the respective participant-generated scaffold. Word–scaffold pairs were presented in serial order with the list word in uppercase in the center of the screen and the part of the participant-generated scaffold below. Participants saw one word–scaffold pair on each screen and were instructed to press ENTER after they were satisfied that they had associated the word with the part of the scaffold to get to the next one. Participants in the Control Group received the filler instruction to read the words aloud to (supposedly) make remembering easier (Bodner & Taikh, 2012). Whether or not this enhances memory is, in fact, controversial (see the editorial for the special issue by Bodner & MacLeod, 2016). After participants in all Groups had studied a list with the respective method, they were asked to recall the list in serial order as in the baseline phase, without displaying the scaffold parts.

**Strategy verification, scaffold-cued recall** Instructions to apply a mnemonic strategy do not guarantee that participants use the strategy to study and recall list items (Bellezza, 1981). Previous studies have assumed compliance of participants (e.g., Roediger, 1980; Bouffard et al., 2017). This is problematic because self-reported compliance rates of using the instructed strategy cannot be expected to be particularly high (Sahadevan et al., 2021). For example, self-reported compliance rate in a study of the Method of Loci was only 40 and 58% in the two strategy groups respectively (Legge et al., 2012). Thus, in addition to concerns with the validity of subjective report of instruction compliance, we were concerned that our comparison of the effectiveness of our four mnemonic scaffolds could be confounded by including participants who do not apply the strategy as instructed. We therefore included the scaffold-cued recall task, where participants were tested directly for memory for scaffold–word associations (regardless of whether or not these had just supported their serial recall). This gives us the unique opportunity to check whether participants were actually using the scaffold strategies as instructed, and actually forming scaffold–word associations, and secondly, whether success in scaffold–word memory, itself, might largely explain the differences in serial recall success across scaffolds (for a related tasks, see Bellezza, 1984a and Sahadevan et al., 2021). As the Control Group did not use mnemonic scaffolds, the scaffold-cued recall task only applied to the Scaffold Groups. After each serial recall phase of each list, the parts of the scaffolds were displayed as cues in a new random order, in the center of the screen and participants were asked to type the word they had associated with the particular scaffold-cue. After scaffold-cued recall of each list, participants in the Scaffold Groups were informed how many lists remained to be studied; participants in the Control Group received this information after serial recall of each list.

**Individual-differences questionnaires** Three questionnaires were administered to test whether some variance in the effectiveness of particular scaffolds might be explained by potentially relevant individual differences. These questionnaires were always administered in the same order, as follows. To measure subjective vividness of visual imagery, we used the Vividness of Visual Imagery Questionnaire (VVIQ; Marks, 1973). Visual imagery is defined as a "combination of clarity and liveliness; [the] more vivid an image the closer it approximates the actual precept" (Marks, 1972b, p. 82). The questionnaire consists of four groups of four items. Participants are asked to consider the image formed in thinking about specific scenes (e.g., a sunset) and situations (e.g, encountering a friend). The vividness of the image is rated along a 5-point scale. While the VVIQ is still, to date, the most widely used tool to measure visual imagery, it has drawn some criticism regarding its validity (for a systematic overview see McKelvey, 1995).

The Paper Folding Task (PFT; French et al., 1963), consists of 20 problems that get progressively more difficult to solve. For each item, participants are asked to imagine the folding of a square piece of paper, with at least one hole punched through the paper at a given point. Participants select one of five displayed options illustrating how the paper would look after being unfolded. This task is considered to measure visuospatial aptitude and has been used to predict the usefulness of the Method of Loci (Sanchez, 2019).

The PFT was scored so that high PFT scores reflect high visuospatial aptitude.

Body responsiveness, defined as the "the tendency to integrate body sensations into conscious awareness to guide decision making and behavior and not suppress or react impulsively to them" (Daubenmier et al., 2013, p.9 ) was assessed by the 7-item, Body Responsiveness Questionnaire (BRQ; Daubenmier, 2005). The BRQ was scored so that high scores reflect high self-reported vividness of visual imagery. We suspected BRQ scores could drive successful application of the Body Scaffold. The BRQ was scored so that high BRQ scores reflect high body responsiveness.

### Data analyses

**Statistical analysis** Statistical analyses were conducted in JASP (JASP Team, 2019) using simple linear regressions or analyses of variance (ANOVA) whenever comparing means of two or more independent groups of data, and analysis of covariance (ANCOVA), to test main and interaction effects of categorical variables, controlling for the effects of selected variables, which co-vary with the dependent variables. We call the main factor "Group" when the Control is included, and "Scaffold" when the Scaffold Groups are compared without the Control. The Greenhouse-Geisser correction was applied where sphericity was violated. When conducting post-hoc tests on significant group effects, post-hoc Tukey's Honest Significant Difference tests were used. We also conducted Bayesian ANOVAs and Bayesian linear regressions, which produce a Bayes factor. Bayesian model comparison assesses support for one model over another, in contrast to classical hypothesis testing, which seeks for evidence against only one model (the null hypothesis). The Bayes factor is the ratio of Bayesian probabilities for the alternative and null hypotheses; $\mathrm{BF}_{10} = \mathrm{p(H1)/p(H0)}$. By convention (Raftery & Kass, 1995), there is "some" evidence for the null when $\mathrm{BF} < 0.3$, and correspondingly, "some" evidence for the alternate hypothesis when $\mathrm{BF} > 3$. "Strong" evidence is inferred when $\mathrm{BF} < 0.1$ or $> 10$. For ANOVAs, we report BFs for including the effect of the model, and for t-test and linear regressions, we report BFs excluding the effect from the model. If Bayes Factors of linear regressions were in the inconclusive zone between 0.3 and 3, we followed up with pairwise Bayesian correlations. $BF_{+0}$ tests the constrained hypothesis that the correlation is positive-only, and $BF_{-0}$ tests the constrained hypothesis that the correlation is negative-only. When serial position curves are plotted, error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994)

**Serial recall scoring** Serial recall accuracy was scored in two ways: a) strict scoring for order memory, in which a word was correct if it was recalled in the position it was presented, sensitive to order-errors, and b) lenient scoring for item memory, in which a word was scored as correct if it came from the current list, regardless of order. Given that mnemonic strategies which require forming associations between existing memories and verbal study items are especially superior in aiding recall in the exact order that items were presented (Roediger, 1980; Bouffard et al., 2017; Yates, 1966), we focus on measures of memory accuracy based on a strict scoring criterion. We conducted the same analyses we report for strict scoring for lenient scoring to investigate the effects of mnemonic scaffolds on memory for items regardless of order. With lenient scoring, many analyses were non-conclusive with $p > 0.05$, and 0.3¿BF¡3, and therefore we primarily report lenient analyses that fall in the conclusive range.

### 4.2.2 Results and Discussion

First, we verified the absence of a subject sampling bias across groups (Table 4.5 and supplementary materials). We also verified the absence of a learning-to-learn effect, suggesting that accuracy does not increase from the first to the last half of the session simply due to practice effects (Table 4.5 and supplementary materials). Consequently, our central analyses of serial recall will focus on post-instruction accuracy. Interestingly, the Control group showed no evidence of a recency effect (Figure 4.3e). Recency effects are often absent in serial recall, particularly for visual presentation (e.g., Drewnowski & Murdock, 1980). This makes it particularly interesting that the advantages of the scaffolds primarily occurred at late list-positions, which we test in the next sections and revisit in the Discussion.

#### Effect of Scaffold: Comparison of pre- and post-instruction recall accuracy within groups

To test whether there was a memory benefit provided by any Scaffold, we compared pre-instruction to post-instruction recall accuracy (Table 4.5 and Figure 4.3a-f) within each group using paired-samples t-tests for both strict scoring (order memory) and lenient scoring (item memory), collapsing across serial position.

For strict scoring, there was a significant memory improvement for the Body, Loci and Story Groups (Figure 4.3a–c, respectively), but not the Activity Group, with the Bayes factor

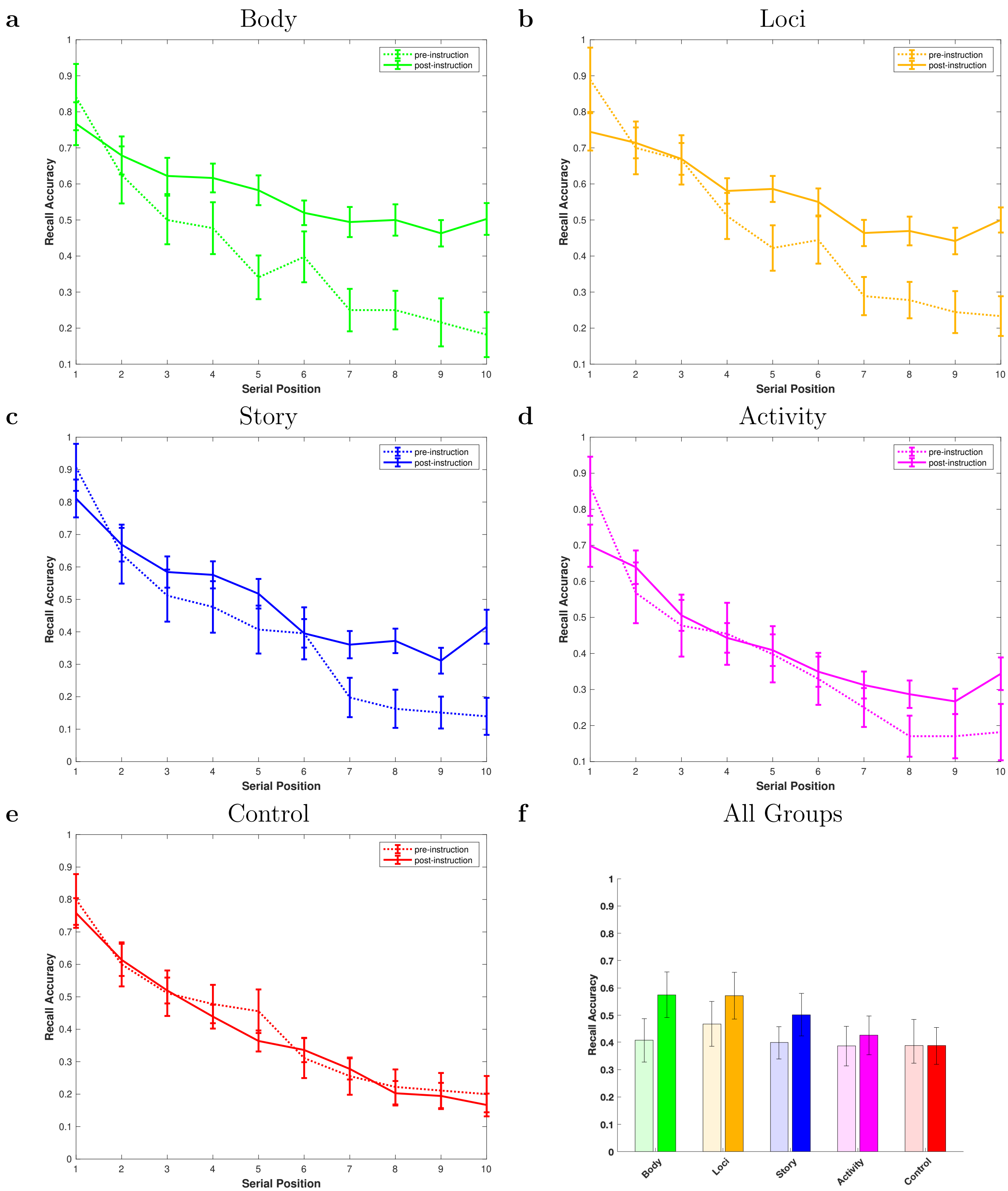



Figure 4.3: Experiment 1: Serial Position Curves for individual groups and bar graph for all groups comparing pre- and post-instruction serial recall accuracy (strict scoring). Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994).

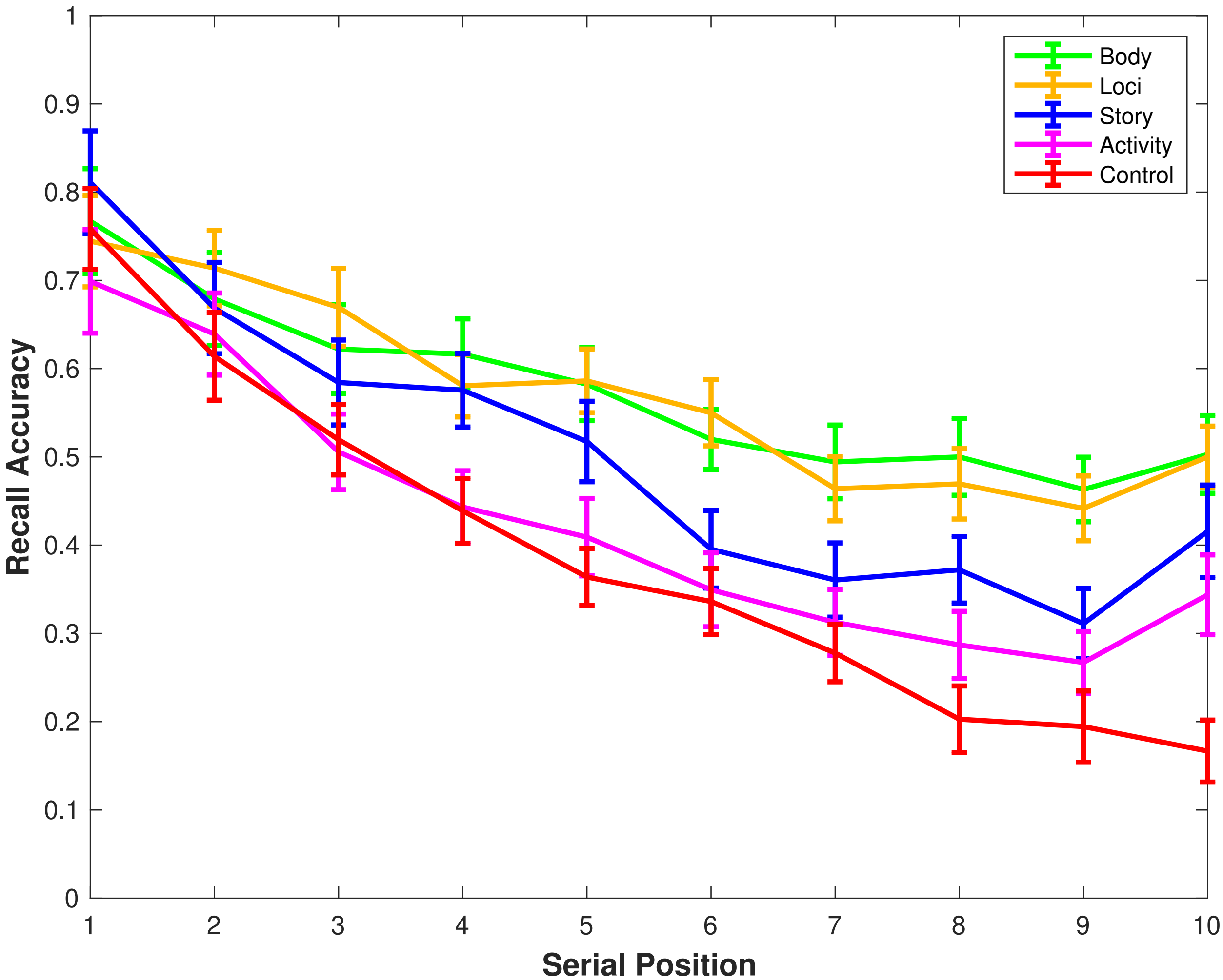



Figure 4.4: Experiment 1: Post-instruction accuracy as a function of serial position (strict scoring), as a function of group. Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994).

providing evidence for a null effect (Figure 4.3d). With lenient scoring, paired-samples t-tests did not confirm memory improvements, and post-instruction recall accuracy was significantly lower than pre-instruction recall accuracy when using the Routine Activity Scaffold. In sum, memory benefits were scaffold-specific and clearly benefited order memory, with no evidence of an effect on memory for items regardless of order.

### Comparison of the effectiveness of the mnemonic scaffolds between groups

To evaluate whether participants were successful in forming word–scaffold associations (as instructed) we compared the proportion of correctly reported word–scaffold associations from the scaffold-cued recall test, across the four Scaffolds (Figure 4.9) As reported in the supplementary materials, scaffold-cued recall accuracy was significantly higher in both the Body Group and the Loci Group than in the Routine Activity Group, and the effect of Scaffold was not modulated by Serial Position, suggesting that participants in the Body and Loci Group were more successful in associating the study words with parts of their scaffold than participants in the Routine Activity Group. In addition, we were interested in whether scaffold–word associations may have relied upon serial recall. A one–way ANCOVA with scaffold-cued recall accuracy as a covariate reported in supplementary material section (cautiously) revealed a non-significant main effect of Group after controlling for scaffold-cued recall accuracy, suggesting that scaffold–word associations may, indeed, have been relied upon the scaffolds during serial recall, itself.

We were interested in whether the effectiveness of the Body Scaffold, Loci Scaffold, and Story Scaffold in facilitating order memory observed in the within-subject, pre-post comparison is also observed when comparing the Scaffolds between subjects. To ask whether the Scaffolds differed from one another and from Control as a function of Serial Position, we conducted a 5 (Body, Loci, Story, Activity, Control) $\times$ 10 (Serial Position 1–10) mixed ANOVA on strict- (Figures 4.4) and lenient- (Figures 4.12) scored post-instruction recall accuracy (Table 4.6). This revealed a significant effect of Group and a significant Group $\times$ Serial Position interaction with strict scoring. Tukey's post-hoc tests revealed that accuracy of the Body and Loci Group was significantly greater than Control. The advantage of the Body Group over the Routine Activity Group was almost significant. To follow up on the significant interaction of Scaffold $\times$ Serial Position, we conducted one-way ANOVAs on Scaffold at each Serial Position (Table 4.6). This indicates that the significant interaction effect of Group $\times$ Serial Position with strict scoring is mainly characterized by the advantage of the Body and Loci Group over the Control in several positions.

The analyses for lenient scoring (Table 4.6, and see supplementary materials), found no significant difference between pre- and post-instruction accuracy, nor a reliable overall advantage for any Scaffold over Control, but at particular serial positions, the Body and Loci scaffolds were superior to Control. In the Discussion, we revisit the interesting pattern that the conditions tend to differentiate more at later than at earlier serial positions.

### Effects of individual differences

Our final objectives were to test whether individual-difference measures could explain subject variability in the effectiveness of the four Scaffolds. Our specific hypotheses were that a) higher VVIQ scores would correlate with higher serial recall in all Scaffolds, as they are all visually rich (but see the null correlation with Method of Loci performance reported by Kliegl et al. (1990) and by McKellar, Marks and Barron (reported by Marks, 1972b); b) higher PFT scores would correlate with higher serial recall accuracy in the Loci Group (a conceptual replication of the report by Sanchez, 2019), and c) higher BRQ scores would correlate with higher serial recall accuracy in the Body Group, if embodiment, in part, underlies the effectiveness of this scaffold. First, we verified the absence of a sampling bias for all three individual differences measures as reported in the Supplementary Materials. Second, to test our planned comparisons, we conducted linear and Bayesian linear regressions to predict post-instruction recall accuracy for order memory based on the individual differences measures. Further follow-up, exploratory analyses are described below.

**VVIQ** No significant correlations were found in our planned comparisons in the individual Scaffold Groups. When combining all Scaffolds the correlation approached significance (Table 4.1). In all Scaffold Groups except for the Routine Activity Group and when combining all Scaffolds, the correlation was nominally in the opposite direction of our hypothesis, with lower VVIQ scores (less vivid) predicting higher recall accuracy. Since the corresponding Bayes Factors were in the inconclusive zone, we followed this up with pairwise Bayesian correlation pairs. For all Groups, $BF_{+0}$ (i.e., the constrained hypothesis that the correlation is positive-only) provides evidence that the correlation is non-positive ($BF_{+0} < 0.3$), challenging the hypothesis that better visualization vividness translates into better performance with this strategy.

**PFT** To test our planned comparison whether high PFT scores predict high recall accuracy when using the Loci Scaffold, we conducted linear and Bayesian linear regressions to predict post-instruction recall accuracy (strict scoring) based on PFT scores in the Loci Group (Table 4.2 and Figure 4.5). The regression was significant ($p < 0.05$). Since the Bayes factor was in the inconclusive range, we followed this up with pairwise Bayesian correlations. $BF_{+0}$, which provided evidence that the correlation is positive ($BF_{+0} > 3$). This suggests

| *Group* | $R^2$ | $\beta$ | $BF_{10}$ | $BF_{+0}$ | $BF_{-0}$ |
|---|---|---|---|---|---|
| **Body** | **0.07** | **-0.26** | **1.03** | **0.07** | **1.52** |
| **Loci** | **0.01** | **-0.12** | **0.38** | **0.11** | **0.37** |
| **Story** | **0.07** | **-0.26** | **0.97** | **0.08** | **1.41** |
| **Activity** | **0.04** | **0.19** | **0.55** | **0.68** | **0.09** |
| **All Scaffolds** | **0.02** | **-0.14** | **0.87** | **0.03** | **1.07** |
| **Control** | 0.00 | 0.04 | 0.30 | 0.23 | 0.15 |

Table 4.1: Experiment 1: Linear regressions, Bayesian linear regressions and pairwise Bayesian correlations to predict post-instruction recall accuracy (strict scoring) based on VVIQ scores; $*p < .05$; $BF_{+0}$ assumes the constrained hypothesis that the correlation is positive-only, and $BF_{-0}$ assumes the constrained hypothesis that the correlation is negative-only. Planned comparisons are depicted in boldface, and all other results are exploratory and should be interpreted with caution.

that higher PFT scores are associated with higher recall accuracy.

Due to the preexisting positive relationship between PFT scores and pre-instruction recall accuracy (specified in the supplementary materials section), we conducted several exploratory, follow-up analyses, reported in Table 4.2. We were interested in whether the positive relationship between PFT scores and recall accuracy in the Loci Group is also observed in the remaining four groups. Linear and Bayesian linear regressions and pairwise Bayesian correlations revealed that higher PFT scores significantly ($p < 0.05$, $BF_{+0} > 3$) predicted higher recall accuracy in all groups including Control with exception of the Story Group. This indicates that PFT scores are not only correlated with higher post-instruction recall accuracy in the Loci Group, but also with higher post-instruction recall accuracy in general, including the no-scaffold (read-aloud), Control group. Taken together with the positive relationship between PFT scores and pre-instruction recall accuracy, this casts doubt on Sanchez' (2019) conclusion that participants with high visuospatial aptitude measured by the PFT benefit more from the Method of Loci than participants with low PFT scores.

In our next follow-up, exploratory analysis, we were interested in whether PFT scores may be a significant ($p < 0.05$, $BF_{10} < 3$, or $BF_{-0} > 3$) predictor of study times. Linear and Bayesian linear regressions and pairwise Bayesian correlations revealed that this was indeed the case for all Groups (Table 4.2, and Figure 4.5). If visual imagery skills were beneficial to using a scaffold, we would expect that higher PFT scores would be associated with higher accuracy and shorter response times because better visual imagery skills should enable the memorizer to form adequate images faster. Instead, we found a speed-accuracy

| *Group* | $R^2$ | $\beta$ | $BF_{10}$ | $BF_{+0}$ | $BF_{-0}$ |
|---|---|---|---|---|---|
| **Body** | | | | | |
| Post-Instruction | 0.10 | 0.32* | 1.87 | 3.08 | 0.06 |
| Study Time | 0.52 | 0.72* | >100 | 1.21 | 0.08 |
| **Loci** | | | | | |
| Post-Instruction | **0.11** | **0.33*** | **2.35** | **4.00** | **0.06** |
| Study Time | 0.16 | 0.40* | 6.55 | 3.22 | 0.06 |
| **Story** | | | | | |
| Post-Instruction | 0.07 | 0.26 | 0.98 | 1.35 | 0.08 |
| Study Time | 0.23 | 0.48* | 27.51 | 0.48 | 0.10 |
| **Activity** | | | | | |
| Post-Instruction | 0.17 | 0.40* | 7.07 | 13.92 | 0.05 |
| Study Time | 0.47 | 0.68* | >100 | 4.58 | 0.06 |
| **Control** | | | | | |
| Post-Instruction | 0.29 | 0.53* | >100 | >100 | 0.04 |
| Study Time | 0.48 | 0.69* | >100 | >100 | 0.04 |

Table 4.2: Experiment 1: Linear regressions, Bayesian linear regressions and pairwise Bayesian correlations to predict post-instruction recall accuracy (strict scoring) and study time based on PFT scores; $*p < .05$; $BF_{+0}$ tests the constrained hypothesis that the correlation is positive-only, and $BF_{-0}$ assumes the constrained hypothesis that the correlation is negative-only. Planned comparisons are depicted in boldface, and all other results are exploratory and should be interpreted with caution.

trade-off, where PFT score was associated with both greater accuracy and longer response times. More importantly, these relationships were not specific to the use of the scaffolds but also observed in pre-instruction accuracy and Control. This suggests that participants who put in more effort in solving the PFT task problems may also try harder to memorize the words. In other words, the PFT may largely be sensitive to motivation and engagement of the participants throughout the whole experiment.

**BRQ** Following our planned comparison, we conducted classical and Bayesian linear regressions to predict post-instruction recall accuracy (strict scoring) based on BRQ scores in the Body Group. The BRQ did not predict post-instruction recall accuracy in the Body Group ($p > 0.05$, $BF_{10} < 0.3$). This suggests that body responsiveness does not contribute to the success of the Body Scaffold.

### 4.2.3 Interim summary

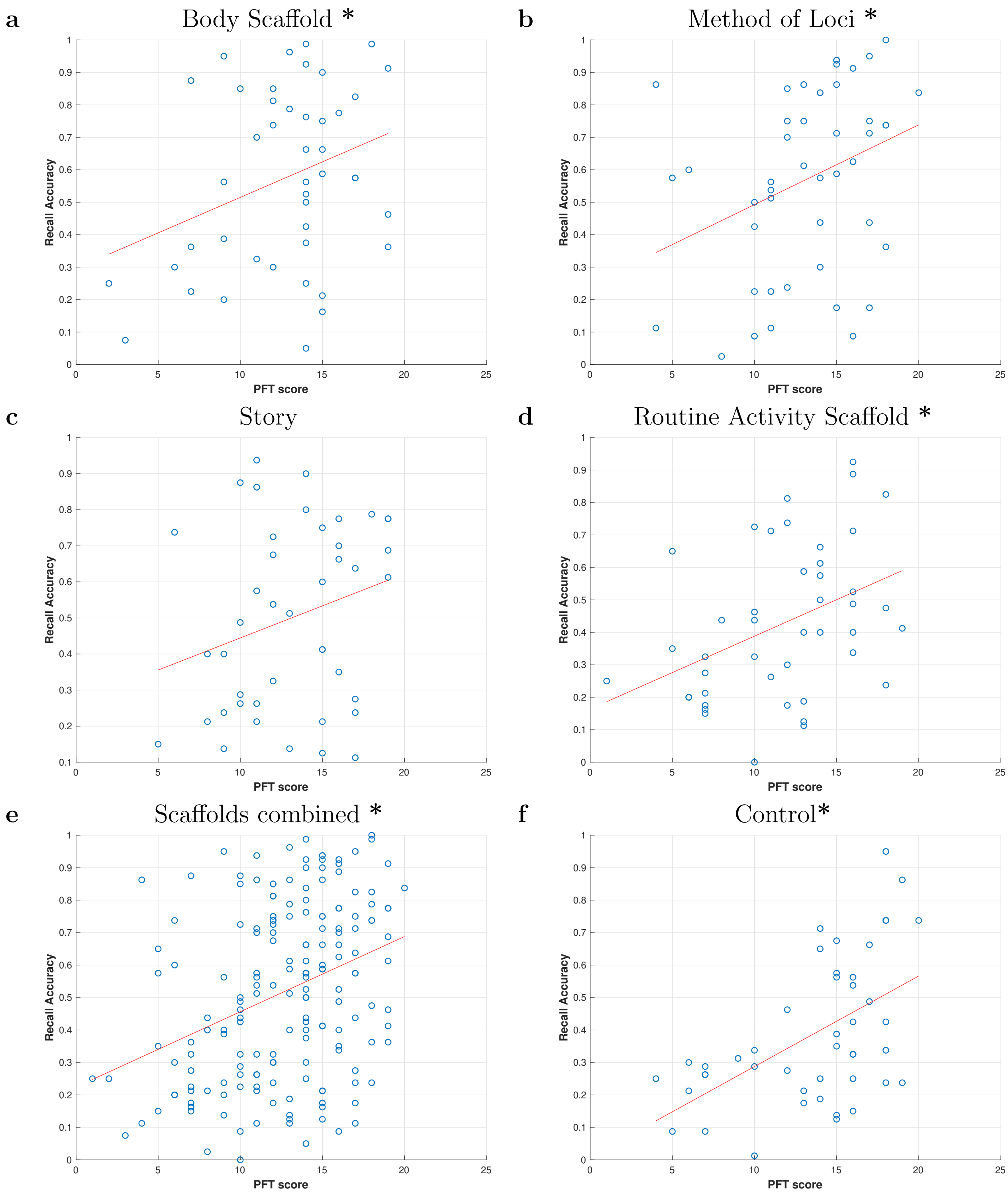



Figure 4.5: Experiment 1: Average serial recall by PFT scores. Regressions in groups marked with * are statistically significant ($p < 0.05$).

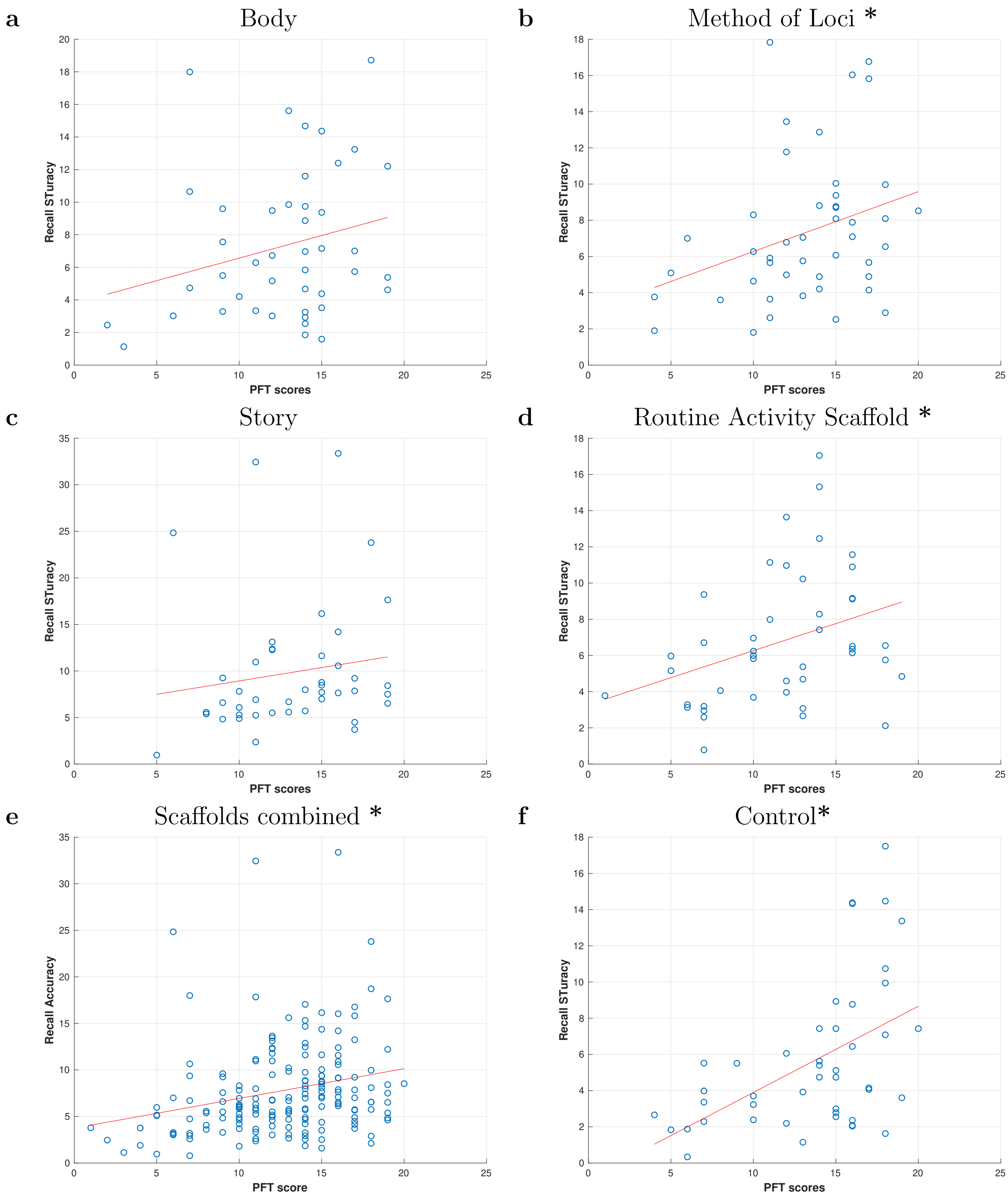



Figure 4.6: Experiment 1: Post-instruction study times by PFT scores. Regressions in groups marked with * are statistically significant.

Experiment 1 showed that some scaffolds are more effective than others. Surprisingly, the Body Scaffold was on par with the Method of Loci. The Autobiographical Story Scaffold provided a significant mnemonic benefit when compared to uninstructed baseline memory. The advantage of the Autobiographical Story Scaffold over the Control was not significant. The Routine Activity Scaffold did not provide a mnemonic benefit. The scaffold-cued recall data suggest that participants have relied upon the scaffolds during serial recall, itself. The mnemonic advantage of the Body, Loci, and Autobiographical Story Scaffold was present in predominantly recency positions.

## 4.3 Experiment 2

Surprised by the high level of success of the Body Scaffold, we wondered whether embodiment might contribute to its mnemonic benefit. To test that, we compared three variants of the Body Scaffold with different levels of attention drawn to the human body to a Control in Experiment 2. All methods were identical to those of Experiment 1 except where noted, as follows.

### 4.3.1 Method

**Participants**

Participants ($N = 147$) were recruited for Experiment 2. The recruitment process and requirements were the same as for Experiment 1. There were 44, 45, 44, 43 participants in the Sticker-on-Body Group, Sticker-on-Table Group, No-Sticker Group, and Control, respectively. The mean age of the participants (no omissions) was 19.03 years.

As in Experiment 1, the sample size was selected to be close to related studies on mnemonic scaffolds (Bouffard et al., 2017; Legge et al., 2012; Roediger, 1980). We used G*power (Faul et al., 2007) as a post-hoc justification of the sample size with the same parameters we used for Experiment 1. The required sample size for Experiment 2 is 100. We exceeded this required sample size and we also relied on Bayes Factors as an indication whether our sample size was sufficient. .

**Materials**

We used the same materials as for Experiment 1.

### Procedures

The basic paradigm was as in Experiment 1 (Figure 4.1). There were four groups: a Control Group, identical to the Control Group in Experiment 1, and three experimental groups that used different variants of the Body Scaffold with varying levels of bodily engagement. The purpose of the experimental groups was to control the level of embodiment in three variants of the Body Scaffold. The Sticker-on-Body Group had the highest level of bodily engagement, which involved physical engagement with the body parts. That is, participants in this group were prompted to touch the body parts from their typed scaffold during the study phase by attaching stickers to their body parts. The Sticker-on-Table Group had a lower level of bodily engagement, i.e., physical engagement of the hands, but no other part of the body. Participants in this Group were asked to repetitively attach the sticker on the edge of the table during the study phase. The No-Sticker Group did not involve physical engagement of the body and received the same instructions as the Body Group in Experiment 1. Because Experiment 2 had a shorter time requirement than Experiment 1 and lasted no longer than 50 minutes, it had fewer lists than Experiment 1; three lists in the pre-instruction baseline phase and six lists using a variant of the Body Scaffold or the saying words aloud (Control).

Pre-Instruction baseline serial recall, scaffold generation, encoding, serial recall, and scaffold-cued recall were as in Experiment 1. The instructions for the Control and No-Sticker groups were identical to the instructions for the Control and Body Scaffold groups of Experiment 1. Participants in all groups studied six lists of ten words using the strategy.

This was because we designed the experiment to be worth one participation credit, and had to last up to 50 minutes, shorter than Experiment 1.

Participants in the Sticker-on-Body Group were given a blank sticker. They were instructed to stick and remove the sticker to the respective body part whenever a new word-body part pair appeared on the screen while making the association between the word and the body part before they pressed ENTER to see the next pair. This motion of touching the body parts by attaching and removing the sticker was repeated for each word-body pair and served the purpose of prompting participants to physically engage their body parts during study and to evoke sensorimotor perceptions.

Participants in the Sticker-on-Table Group were also given a sticker. Instead of attaching the sticker to their body parts, they were asked to attach the sticker to and remove it from the edge of the table each time they studied a word-body part pair. We included this group as a Control condition to test whether repetitive hand motion that does not involve tactile perception on other body parts influences the success of the Body Scaffold in a different way than active engagement of the body parts as in the Sticker-on-Body Group.

After studying a list with the respective method, participants were asked to recall the list in serial order. As in Experiment 1, participants in the Body Strategy groups also completed a scaffold-cued recall task, cued with each self-generated body part in a random order and recalling the associated list-word.

As in Experiment 1, all participants completed the VVIQ (Marks, 1973), PFT (French et al., 1963) and BRQ (Daubenmier, 2005) in the same order. At the end of Experiment 2, we asked participants three self-report questions on strategy use: 1. Did you associate the list words with your body parts when studying them? (possible answers: always, sometimes, mostly, never) 2. If so, did connecting the words to parts of your body make remembering the words easier? (possible answers: always, sometimes, mostly, never) 3. Have you used this memorization technique before? (possible answers: yes/no).

### 4.3.2 Results and Discussion

The goal of Experiment 2 was to investigate further whether embodiment might be a driving factor behind the success of the Body Scaffold and whether additional attention drawn to one's body is associated with higher recall accuracy. Although the lack of a correlation of serial recall accuracy with the BRQ in Experiment 1 failed to support our initial prediction as tested through the lens of individual differences, the BRQ relies on self-report, and thus, might simply have lacked the sensitivity to individual differences in embodiment. Alternatively, embodiment might be an important factor completely apart from individual differences in overall impressions of one's body. Taking an orthogonal approach, we wondered if increasing the level of embodiment of the Body Scaffold procedure, itself, might improve serial recall performance, and thus reveal a positive influence of embodiment underlying the Body Scaffold. We first report a replication of the high recall accuracy of the Body Group in Experiment 1. Then, we follow the same order of analyses as in Experiment 1. Foreshadowing our results, we found supported null effects of strategy variant for most analyses, further challenging the idea that embodiment is a driving factor behind the success of the Body Scaffold.

### Replication of the high accuracy of the Body Group

As described earlier, the experiments had a common condition where participants were instructed to use the Body Scaffold without physically engaging their body. The instructions for both groups were identical; the only difference was the number of lists. In Experiment 1, two pre- and eight post-instruction lists were studied, and Experiment 2 had three pre- and six post-instruction lists because of different time requirements for the experiments. Although comparisons across experiments should be interpreted with caution, to compare whether recall accuracy of the Body Scaffold was consistent across the two experiments, we conducted a one-way 2 factor (Body Group of Experiment 1, No-Sticker Group of Experiment 2) ANOVA and Bayesian ANOVA on post-instruction recall accuracy for each serial position. This resulted in a non-significant effect of Group and a Bayes factors favouring a null effect for both post-instruction, $F(1, 80) = 0.17$, $p = 0.683$, $\eta_p^2 < 0.01$, $BF_{inclusion} = 0.25$) confirming that the success of the Body Scaffold is roughly consistent across the two experiments.

### Effect of Scaffold: Comparison of pre- and post-instruction recall accuracy within groups

To test whether there was a memory benefit provided by any variant of the Body Scaffold, we compared strict scoring pre-instruction to post-instruction recall accuracy within each group using paired-samples t-tests (Table 4.4, and Figure 4.7a-c). We averaged across serial position because we were interested in a mnemonic effect of the Scaffolds on the whole list. Under the strict scoring criterion this analysis revealed a significant memory benefit in all Body Scaffold Variants. With lenient scoring, the memory improvement approached significance in the Sticker-on-Body Group, and was non-significant for the remaining groups (Table 4.4).

### Comparison of the effectiveness of the Body Scaffold variants between groups

While reporting post-instruction serial recall data (strict scoring: Figure 4.8; lenient scoring: Figure 4.16), we ask whether the variants of the Body Scaffold differ in their effectiveness in facilitating serial recall accuracy. A one-way ANOVA and Bayesian ANOVA on Body Scaffold Variant (strict scoring) revealed that the Body Scaffold Variants did not differ from one another ($F(2, 107 = 0.44)$, $p = 0.644$, $\eta_p^2 = 0.01$, $BF_{inclusion} = 0.12$). When including the Control, a one-way ANOVA revealed a significant effect of Group (Table 4.4, Figure 4.8).

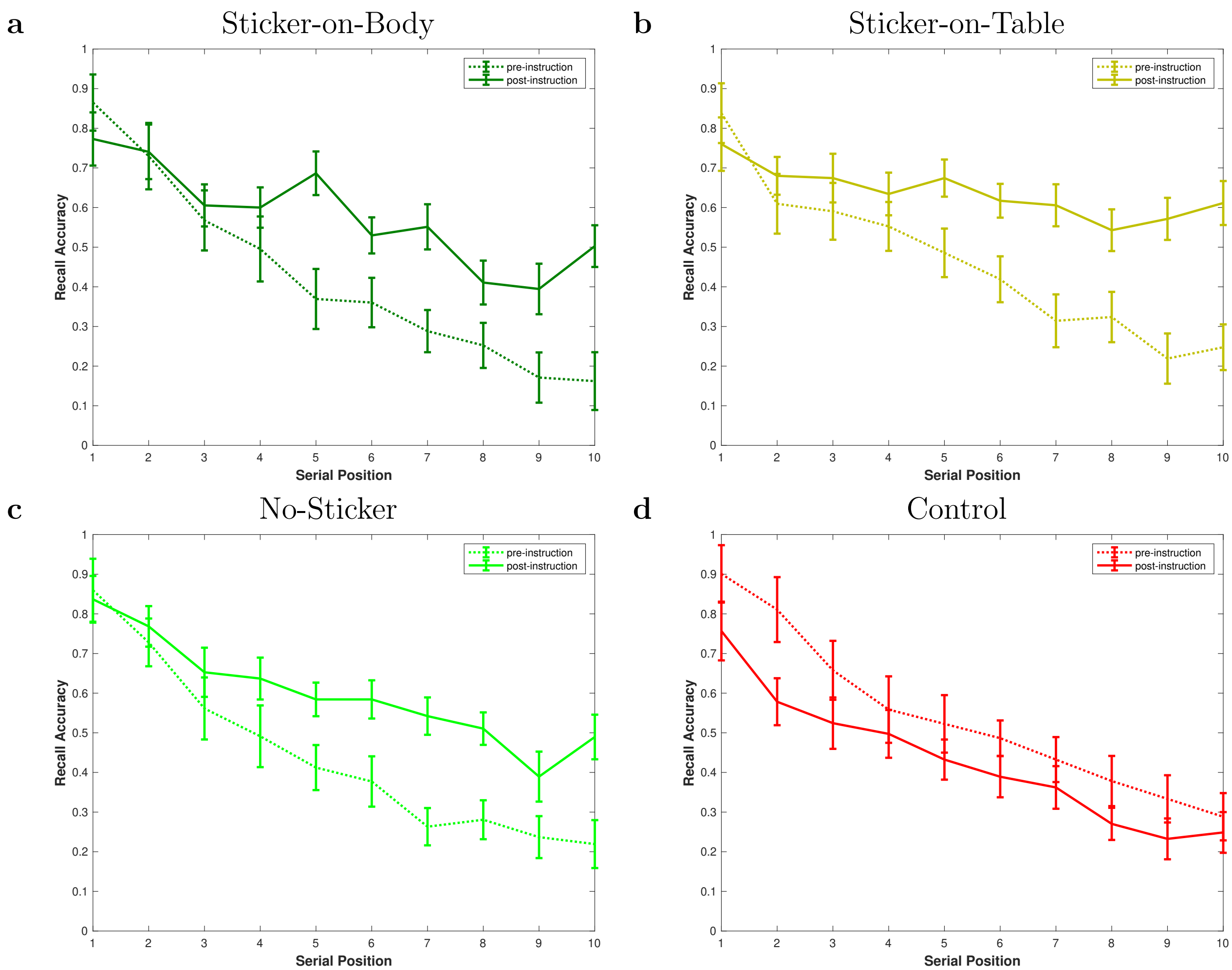



Figure 4.7: Experiment 2: Serial Position Curves comparing pre- and post-instruction recall accuracy (strict scoring). Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994)

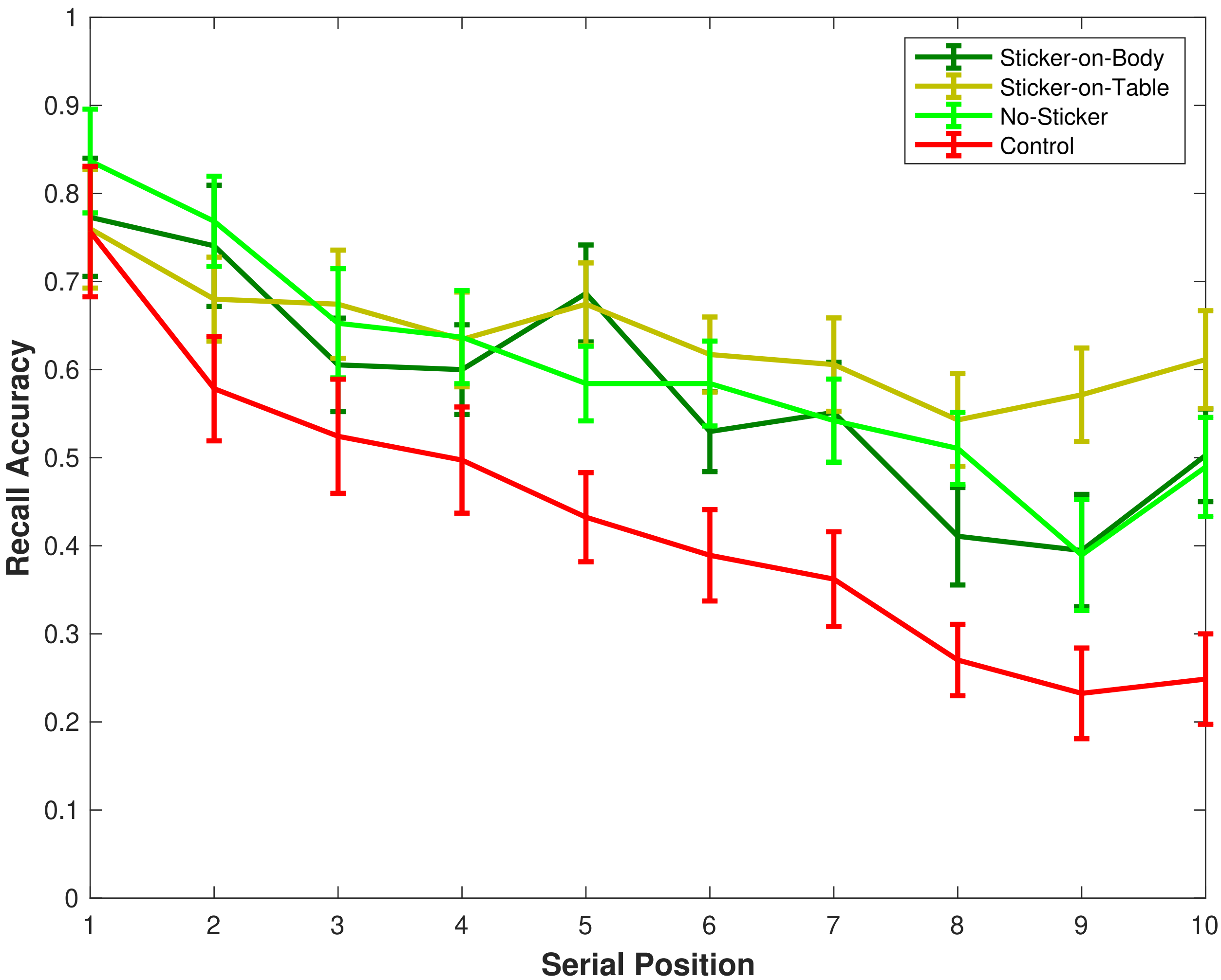



Figure 4.8: Experiment 2 - order memory (strict scoring): Post-instruction recall accuracy for each group. Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994).

Tukey's post-hoc tests revealed that serial recall accuracy of the Sticker-on-Table and No-Sticker Group was significantly higher than Control. Recall accuracy in the Sticker-on-Body-Group approached significance (Table 4.4). To ask whether Body Scaffold Variants differed from one another as a function of Serial Position, we conducted a 4 (Sticker-on-Body, Sticker-on-Table, No-Sticker, Control) $\times$ 10 (Serial Position 1–10) mixed ANOVA on strict scoring of post-instruction recall accuracy (Table 4.4). This revealed a significant interaction effect of Body Scaffold Variant $\times$ Serial Position, strongly supported by the Bayes Factor. To follow up on the significant interaction effect and the contradicting Bayes Factor of the main effect, we conducted one-way ANOVAs on Body Scaffold Variant at each Serial Position (Table 4.4). The Bayes Factor contradicting the non-significant $p$-value of the main effect of Body Scaffold Variant does not necessarily mean that there is no null effect. This is because the JASP algorithm requires the constituent main effects to be included in the model if the interaction is included. Hence, we conducted a one-way Bayesian ANOVA collapsing across serial position. This produced a Bayes factor of 0.13, providing strong evidence for the null, indicating that the variants of the Body Scaffold do not differ among themselves.

As in Experiment 1, we were interested in whether a different pattern could be observed with lenient scoring (Table 4.4, Figure 4.16). This analysis confirms the findings of Experiment 1 that for item memory regardless of order, the advantage of Scaffolds over Control is only observed towards the end of the list.

### Individual Differences

Since we did not have any a priori hypotheses regarding group differences in the relationship of recall accuracy and the individual differences measures, we averaged across Body Scaffold Variants and conducted classical and Bayesian linear regressions to predict post-instruction recall accuracy based on the scores in the individual differences questionnaires.

As in Experiment 1, we hypothesized a) higher VVIQ scores would correlate with higher serial recall accuracy when using the Body Scaffold due to its visual component and b) higher BRQ scores would correlate with higher serial recall accuracy when using the Body Scaffold due to the embodiment component. The PFT was used for follow-up exploratory analyses. First, we verified the absence of a sampling bias for all three individual differences measures and found a significant correlation between PFT scores and pre-instruction serial recall accuracy as specified in the supplementary material section. Then, we conducted classical and Bayesian linear regressions to predict post-instruction recall accuracy based on BRQ and VVIQ scores for our planned comparisons and post-instruction recall accuracy based on PFT scores for a follow-up, exploratory analysis.

**BRQ** Following our planned comparison, we conducted classical and Bayesian linear regressions to predict post-instruction recall accuracy (strict scoring) based on BRQ scores (self-reported body responsiveness) for all variants of the Body Scaffold combined illustrated in the supplementary materials. The BRQ did not predict post-instruction recall accuracy in the Body Scaffold Variants ($R^2 = 0.01$, $\beta = 0.03$, $p = 0.792$, $BF_{10} = 0.21$). This confirms the null effect of BRQ scores on the Body Scaffold found in Experiment 1 and provides evidence against our hypothesis that participants with high BRQ scores benefit more from using the Body Scaffold.

**VVIQ** To test our planned comparison, we conducted classical and Bayesian linear regressions to predict post-instruction recall accuracy (strict scoring) based on VVIQ scores (self-reported vividness of visual imagery) for all variants of the Body Scaffold combined, illustrated in Figure 4.18. This produced null effects, strongly supported by Bayes Factors for both post-instruction recall accuracy ($R^2 = 0.00$, $\beta = 0.07$, $p = 0.497$, $BF_{10} = 0.25$) suggesting that VVIQ scores have no bearing on effective use of the Body Scaffold. This add to the evidence that our hypothesis that high-imagery participants benefit more from mnemonic scaffolds than low-imagery participants can be rejected.

**PFT** Consistent with Experiment 1, heightened PFT scores significantly ($p < 0.05$) predicted heightened pre-instruction recall. As a further follow-up, exploratory analysis, classical and Bayesian linear regressions identified a positive relationship between PFT scores and post-instruction recall accuracy ($R^2 = 0.14$, $\beta = 0.37$, $p < 0.001$, $BF_{10} > 100$) for the variants of the Body Scaffolds combined, as illustrated in Figure 4.19. Since the Body Scaffold, in contrast to the Method of Loci, is thought to be independent of visuospatial cognitive ability, the positive relationship between PFT scores and recall accuracy in participants using the Body Scaffold further indicates that the PFT may primarily reflect effects due to task engagement and motivational factors rather than visuospatial ability.

## 4.4 General Discussion

Very little was previously known about how and even whether mnemonic scaffolds might differ in their effectiveness in supporting serial recall. Experiment 1 showed that some scaffolds are, in fact, clearly superior to others. Unexpectedly, the Body Scaffold was equally effective as the Method of Loci. The Autobiographical Story Scaffold provided a significant mnemonic benefit when compared to uninstructed baseline memory, but the advantage over the Control was not significant. The Routine Routine Activity Scaffold did not increase memory. With the strict scoring criterion, the mnemonic advantage of the Body, Loci, and Autobiographical Story Scaffold was observed for the whole list, while with the lenient scoring criterion, the mnemonic advantage of these was only present in some (predominantly recency) positions. Experiment 2 tested the hypothesis that the Body Scaffold might have been effective to the extent that it draws attention to the human body. However, a manipulation targeting embodiment failed to affect the effectiveness of the Body Scaffold, suggesting that attention drawn to the body is not a driving factor. Conceivably, differences across groups might have been due to participants in the Control group or some Scaffold conditions getting bored over the course of the session. However, arguing against this, when List Number was included in additional ANOVAs, its main effect and interactions were all non-significant (and pre- and post-instruction accuracy was virtually unchanged for the Control group in Experiment 1; Figure 4.5e). Finally, individual-differences analyses showed no reliable relationship between visual imagery skill and body responsiveness and the usefulness of our four mnemonic scaffolds.

### 4.4.1 The prior knowledge effect does not apply to all mnemonic scaffolds

The failure of the Routine Routine Activity Scaffold to exceed Control challenges our general hypothesis that all scaffolds generated from prior knowledge should support serial recall because they enable anchoring of new information with prior knowledge. In a sister paradigm, cued recall of verbal associations, Sahadevan et al. (2021) found that any benefit prior knowledge afforded by a peg list could still not raise performance to the level achieved by participants forming direct inter-item imagery (scaffold-free). This suggests that the level to which prior knowledge facilitates memory may be dependent on task-relevant characteristics of the scaffold. The fact that the Routine Routine Activity Scaffold did not show a mnemonic benefit seems surprising in light of Script Theory (Bartlett, 1932; Schank & Abelson, 1977; Abelson, 1981). This is because this theory suggests that that routine activities facilitate memory, as part of our knowledge and cognitive processes is organized around hundreds of stereotypical situations (Bower, 1970). Importantly, there is a key difference between routine activities facilitating memory in the context of Script Theory, and when used as mnemonic devices. In Script Theory, routine activities serve as a base for elaborations surrounding a topic (Bower, 1970), meaning that memory for related information on a certain topic is increased. This is quite the opposite of using routine activities as anchors for unrelated new knowledge. Thus, our findings are not at odds with Script Theory, but they indicate that routine activities do not enhance memory for *unrelated* information. This might be because routine activities consist of actions, which in contrast to locations, body parts or objects are dynamic, abstract, and difficult use as anchors.

In sum, the success of the other scaffolds suggests that a modified version of the anchoring hypothesis may still be tenable: If anchoring to prior knowledge, in itself, provides a benefit to scaffold-based strategies, these benefits may not apply to all scaffolds. The success of a particular scaffold might depend on particular characteristics of the scaffold, as we elaborate below.

### 4.4.2 Mnemonic scaffolds primarily affect memory for order, not items

Before we discuss the scaffolds individually, it is important to note that the mnemonic advantage of the Scaffolds over Control applied predominantly to memory for items in their presented order, which we investigated with the strict scoring criterion. This resonates with previous research that the strength of associative encoding techniques lies in facilitating memory for ordered information (Bouffard et al., 2017; Ericsson et al., 1980; Foer, 2011; Roediger, 1980; Yates, 1966). With the lenient scoring criterion used to investigate item memory, we only found a significant advantage of the Body and Loci Scaffold for some positions, with a tendency towards the end of the list. With regard to well-known serial position effects (recall accuracy is generally higher in early and late list positions, Lashley, 1951; Murdock, 1974), the equality of the Groups in primacy positions is likely due to a ceiling effect, and group differences are therefore more likely to materialize later in the list.

We had not anticipated the benefit of scaffolds being particularly strong late in the list, so we can only speculate as to the cause. Perhaps the scaffolds afford direct access to the study items so that when recall halts, the learner has a chance to pick them up by cueing with a later part of the scaffold. In other words, learners can skip previous anchors and access anchors later in the list to recall the study items. A caveat is that this is a post-hoc explanation. It is conceivable that the mnemonic benefit for item memory becomes more pronounced as the list progresses and more items are exempt from primacy effects. Thus, longer list lengths could reveal a more pronounced overall mnemonic benefit compared to Control or more pronounced group differences between the mnemonic scaffolds. [4] Ordered anchors that study items are associated with support memory in two ways. First, ordered anchors, even when order is not emphasized in the task, encourage learners to systematically recall an entire list without backtracking, skipping or repeating items. This is where the scaffold metaphor seems particularly fitting. Like in a climbing scaffold, learners make their way from one anchor to the next. Second, ordered anchors support serial recall or sometimes even direct access, through absolute positional retrieval. Additionally, ordered anchors may also provide relative coding. Absolute positional coding is best exemplified by the numerical rhyming peg list. In this method, a set of objects, each rhyming with a number (e.g., 1–BUN, 2–SHOE, 3–TREE, etc.) provides a scaffold, where the number peg–pairings allow immediately accessible direct correspondence to the number system (e.g., Bower, 1970; Lieberman, 2011). Relative coding is used in the Method of Loci, for example, when neighbouring words are associated with objects encountered in one subordinate unit, such as a room in a building or a street in a city. Rather than having to re-walk the whole route, learners, in theory, can access those units to retrieve study items according to their relative proximity.

While this warrants further research, it seems plausible that the Body and Loci Scaffolds have an advantage over the Autobiographical Story Scaffold as a result of internal order. This is because the order of anchors in the Body Scaffold, where the order of body parts can be retrieved by looking at one's body, and the Loci Scaffold, where the order can be retrieved by following a fixed route without backtracking, is more stable than the order of sentences in an autobiographical story as autobiographical events are not reliably retrieved in a chronological order (e.g., E. F. Loftus & Fathi, 1985).

[4] This was not observed in Chapter 3.

### 4.4.3 The Body Scaffold as an alternative to the Method of Loci

Our finding that the Method of Loci is on par with the Body Scaffold may converge with findings suggesting imagined navigation occurring during the use of the Method of Loci is epiphenomenal and may not be relevant for its memory benefit (Bouffard et al., 2017; Bower, 1970; Caplan et al., 2019; Carey, 2014). Thus, our findings may be at odds with the notion that visuospatial navigation and the engagement of the medial temporal lobe system are a determining factor in the memory benefit provided by this method, due to the dual role of this network in navigation and episodic memory (e.g., Rolls, 2017; Moser et al., 2015; Fellner et al., 2016). In fact, the visuospatial environment of the Method of Loci may be best viewed as a type of mnemonic scaffold that, in its mnemonic characteristics, does not differ from numerous other mnemonic scaffolds. The direct comparison between the Method of Loci and the Body Scaffold offers some insight as to what those "mnemonic core characteristics" of mnemonic scaffolds might be. The shared characteristic of the Body and Loci Scaffold is that they consist of single-unit anchors (body parts and objects/locations). Thus, each study item can be mapped onto a single-unit anchor. In contrast, the Story and Routine Activity Scaffolds comprise multi-word phrases, each offering multiple anchor points for association. Having to choose which anchor point to use may divert attention away from the study items and increase cognitive load. The advantage of the Body and Loci Scaffold over scaffolds consisting of sentences or phrases may therefore be due to direct association of each anchor with each study item. An alternative view is that imagined navigation occurs with both the Method of Loci and the Body Scaffold and explains the effectiveness of both (e.g., Rolls, 2017).

The anchors of the Body and Loci Scaffold further have in common that they consist of single-concept units. As Bellezza (1984) points out, body parts are constructible memory cues because the body forms an integrated and limited physical unit. This reasoning is in line with previous research comparing the Method of Loci to the numerical peg system, which found that both methods performed almost equally well (Roediger, 1980). As in the Method of Loci and the Body Scaffold, numerical peg systems have single-concept units as anchors. Together, the shared characteristics of the Body Scaffold, the Method of Loci, and the numerical peg system suggest that single-unit scaffolds are superior to multi-word scaffolds. The Body Scaffold and the Method of Loci are preferred over the numerical peg system for practical reasons; no pre-learned system is needed and, provided that the learner is not running out of body parts used as anchors, there are no constrains on the number of study items that can be memorized.

### 4.4.4 Autobiographical Stories as mnemonic scaffolds

The Autobiographical Story Scaffold was less effective than the Body Scaffold and the Method of Loci and did not yield a significant advantage over the Control. Yet, the mnemonic benefit was significant when compared to uninstructed baseline performance. This suggests that the Autobiographical Story Scaffold can be used to increase memory while further research and fine-tuning for the technique to be used successfully in applied settings is needed.

One disadvantage of using autobiographical stories as mnemonic scaffolds may be that recall of autobiographical narratives is not stable and based on the narrative and its circumstances (e.g., Greenberg & Rubin, 2003; Habermas, 2018; Hirst & Echterhoff, 2011; McAdams & McLean, 2013) and on individual abilities (Rubin, 2020, 2021). Therefore, using autobiographical stories that participants recall reliably rather than asking them to use any autobiographical event they might recall for the first time may increase the effectiveness of the technique.

Even if there is theoretical support for the idea that autobiographical material enhances memory (for a meta-analysis, see Symons & Johnson, 1997), it is unclear whether the mnemonic benefit we observed is due to the stories being autobiographical. Considering our general hypotheses that mnemonic scaffolds enhance memory because they allow for the anchoring of new information with prior knowledge it and given that the non-autobiographical Body and Loci Scaffolds were superior to the Story Scaffold when controlling for study time, it seems plausible that known fictional stories might be equally effective scaffolds as autobiographical ones.

Future studies could consider adapting the story scaffold for older participants. Though the Autobiographical Story Scaffold has not been tested with older adults, there are both motivational and neurocognitive factors suggesting that the Autobiographical Story Scaffold might be particularly well suited for older adults. Multiple studies, meta-analyses and review papers (Anschutz et al., 1987; Baltes & Kliegl, 1992; Gross & Rebok, 2011; Kliegl et al., 1990; Karbach & Verhaeghen, 2014; Yesavage et al., 1989) have shown that, in contrast to younger adults, older adults may not benefit from memory training using the Method of Loci in daily life. From a motivational perspective, we wonder if this could, in part, be because the navigational metaphor of the Method of Loci might induce a stereotype effect related to the fear of getting lost with increasing age (Levy, 2003). Autobiographical stories, in contrast, may increase motivation if older adults feel motivated by recalling stories from their own lives, possibly increasing self-efficacy and memory success with the Autobiographical Story Scaffold . From a neurocognitive perspective, the Autobiographical Story Scaffold may be better suited for older adults than the Method of Loci, which engages the hippocampus (e.g., Fellner et al., 2016) If hippocampal engagement is not the reason for the memory-enhancing effect of the Method of Loci, as explained above, invoking the hippocampus may be counterproductive. This is because the hippocampus is one of the first areas affected by age-related memory decline (Galton et al., 2001; den Heijer et al., 2010; Apostolova et al., 2006), while remote autobiographical memories are among the longest preserved in aging, even in age-related cognitive decline and early-stage Alzheimer's Disease because those memories are retrieved by the neocortex rather than the hippocampus (Cabeza & St Jacques, 2007). Given these factors, it might be promising to conduct further research on the Autobiographical Story Scaffold not only in younger adults, but particularly comparing it to the Method of Loci in older adults.

### 4.4.5 Attention drawn to the body does not contribute to the effectiveness of the Body Scaffold

Given the embodiment component of the Body Scaffold, we hypothesized that participants with high scores in the BRQ would benefit more from this method than participants with low BRQ scores. This was not confirmed by our analysis. BRQ scores did not predict any measure of scaffold-dependent recall accuracy in any group neither in Experiment 1 nor in Experiment 2. Similarly, as shown in Experiment 2, the Body Scaffold is robust to variations in bodily engagement. In other words, the Body Scaffold is effective regardless of whether learners sit still or engage their bodies. These results overlap with findings from a study on embodiment in a virtual environment in which participants performed a free recall and item recognition task recalling details of a virtual environment, which they explored either by a) marching in place and touching their forehead to turn while an avatar performed their movements to navigate a virtual environment, b) by initiating the movement performed by an avatar that was controlled by the experimenter, or c) standing still and watching the avatar controlled by the experimenter, and d) watching the environment pass by without an avatar on a head-mounted display (Tuena et al., 2017). The groups did not differ significantly in free recall nor item recognition, suggesting that different levels of embodiment do not affect recall in a virtual environment. However, the question of whether active movement during study enhances recall remains unresolved, as other virtual reality studies found a positive effect of active movement on recall, endorsing embodied cognition theories (Plancher et al., 2013; Jebara et al., 2014; Sauzéon et al., 2011). It is important to note that the similarity of those studies to ours is limited to embodiment in the sense of movement during encoding. In contrast to our experiment, previous studies did not involve explicit binding of study items and body parts.

A number of studies on embodied cognition have shown that verbal processing of concrete stimuli implicates the same sensorimotor neural correlates that are active during physical interaction with the object or entity itself (Zwaan, 2004; Pulvermüller, 2005; Barsalou, 2008; Fischer & Zwaan, 2008; Toni et al., 2008; Sakreida et al., 2013). Our results show that adding sensorimotor perception to verbal processing of the study items does not result in a mnemonic benefit. This can be seen as support for the view that simulation, the process of internally representing a verbal stimulus (Gallese, 2008; Zwaan, 2004), is inherent in embodied cognition, as additional sensorimotor perception is not required for a stimulus to be mentally processed. Therefore, strengthening the sensorimotor component of the Body Scaffold through active engagement of the body is not required for the same processes to be performed when the task is performed without active bodily engagement. It is important to note that there might be a trade-off effect of directing attention to the body parts and away from the study items shown on the screen. More time spent interacting with the body parts means less engagement with the study item itself, and while it is assumed that the study items is held in working memory while touching the body parts, we cannot ascertain this. If there was a hypothetical benefit of touching the body parts, it might be cancelled out by the fact that the body parts might have received proportionally more attention than the study items. In sum, viewing the Body Scaffold as similar to the Method of Loci, attention drawn to the body might be entirely irrelevant to the effectiveness of the Body Scaffold in much the same way that imagined navigation might be irrelevant to the effectiveness of the Method of Loci (Bower, 1970; Caplan et al., 2019).

### 4.4.6 Visual imagery aptitude does not contribute to the effectiveness of mnemonic scaffolds

In both experiments, individual differences did not differentiate levels of recall accuracy. This leaves open the possibility that imagined navigation, visual imagery and embodiment are all necessary, but at such a minimal level that the corresponding domain-skill or domain-affinity makes little difference. For novices, recall accuracy is not as high as for experienced memorizers. There might be a self-selection effect, where people are drawn to a scaffold-technique given their own skills and the ideal mnemonic scaffold may need to be customized to the skills and affinities of each individual.

The notion that the Method of Loci and the Body Scaffold share characteristics that underlie their equal effectiveness is corroborated by our findings that visual imagery skills are not responsible for the effectiveness of mnemonic scaffolds. According to common advice of memory athletes forming vivid images of the study items or "thinking in pictures" (e.g., Konrad, 2013) is key to successful application of mnemonic scaffolds (Foer, 2011; Konrad, 2013; Müller et al., 2018). During the Method of Loci, for example, memory athletes reportedly transform the to-be-remembered information into a vivid image which is then associated with one the loci of their familiar route.

We therefore hypothesized that participants who visualize objects and scenes in much detail are inclined to adopt this advice naturally, and therefore show higher post-instruction recall accuracy than participants who report low vividness in visual imagery. This, however, was not observed in our data. VVIQ scores did not predict serial recall accuracy. This resonates with two previous studies on the Method of Loci. McKellar Marks and Barron (1972, reported by Marks, 1972b) found that both high and low visualizers benefited from instructions in the Method of Loci, and VVIQ scores had no effect on memory improvement. A study by Kliegl et al. (1990) that investigated visual imagery skill and the effectiveness of the Method of Loci in older adults also found no relationship between VVIQ scores and effectiveness of the Method of Loci. Despite the fact that the VVIQ is widely used to assess self-reported vividness of visual imagery, including to verify self-diagnosis of aphantasia (e.g., Keogh & Pearson, 2018), its construct validity has been challenged (McKelvey, 1995), and future studies should use additional ways to measure individual differences in vividness of visual imagery.

In addition to the VVIQ, we administered the PFT as an objective measure of spatial visual imagery ability. One advantage is that it does not rely on self-report. Instead, it consists of problems that get progressively more difficult and whose solution likely requires spatial visualization. Sanchez (2019) found that PFT scores predicted participants' effective use of the Method of Loci. While our experiment is a follow-up of Sanchez' (2019) findings rather than a direct replication attempt, it is important to note some critical factors in which Sanchez' (2019) and our experiments differ with regard to measuring serial recall. First, Sanchez (2019) had timed (30 s study time), simultaneous presentation of the lists during study, while our experiment was self-paced and words were presented one at a time. Second, Sanchez' (2019) measure of recall accuracy was confounded by imagery of the study items, i.e., the difference between post- and pre-instruction accuracy, whereby the pre-instruction lists comprised exclusively low-imageability words and the post-instruction lists comprised exclusively moderate to high-imageability words. In contrast, all of our lists included high, moderate and low-imageability words. Third, Sanchez' (2019) participants were asked to recall the words in any order, while our participants were asked to recall the words in serial order. These differences in experimental design restrain us from viewing ours as a failure to replicate Sanchez' (2019) findings. However, we find it plausible that the shift from low- to high-imageability stimuli may have affected how participants with low versus high visuospatial imagery ability adopted the Method of Loci.

Critically, we found that PFT scores predicted both elevated pre-instruction recall accuracy and post-instruction recall accuracy, including Control. In addition, we found that PFT scores predicted longer study times in some Groups. We therefore suppose that PFT scores reflect motivational factors and compliance levels rather than a relationship between visuospatial aptitude and mnemonic benefit. Nonetheless, it is quite possible that visual imagery subjective experience or objective ability do, in fact, determine success with mnemonic scaffold strategies, but that the VVIQ and PFT are simply not sensitive to those most relevant aspects of visual imagery.

In addition to the absent correlation between visual imagery skills measured by the VVIQ and PFT and successful use of mnemonic scaffolds, findings that congenitally blind participants can perform well with the Method of Loci (De Beni & Cornoldi, 1985) corroborate the notion that the effectiveness of mnemonic scaffolds does not rely on forming vivid mental images. It is a compelling idea that forming mental images is superfluous to the effective use of mnemonic scaffolds in the same way that imagined navigation may be epiphenomenal, or at least not necessary for the mnemonic benefit of the Method of Loci (Bouffard et al., 2017; Bower, 1970; Caplan et al., 2019).

### 4.4.7 Methodological and theoretical contributions

We systematically compared three mnemonic scaffolds to the Method of Loci. Our findings challenge three widely-held conceptions about mnemonic techniques. First, we have shown that the Body Scaffold, a mnemonic scaffold that has previously been recommended by historical and modern memory training authorities (Gesualdo, 1592; Konrad, 2013) yet not empirically investigated, is equally effective as the famous Method of Loci. This not only calls for further studies and applications of the Body Scaffold in applied settings. It also highlights the need of broadening the narrow focus in the field of memory enhancement on the Method of Loci. To understand the mechanisms by which superior memory strategies operate, the field must move beyond the special status of the Method of Loci and fill in the research gap on mnemonic scaffolds that are equally effective and share underlying cognitive mechanisms.

Second, our finding that the Body Scaffold is equally effective as the Method of Loci provides evidence against the conception that the effectiveness of the Method of Loci is driven by imagined navigation (Fellner et al., 2016; Maguire et al., 2003). In line with accumulating evidence that imagined navigation (Bouffard et al., 2017; Bower, 1970; Caplan et al., 2019) may not be relevant for successful use of the Method of Loci, our findings suggest that future applications of the Method of Loci are unlikely to benefit from emphasizing the navigation metaphor.

Third, our findings challenge the common assumption that creating vivid visual images is key to successful use of mnemonic techniques (Foer, 2011; Konrad, 2013). We did not find evidence that individual differences in vividness of visual imagery predict the success with which participants use mnemonic scaffolds. This is not only relevant when teaching mnemonic techniques to novice learners, it also reveals how common advice from memory training authorities and empirical findings of factors that underlie successful use of mnemonic scaffolds diverge.

The scaffold-cued recall task as an objective measure to compare compliance between the mnemonic scaffolds is a methodological advancement from previous studies which have either assumed compliance or relied on self-report (Bouffard et al., 2017; Roediger, 1980). The scaffold-cued recall task provides the unique opportunity to check compliance (Sahadevan et al., 2021) and whether success in scaffold–word memory, itself, might largely explain the differences in successful use of mnemonic scaffolds (see also, Bellezza & Bower, 1981).

### 4.4.8 Practical Implications

Learners can dramatically improve their memory performance with mnemonic techniques (Dresler et al., 2017; Ericsson, 2003; Ericsson et al., 1980; Maguire et al., 2003). Yet, the memory-boosting potential of mnemonic techniques in educational settings and for memory-impaired individuals is vastly under-exploited. In educational settings with younger adults, many observational and quasi-experimental studies across several decades have produced favourable results (Cornoldi & De Beni, 1991; Dresler et al., 2017; Groninger, 1971; Lea, 1975; Maguire et al., 2003; McCabe, 2015; Ross & Lawrence, 1968). However, all of these studies have focused on the Method of Loci. In memory training settings for older adults, the Method of Loci has been shown to be unsuitable (Anschutz et al., 1987; Baltes & Kliegl, 1992; Gross & Rebok, 2011; Kliegl et al., 1990; Karbach & Verhaeghen, 2014; Yesavage et al., 1989), reinforcing the importance of alternatives to the Method of Loci.

Surprisingly, our findings suggest that the Body Scaffold may be a strong alternative to the Method of Loci. Experiment 1 and 2 have shown that the human body can facilitate learning in a similar way as a route through a familiar environment, and that no interaction with the body parts is needed for the Body Scaffold to be effective. This raises the interesting possibility that the Body Scaffold may even be better than the Method of Loci for people with poor (self-perceived) navigation skills and ageing populations in which the use of the Method of Loci who may experience an age-related decline in self-efficacy of navigation. Our individual differences analyses suggest that the Method of Loci, the Body Scaffold, and the Autobiographical Story Scaffold hold promise to facilitate memory in learners regardless of their individual visuospatial aptitude or body responsiveness. Finally, our findings suggest it would be fruitful to conduct further studies on the Body Scaffold and some fine-tuning of the Autobiographical Story Scaffold for learners of all ages.

| | All Groups | Sticker-on-Body | Sticker-on-Table | No-Sticker | Control | Body Scaffold Variants |
|---|---|---|---|---|---|---|
| **Strict** | Main Effect:<br>$F(3, 143) = 4.48$<br>$\eta_p^2 = 0.09$<br>$p = 0.005$<br>$BF_i = 6.70$ | Across Positions:<br>¿$Control$<br>0.15(0.06), $d = 0.61$<br>$p = 0.068$ | Across Positions:<br>¿$Control$<br>0.221(0.06), $d = 0.78$<br>$p = 0.005$ | Across Positions:<br>¿$Control$<br>0.17(0.06), $d = 0.65$<br>$p = 0.027$ | Across Positions:<br>see left columns | Main Effect:<br>$F(2, 107) = 0.44$<br>$\eta_p^2 = 0.01$<br>$p = 0.644$<br>$BF_i = 0.12$ |
| | Interaction Effect:<br>$F(18.91, 901.79) = 3.30$<br>$\eta_p^2 = 0.07$<br>$p$¡0.001<br>$BF_i$¿100 | Individual Positions:<br>¿$Control$<br>at P5, and 10 ($p$¡0.05) | Individual Positions:<br>¿$Control$<br>at P3, and $5 - 10$ ($p$¡0.05) | Individual Positions:<br>¿$Control$<br>at P 2, 8, and 10 ($p$¡0.05) | Individual Positions:<br>see left columns | Interaction Effect:<br>$F(12.91, 690.86) = 2.92$<br>$\eta_p^2 = 0.05$<br>$p$¡0.001 |
| **Lenient** | Main Effect:<br>$F(3, 143) = 2.09$<br>$\eta_p^2 = 0.04$<br>$p = 0.103$<br>$BF_i = 0.41$ | Across Positions:<br>no significant<br>differences | Across Positions:<br>no significant<br>differences | Across Positions:<br>no significant<br>differences | Across Positions:<br>see left columns<br>differences | Main Effect:<br>$F(2, 107) = 0.89$<br>$\eta_p^2 = 0.02$<br>$p = 0.413$<br>$BF_i = 0.23$ |
| | Interaction Effect:<br>$F(24.29, 1157.84) = 1.92$<br>$\eta_p^2 = 0.04$<br>$p = 0.005$<br>$BF_i$¿100 | Individual Positions:<br>no significant<br>differences | Individual Positions:<br>¿$Control$<br>at P 9 and 10 ($p$¡0.05) | Individual Positions:<br>¿$Control$<br>at P 8 ($p = 0.051$) | Individual Positions:<br>see left columns | Interaction Effect:<br>$F(15.66, 837.62) = 1.92$<br>$\eta_p^2 = 0.04$<br>$p = 0.012$<br>$BF_i = 0.50$ |

Table 4.3: Experiment 2: Comparison of the mnemonic scaffolds and Control. The median and standard deviation for pre- and post-instruction recall accuracy is reported as M(SD); $BF_i$ is short for $BF_{inclusion}$, P is short for position

| | Sticker-on-Body | | Sticker-on-Table | | No-Sticker | | Control | |
|---|---|---|---|---|---|---|---|---|
| | Pre | Post | Pre | Post | Pre | Post | Pre | Post |
| **Strict** | | | | | | | | |
| | 0.43(0.23) | 0.58(0.24) | 0.46(0.26) | 0.64(0.28) | 0.44(0.23) | 0.59(0.27) | 0.54(0.24) | 0.43(0.25) |
| | $t(36) = -4.49$ | | $t(34) = -3.69$ | | $t(37) = -4.73$ | | $t(36) = 3.10$ | |
| | $p$¡0.001 | | $p$¡0.001 | | $p$¡0.001 | | $p = 0.004$ | |
| | $BF_{10}$¿100 | | $BF_{10} = 39.83$ | | $BF_{10}$¿100 | | $BF_{10} = 9.79$ | |

Table 4.4: Experiment 2: Comparison of pre- versus post instruction recall accuracy for strict scoring. The median and standard deviation for pre- and post-instruction recall accuracy is reported as M(SD). With lenient scoring, the memory improvement approached significance in the Sticker-on-Body Group (p= 0.056), and was non-significant for the remaining groups.

| | Body | | Loci | | Story | | Activity | | Control | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Pre | Post | Pre | Post | Pre | Post | Pre | Post | Pre | Post |
| **Strict** | | | | | | | | | | |
| | 0.41(0.26) | 0.58(0.28) | 0.47(0.27) | 0.57(0.28) | 0.40(0.19) | 0.50(0.25) | 0.38(0.43) | 0.43(0.23) | 0.40(0.27) | 0.39(0.23) |
| | $t(43) = 3.69$ | | $t(44) = 2.75$ | | $t(42) = 2.56$ | | $t(43) = 0.91$ | | $t(44) = -0.52$ | |
| | $p$¡0.001 | | $p = 0.009$ | | $p = 0.015$ | | $p = 0.371$ | | $p = 0.608$ | |
| | $BF_{10} = 45.20$ | | $BF_{10} = 4.45$ | | $BF_{10} = 2.87$ | | $BF_{10} = 0.24$ | | $BF_{10} = 0.18$ | |
| **Lenient** | | | | | | | | | | |
| | 0.62(0.22) | 0.65(0.25) | 0.67(0.20) | 0.67(0.25) | 0.64(0.18) | 0.62(0.23) | 0.61(0.21) | 0.54(0.22) | 0.60(0.22) | 0.57(0.19) |
| | $t(43) = 0.94$ | | $t(44) = -0.24$ | | $t(42) = -0.63$ | | $t(43) = -2.02$ | | $t(44) = -1.29$ | |
| | $p = 0.355$ | | $p = 0.740$ | | $p = 0.530$ | | $p = 0.05$ | | $p = 0.202$ | |
| | $BF_{10} = 0.25$ | | $BF_{10} = 0.17$ | | $BF_{10} = 0.20$ | | $BF_{10} = 1.03$ | | $BF_{10} = 0.35$ | |

Table 4.5: Experiment 1: Comparison of pre- and post instruction recall accuracy in all groups, including the Control, where the null effect confirms the absence of a learning-to learn effect. The mean and standard deviation for pre- and post-instruction recall accuracy are reported as M(SD).

| | All | Body | Loci | Story | Activity | Control |
|---|---|---|---|---|---|---|
| **Strict** | Main Effect:<br>$F(4, 216) = 4.89$<br>$\eta_p^2 = 0.08$<br>$p$¡0.001<br>$BF_i$¿100 | Across Positions:<br>*Body¿Control*<br>0.19(0.05), $d = 0.75$<br>$p = 0.006$ | Across Positions:<br>*Loci¿Control*<br>0.19(0.05), $d = 0.72$<br>$p = 0.007$ | Across Positions:<br>no significant<br>differences | Across Positions:<br>*Activity¡Body*<br>0.19(0.05), $d = 0.72$<br>$p = 0.052$ | Across Positions:<br>*Control¡Body*<br>*Control¡Loci* |
| | Interaction Effect:<br>$F(21.82, 1178.12) = 4.19$<br>$p$¡0.001<br>$\eta_p^2 = 0.07$<br>$BF_i$¿100 | Individual Positions:<br>*Body¿Control*<br>at P $4-10$ ($p$¡0.05)<br>*Body¿Activity*<br>at P 4,8, and 9 ($p$¡0.05) | Individual Positions:<br>*Loci¿Control*<br>at P $5-10$ ($p$¡0.05)<br>*Loci¿Activity*<br>at P 3 and 6 ($p$¡0.05) | Individual Positions:<br>*Story¿Control*<br>at P 10 ($p$¡0.05)<br>*Control¡Activity*<br>at P 10 ($p$¡0.05) | Individual Positions:<br>*Activity¿Control*<br>at P 10 ($p$¡0.05) | Individual Positions:<br>see left columns |
| **Lenient** | Main Effect:<br>$F(4, 216) = 2.28$<br>$\eta_p^2 = 0.04$<br>$p = 0.062$<br>$BF_i = 0.45$ | Across Positions:<br>no significant<br>differences | Across Positions:<br>no significant<br>differences | Across Positions:<br>no significant<br>differences | Across Positions:<br>no significant<br>differences | Across Positions:<br>no significant<br>differences |
| | Interaction Effect:<br>$F(26.57, 1434.97) = 3.29$<br>$p$¡0.001<br>$\eta_p^2 = 0.06$<br>$BF_i$¿100 | Individual Positions:<br>*Body¿Control*<br>at P 8, and 9 ($p$¡0.05)<br>and P 10 ($P$¡0.001) | Individual Positions:<br>*Loci¿Control*<br>at P 8 ($p = 0.051$)<br>and P 10 ($P$¡0.001) | Individual Positions:<br>no significant<br>differences | Individual Positions:<br>*Activity¿Control*<br>at P 1 ($p$¡0.05) | Individual Positions:<br>see left columns |

Table 4.6: Experiment 1: Comparison of the mnemonic scaffolds and Control. The median and standard deviation for pre- and post-instruction recall accuracy is reported as M(SD); $BF_i$ is short for $BF_{inclusion}$, P is short for position

## 4.5 Supplementary Materials

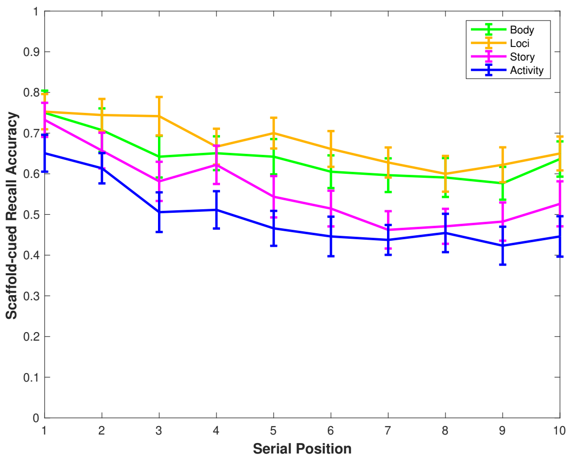
### 4.5.1 Experiment 1

#### Absence of sampling bias

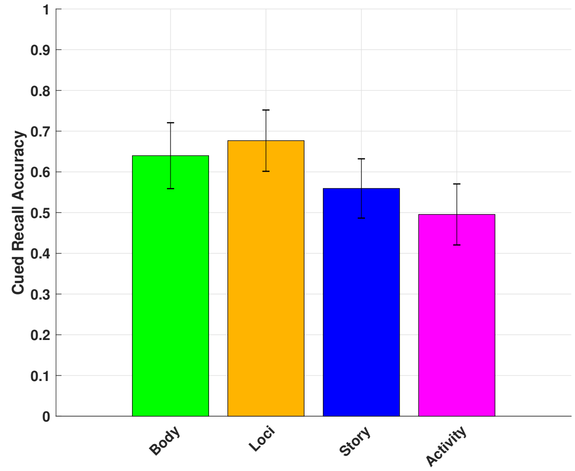
To verify the absence of a subject sampling bias across groups, we conducted a 5 (Body, Loci, Story, Activity, Control) $\times$ 10 (Serial Position 1–10) mixed ANOVA on the average serial recall accuracy for each serial position averaged across the two lists from the pre-instruction baseline memory phase that were studied prior to receiving instructions on mnemonic scaffolds (Figure 4.3, dotted lines). For both scoring methods, this produced neither a significant main effect of Group (strict: $F(4, 216) = 0.72$, $p = 0.580$, $\eta_p^2 = 0.01$, $BF_{\text{inclusion}} = 0.02$, lenient: $F(4, 216) = 0.85$, $p = 0.497$, $\eta_p^2 = 0.02$, $BF_{\text{inclusion}} = 0.01$), nor a significant interaction effect (strict: $F(25.47, 1375.26) = 0.91$, $p = 0.597$, $\eta_p^2 = 0.02$, $BF_{\text{inclusion}} < 0.01$, lenient: $F(33.37, 1802.00) = 1.05$, $p = 0.939$, $\eta_p^2 = 0.02$, $BF_{\text{inclusion}} < 0.01$), with Bayes factors providing strong evidence for the null, confirming the absence of a subject sampling bias.

#### Absence of a learning-to-learn effect in the Control Group

To check whether participants in the Control Group received a memory benefit from the filler instruction (reading the words out loud) we conducted a paired-samples $t$-test comparing pre- with post-instruction recall accuracy in the Control Group (Figure 4.3e). The test was non-significant, with the Bayes factor strongly favouring a null effect for both strict (pre-instruction: $M = 0.40$, $SD = 0.27$, and post-instruction: $M = 0.39$, $SD = 0.23$; $t(44) = -0.52$, $p = 0.608$, $BF_{10} = 0.18$), and lenient scoring (pre-instruction: $M = 0.60$, $SD = 0.22$, and post-instruction: $M = 0.57$, $SD = 0.18$; $t(44) = -1.29$, $p = 0.202$, $BF_{10} = 0.35$). This confirms the absence of a measurable learning-to-learn effect in the Control Group.




(a) Serial position curves of scaffold-cued recall accuracy




(b) Scaffold-cued recall collapsed across serial position

Figure 4.9: Experiment 1: Scaffold-cued recall accuracy for each mnemonic scaffold. Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994).

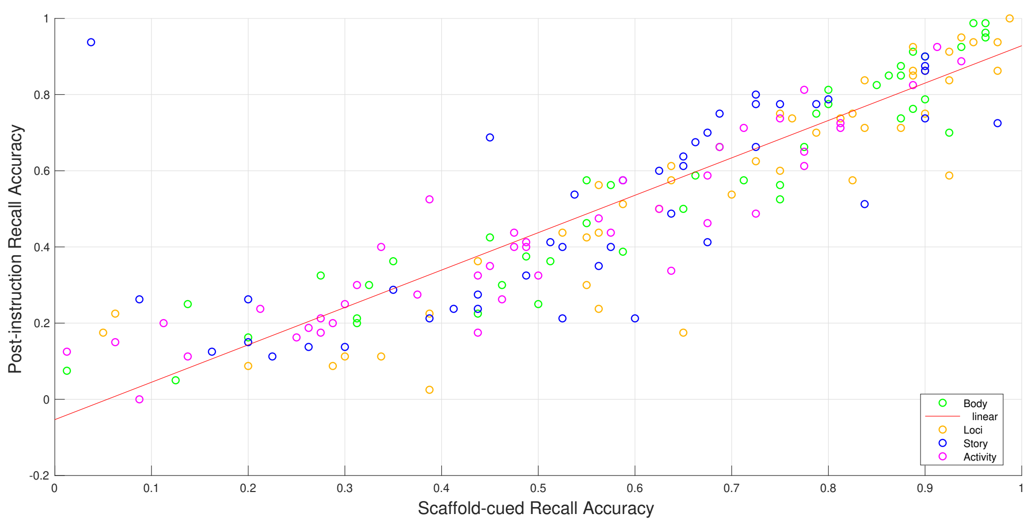



Figure 4.10: Experiment 1: Average post-instruction serial recall accuracy (strict scoring) by scaffold-cued recall accuracy

### Scaffold-cued recall

To evaluate whether participants were successful in forming word–scaffold associations (as instructed) we conducted a linear regression to predict post instruction serial recall accuracy (strict scoring) based on scaffold-cued recall accuracy (Figure 4.10. The scaffold-cued recall did indeed predict post-instruction recall accuracy in the Scaffold Groups ($R^2 = 0.76$, $\beta = 0.87$, $p < 0.01$, $BF_{10} > 100$). This strong relationship between cued-recall and post-instruction recall suggests that scaffold-cued recall is a good indicator of strategy use. We then compared the proportion of correctly reported word–scaffold associations from the scaffold-cued recall test, across the four Scaffolds (Figure 4.9). We conducted a 4 (Body, Loci, Story, Activity) $\times$ 10 (Serial Position 1–10) mixed ANOVA on the average scaffold-cued recall accuracy for each serial position. This produced a significant main effect of Scaffold ($F(3, 172) = 4.68$, $p = 0.004$, $\eta_p^2 = 0.05$, $BF_{\text{inclusion}} = 7.41$). The interaction was non-significant with the Bayes factor providing strong evidence for a null effect of Serial Position $\times$ Scaffold ($F(21.29, 1220.33) = 1.25$, $p = 0.202$, $\eta_p^2 = 0.02$, $BF_{\text{inclusion}} = 0.01$), indicating that the effect of Scaffold was not modulated by Serial Position. Tukey's post-hoc tests revealed scaffold-cued recall accuracy was significantly higher in both the Body Group [Mean Difference (Body – Activity)=0.14, $SE = 0.05$, $d = 0.20$, $p = 0.037$] and the Loci Group [Mean Difference (Loci – Activity)=0.12, $SE = 0.05$, $d = 0.26$, $p = 0.004$] than in the Routine Activity Group. No significant differences ($p > 0.05$) were detected between the Body Group, the Loci Group, and the Story Group, nor between the Story Group and the Routine Activity Group. This suggests that participants in the Body and Loci Group were more successful in associating the study words with parts of their scaffold than participants in the Routine Activity Group. A Bayesian paired-samples t-test comparing scaffold-cued recall accuracy of the Body Group and Loci Group $t(43) = -0.95$, $p = 0.347$, $BF_{10} = 0.25$) suggested scaffold-cued recall accuracy was equally high for these two scaffolds.

Next, we were interested in whether scaffold–word associations may have relied upon serial recall. Therefore, we collapsed across serial position and included scaffold-cued recall accuracy as a covariate in a one-way ANCOVA to determine significant effects of Scaffold on serial recall accuracy (strict scoring) controlling for scaffold-cued recall accuracy. If scaffold–word associations are the driving force behind serial recall using the scaffolds, the addition of scaffold-cued recall as a covariate would be expected to render the main effect of Scaffold non-significant. The effect of the covariate scaffold-cued recall accuracy itself was significant, strongly supported by the Bayes Factor ($F(1, 171) = 526.08$, $p < 0.001$, $\eta_p^2 = 0.76$, $BF_{\text{inclusion}} > 100$). The main effect of Group after controlling for scaffold-cued recall accuracy was non-significant, with the Bayes Factor strongly favouring a null effect ($F(3, 171) = 0.78$, $p = 0.508$, $\eta_p^2 = 0.01$, $BF_{\text{inclusion}} = 0.07$). In light of the large effect of Group without the covariate reported earlier, this (cautiously) suggests that scaffold–word associations may, indeed, have been relied upon the scaffolds during serial recall, itself.

### Effect of self-paced study time

We were interested in whether study time (i.e., the time participants spent looking at the word and part of the scaffold on each screen during study) might be the mechanism that produced group differences in serial recall accuracy as more time spent engaging with a study item during encoding might translate into higher recall accuracy. First, to test whether study time varied significantly between groups (Figure 4.11), we conducted a one-way ANOVA of Scaffold on study time. This yielded a significant main effect of Group ($F(4, 16) = 4.52$, $p = 0.002$, $\eta_p^2 = 0.08$, $BF_{\text{inclusion}} = 8.70$). Tukey's post-hoc test revealed that study time in the Routine Activity Group was significantly shorter than in the Story Group [Mean Difference (Activity – Story)= $-3.08$, $SE = 1.03$, $d = -0.55$, $p = 0.026$]. Study time in the Story Group was significantly longer than in the Control Group [Mean Difference (Story – Control)= $4.24$, $SE = 1.03$, $d = 0.74$, $p < 0.001$]. This raises the question: does this difference in study time during encoding modulate the group differences in serial recall accuracy? To answer this, we conducted a one-way ANCOVA of Group (Body, Loci, Story, Activity, Control) on serial recall accuracy with study time as a covariate. The covariate study time was significant ($F(1, 215) = 96.57$, $p < 0.001$, $\eta_p^2 = 0.31$, $BF_{\text{inclusion}} > 100$). This

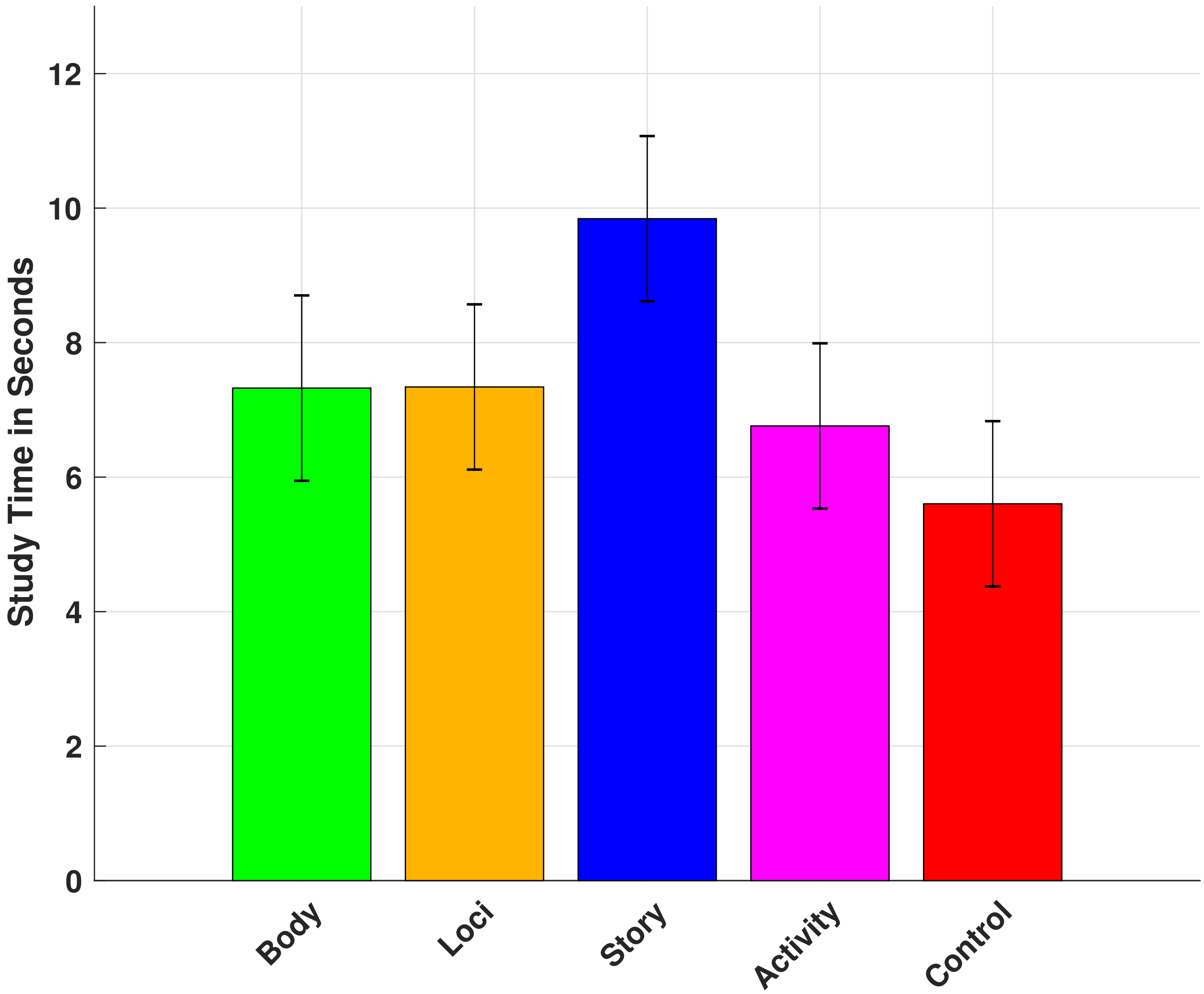



Figure 4.11: Experiment 1: Study times in seconds for each mnemonic scaffolds group. Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994).

analysis revealed a significant effect of Group after controlling for study time ($F(4, 215) = 5.55$, $p < 0.001$, $\eta_p^2 = 0.09$, $BF_{\text{inclusion}} > 100$). A Tukey's post-hoc test revealed that after controlling for study time, participants in the Body Group recalled significantly more words than participants in the Story Group [Mean Difference (Body – Story )=0.15, $SE = 0.05$, $d = 0.56$, $= 0.014$], and than participants in the Routine Activity Group [Mean Difference, Body – Activity=0.13, $SE = 0.05$, $d = 0.52$, $p = 0.03$], and participants in the Control Group [Mean Difference (Body – Control)=0.14, $SE = 0.05$, $d = 0.54$, $p = 0.024$]. Participants in the Loci Group recalled significantly more words than participants in the Story Group [Mean Difference (Loci – Story)=0.15, $SE = 0.05$, $d = 0.54$, $p = 0.016$] and participants in the Control [Mean Difference (Loci – Control)=0.13, $SE = 0.05$, $d = 0.52$, $= 0.029$]. The difference between participants in the Loci and Activity Group after controlling for study time was significant [Mean Difference (Loci – Activity)=0.13, $SE = 0.05$, $d = 0.50$, $p = 0.037$]. These findings support the advantage of the Body and Loci Scaffold over the Routine Activity Scaffold and the Control and further indicate that when accounting for study time, the Body Scaffold and the Loci Scaffold provide an advantage over the Story Scaffold. The longer study time induced by the Story Scaffold did not, apparently, translate directly into a memory advantage.

We also wondered if slow typists might perform worse due to increased output interference during serial recall. However, inter-response time (not reported here) produced null or inconclusive effects involving Scaffold, suggesting that typing speed was not a major factor in these experiments.

**Absence of sampling bias with respect to individual difference measures**

To verify the absence of a subject sampling bias, we first conducted one-way ANOVAs of Group (Body, Loci, Story, Activity, Control) on the scores of the BRQ, VVIQ, and PFT, respectively. These were all non-significant ($p > 0.05$), supported null effects ($BF < 0.3$), indicating that the five Groups were well matched on the three individual differences measures.

### Correlations with pre-instruction serial recall accuracy

To verify that a relationship between the scores in the individual differences tasks and the memory advantage provided by the mnemonic scaffolds was not confounded by a preexisting relationship in pre-instruction memory, we conducted classical and Bayesian linear regressions to predict pre-instruction recall accuracy based on the three individual differences measures. For the BRQ ($R^2 < 0.01$, $\beta = 0.06$, $p = 0.329$, $BF_{10} = 0.23$) and VVIQ ($R^2 = 0.01$, $\beta = -0.02$, $p = 0.728$, $BF_{10} = 0.16$), no such pre-existing relationship between pre-instruction recall accuracy and scores in the individual differences tasks was found. For the PFT, however, the significant $p$-value and the Bayes factor providing strong evidence against a null effect indicate that higher scores in the PFT predicted higher baseline memory ($R^2 = 0.07$, $\beta = 0.27$, $p < 0.001$, $BF_{10} = 335.44$), with a very small proportion of the variance explained. We will further discuss this below.

### Absence of sex differences

We were interested in sex differences in our sample, because if scaffolds differ in effectiveness by sex, that could be relevant to training protocols. Previous research has found a small overall female advantage in most verbal memory tasks, and no sex differences in serial recall (for a review and meta analysis, see Voyer, Saint Aubin, Altman, & Gallant, 2021). In our sample, there were 127 female participants, 90 male participants, and 4 participants who selected "prefer not to answer". To test whether sex had an effect on serial recall (strict scoring), we ran an ANOVA of sex (female, male, prefer not to answer) on pre-instruction serial recall accuracy ($F(2, 218) = 2.10$, $p = 0.125$, $\eta_p^2 = 0.02$, $BF_{inclusion} = 0.433$) and post-instruction serial recall accuracy ($F(2, 218) = 2.29$, $p = 0.104$, $\eta_p^2 = 0.02$, $BF_{inclusion} = 0.474$). In both cases, the effect of sex was non-significant with the Bayes Factor in the inconclusive range. When including Group (Body, Loci, Story, Activity, Control) as a covariates in both analyses, the $p$-values remained non-significant (pre-instruction: 0.104, post-instruction: $p = 0.082$), and the Bayes Factors remained inconclusive (pre-instruction: 0.43, post-instruction: 0.82). Notably, the partial eta squares are tiny, and unlikely to be

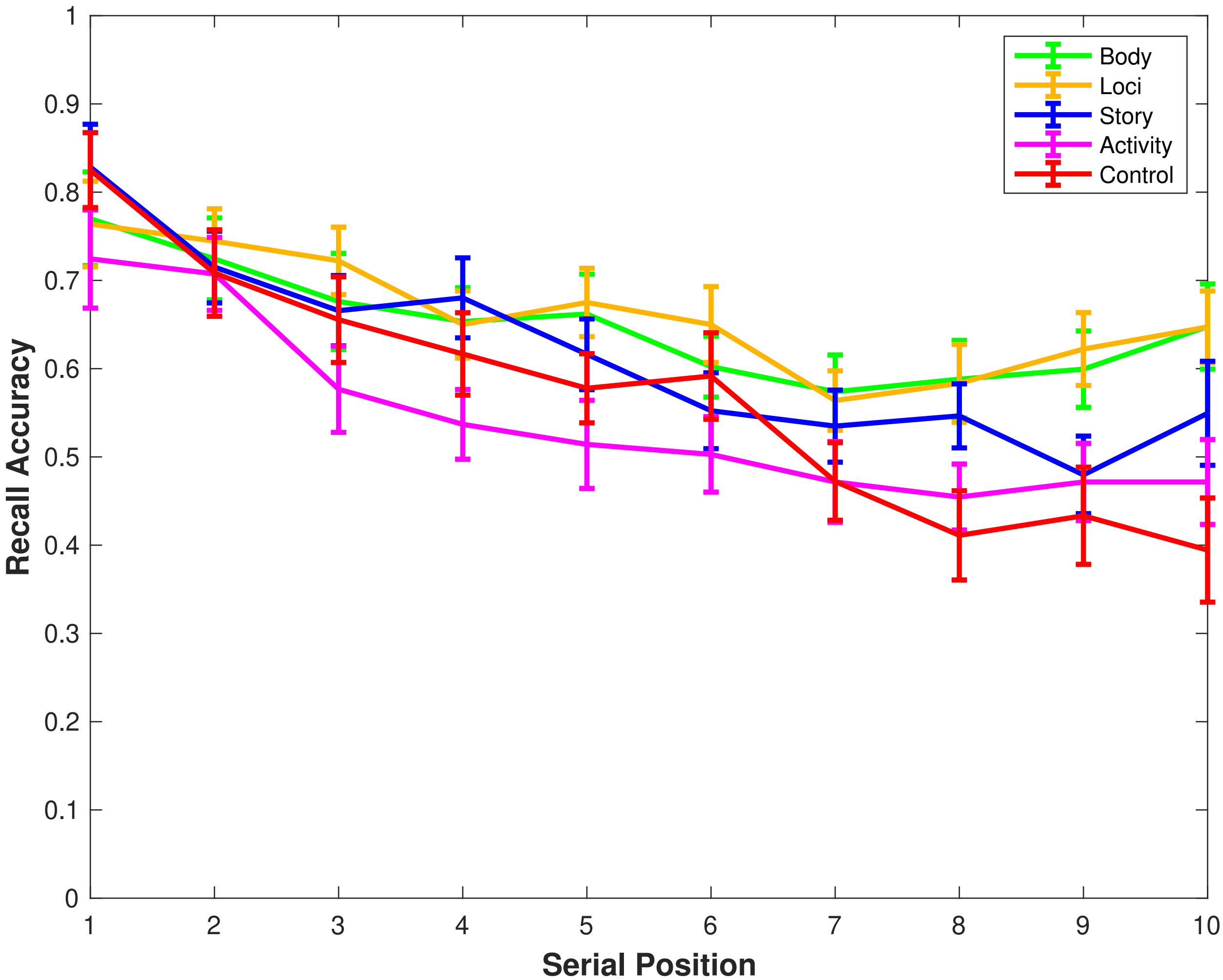



Figure 4.12: Experiment 1: Serial position curve for post-instruction recall accuracy (lenient scoring) as a function of group. Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994).

meaningful at that effect size. This suggests that sex differences remain elusive. In addition, the null interaction effect also suggests that that null sex difference does not emerge for any particular scaffold. Even though the Bayes Factors are inconclusive, our classical ANOVAs echo previous findings noting an absence of sex differences in serial recall (Voyer et al., 2021). In sum, these results speak against the possibility that males and females differ in their ability to adopt scaffold-based strategies

**Serial position curve for lenient scoring**

The serial position curve for lenient scoring of Experiment 1 is reported below in Figure 4.12

### 4.5.2 Experiment 2

**Absence of sampling bias**

Using the same approach as in Experiment 1 to verify the absence of a subject sampling bias, we conducted a 4 (Sticker-on-Body, Sticker-on-Table, No-Sticker, Control) × 10 (Serial Position 1–10) repeated-measures ANOVA on the average pre-instruction recall accuracy for each serial position of the three pre-instruction lists. Neither a significant main effect of Group ($F(3, 143) = 1.52$, $p = 0.213$, $\eta_p^2 = 0.03$, $BF_{inclusion} = 0.17$), nor a significant interaction effect $F(16.95, 807.87) = 0.91$, $p = 0.555$, $\eta_p^2 = 0.02$, $BF_{inclusion} < 0.01$) was found, with Bayes factors providing strong evidence for a null effect. This confirms the absence of a subject sampling bias.

**Absence of a learning-to-learn effect in the Control Group**

In fact, a paired-samples $t$-test comparing pre-instruction recall accuracy to post-instruction recall accuracy revealed that saying the words loud actually lead to significantly lower recall than receiving no instruction (pre-instruction: $M = 0.54$, $SD = 0.24$, and post-instruction: $M = 0.43$, $SD = 0.25$, $t(36) = 3.10$, $p = 0.004$, $BF_{10} = 9.79$), plotted in Figure 4.7d.

**Scaffold-cued recall**

To evaluate whether the Body Scaffold Variants differed in the success with which participants actually formed word-scaffold associations (as instructed) we conducted a 3 (Sticker-on-Body, Sticker-on-Table, No-Sticker) x 10 (Serial Position 1 -10) mixed ANOVA on the average scaffold-cued accuracy for each serial position, plotted in Figure 4.13 and 4.14. This revealed a non-significant main effect of Group ($F(2, 107) = 0.59$, $p = 0.555$, $\eta_p^2 = 0.011$, $BF_{inclusion} = 0.08$). The interaction was significant ($F(15.56, 832.41 = 1.60$, $p = 0.065$, $\eta_p^2 = 0.03$, $BF_{inclusion} = 0.04$), with the Bayes factor strongly favouring the null suggesting that the interaction was negligible in magnitude, and that the Body Scaffold Variants had no effect on cued-recall accuracy.

To check whether the variants of the Body Scaffold are also equal with regard to the

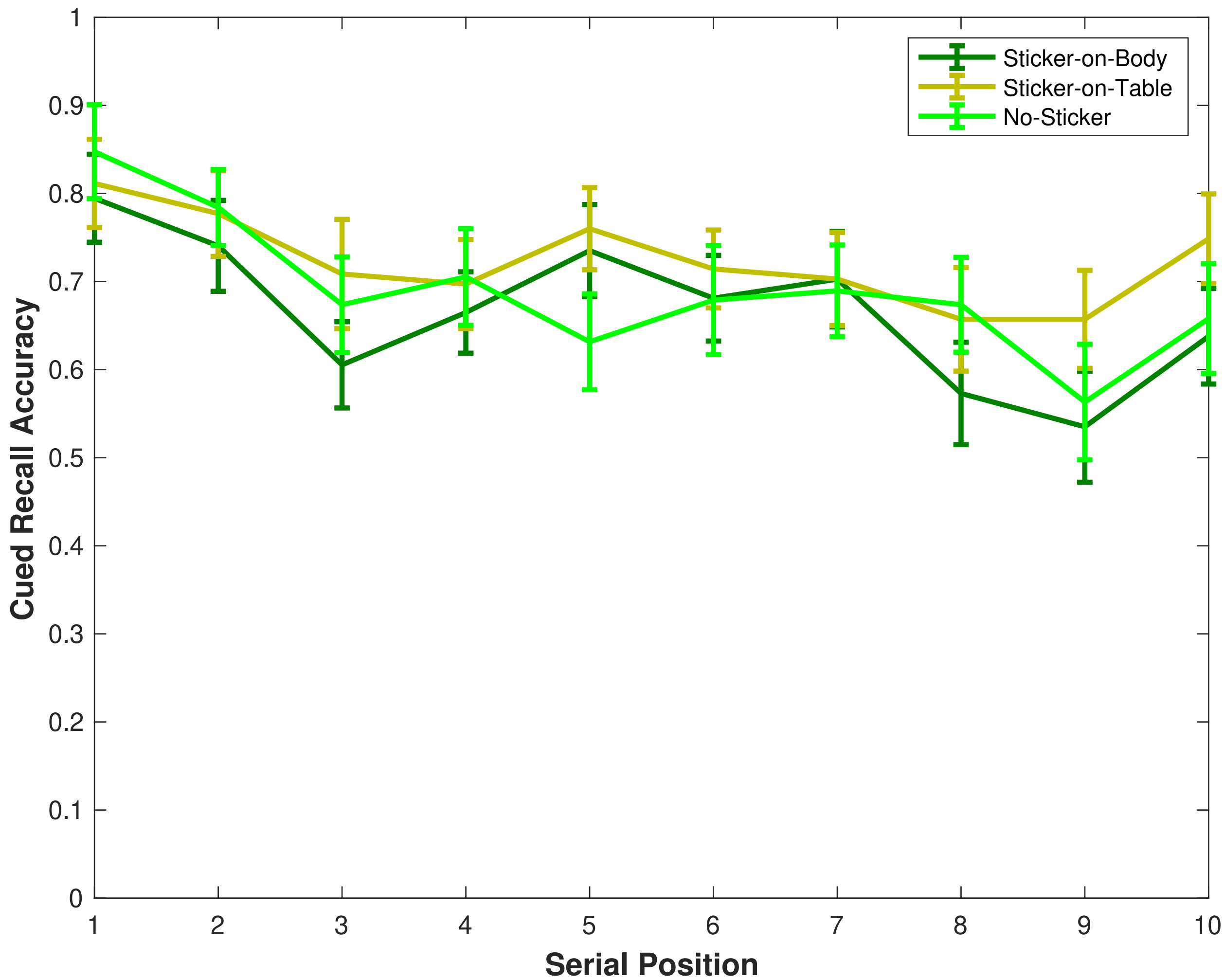



Figure 4.13: Experiment 2: Serial position curves of cued recall accuracy for each group. Error bars are standard errors of the mean corrected for subject variability (Loftus and Masson, 1994).

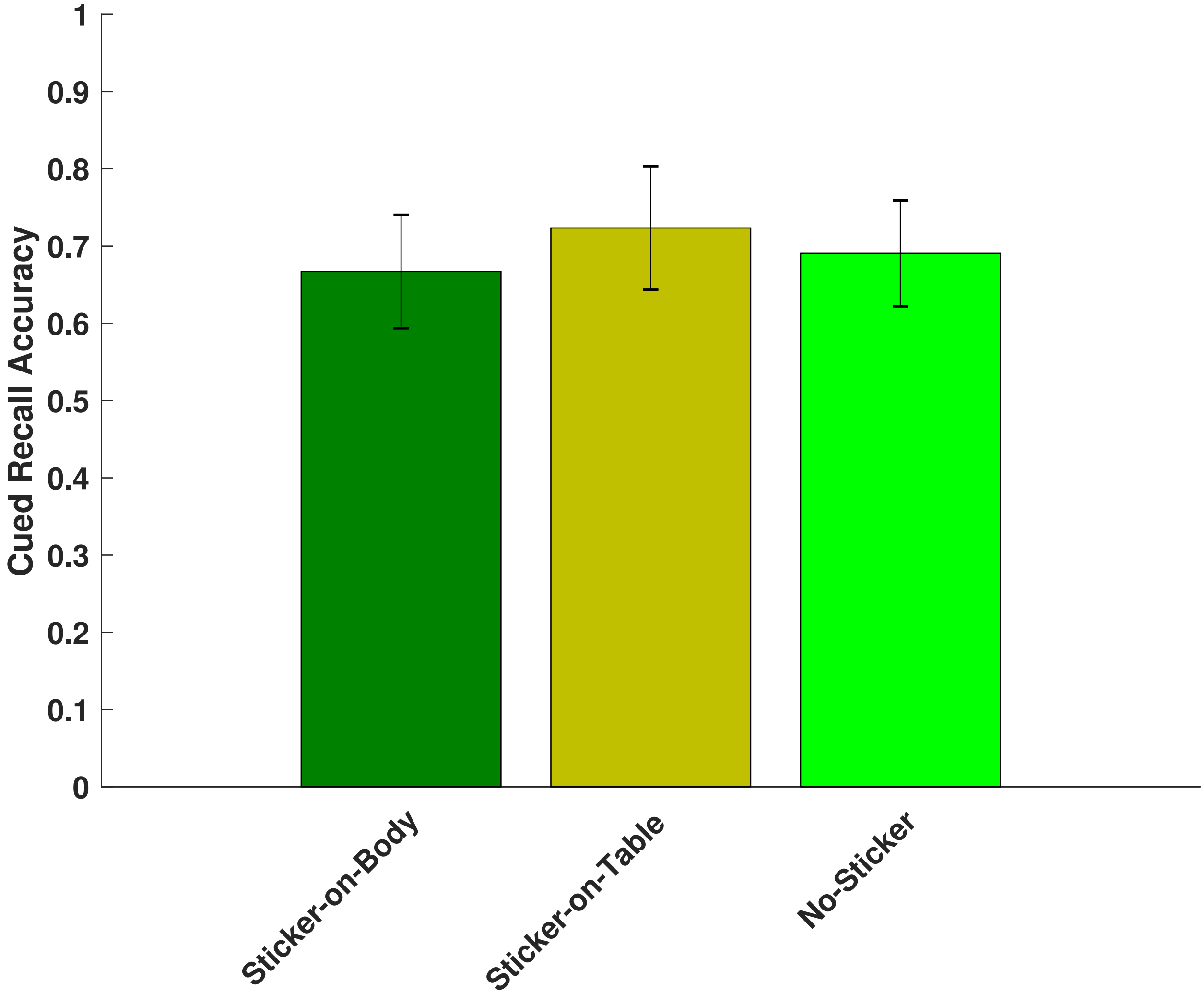



Figure 4.14: Experiment 2: Scaffold-cued recall accuracy for each group. Error bars are standard errors of the mean corrected for subject variability (Loftus and Masson, 1994).

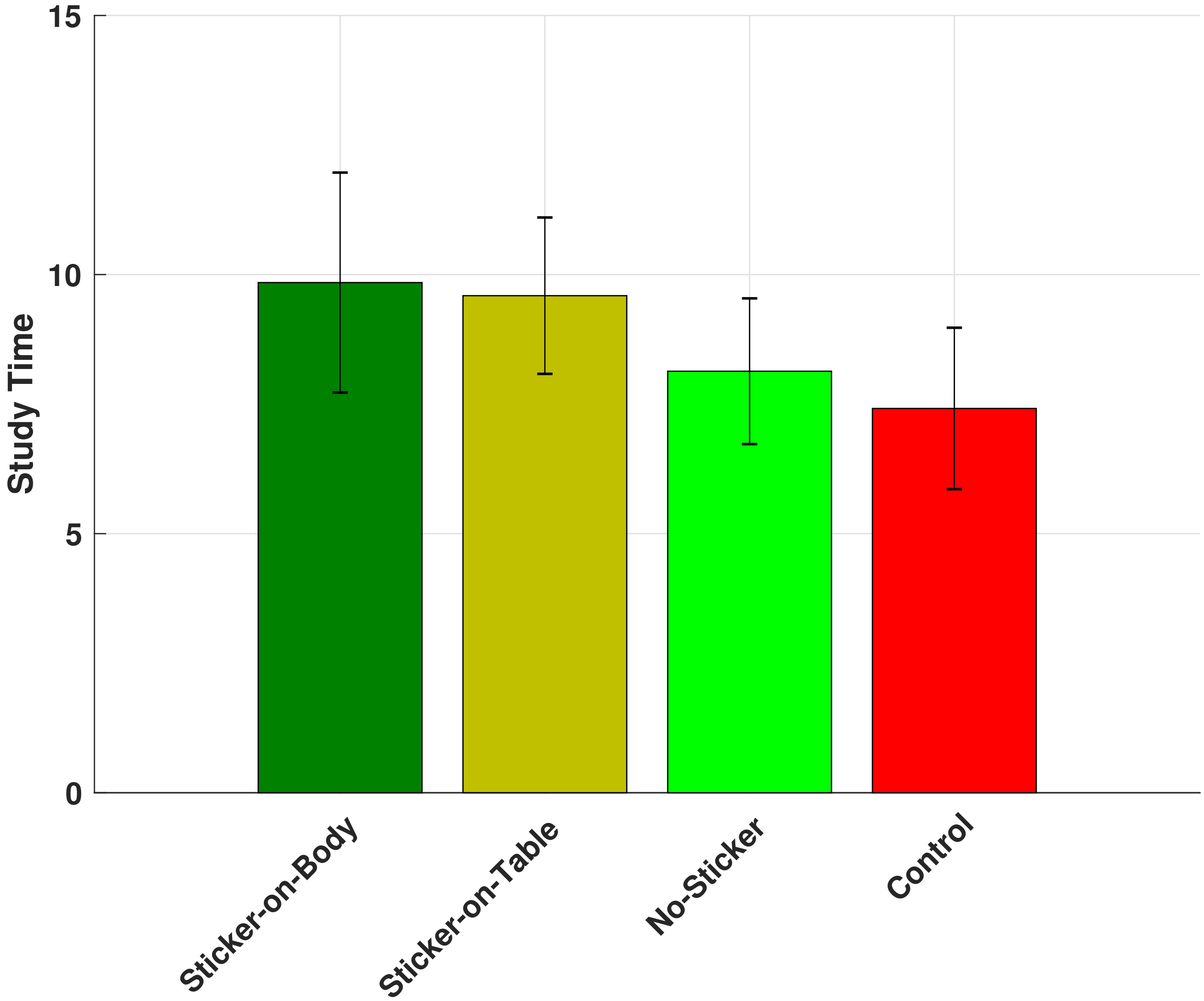



Figure 4.15: Experiment 2: Study time for each group. Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994).

success with which participants formed scaffold–word associations, we conducted a one-way ANOVA of Body Scaffold Variant on scaffold-cued recall accuracy (Figures 4.14 and 4.13). This revealed a non-significant effect of Body Scaffold Variant with the Bayes Factor providing strong evidence for a null effect ($F(2, 107 = 1.38$, $p = 0.257$, $\eta_p^2 = 0.01$, $BF_{inclusion} = 0.14$) indicating that the Body Scaffold Variants do not differ in how they facilitated scaffold–word associations.

### Effect of self-paced study time

We were interested in whether the different instructions with varying degrees of bodily engagement affected study time (Figure 4.15). A one-way ANOVA of Body Scaffold Variant (Sticker-on-Body, Sticker-on-Table, No-Sticker, Control) on study time was non-significant with the Bayes factor providing evidence for a null effect ($F(2, 107) = 0.64$, $p = 0.528$, $\eta_p^2 = 0.01$, $BF_{\text{inclusion}} = 0.14$), indicating that the variants of the Body Scaffold did not vary in study time. We also conducted a one way ANCOVA on post-instruction recall accuracy with study time as a covariate. The effect of the covariate study time was significant ($F(1, 142) = 23.21$, $p < 0.001$, $\eta_p^2 = 0.14$, $BF_{\text{inclusion}} > 100$), and after controlling for study time, the main effect of Group was significant ($F(3, 142) = 6.99$, $p < 0.001$, $\eta_p^2 = 0.13$, $BF_{\text{inclusion}} > 100$). Tukey's post-hoc tests revealed significantly ($p < 0.05$) higher recall accuracy of all Body Scaffolds Variants over Control. This suggests that after controlling for study time, all Body Scaffold Variants outperformed Control, while the variants of the Body Scaffold do not differ among themselves.

### Correlations with pre-instruction serial recall accuracy

To verify that a relationship between the scores in the individual differences questionnaires and the memory advantage provided by the mnemonic scaffolds was not confounded by a pre-existing relationship in pre-instruction baseline memory, we conducted classical and Bayesian linear regressions to predict pre-instruction recall accuracy based on the three individual differences measures. We found the same pattern as in Experiment 1: for the BRQ ($p = 0.931$, $BF_{10} = 0.18$) and VVIQ ($p = 0.829$, $BF_{10} = 0.18$), no such pre-existing relationship between baseline recall and scores in the individual differences tasks was found. For the PFT, the significant p-value and Bayes factor providing strong evidence against a null effect ($F(1, 146) = 11.11$, $\beta = 0.28$, $R^2 = 0.07$, $p < 0.001$, $BF_{10} = 25.74$), indicate that higher scores in the PFT predicted higher pre-instruction accuracy.

Table 4.7: Question 1: Did you associate the list words with your body parts when studying them?

| Group | always | mostly | sometimes | never | Total |
|---|---|---|---|---|---|
| Sticker-on-Body | 13 | 14 | 6 | 4 | 37 |
| Sticker-on-Table | 15 | 13 | 4 | 3 | 35 |
| No-Sticker | 12 | 20 | 5 | 1 | 38 |
| Total | 40 | 47 | 15 | 8 | 110 |

Table 4.8: Question 2: If so, did connecting the words to parts of your body make remembering the words easier?

| Group | yes | no | I don't know | Total |
|---|---|---|---|---|
| Sticker-on-Body | 27 | 5 | 5 | 37 |
| Sticker-on-Table | 26 | 4 | 5 | 35 |
| No-Sticker | 26 | 8 | 4 | 38 |
| Total | 79 | 17 | 14 | 110 |

Table 4.9: Question 3: Have you used this memorization technique before?

| Group | yes | no | Total |
|---|---|---|---|
| Sticker-on-Body | 6 | 31 | 37 |
| Sticker-on-Table | 3 | 32 | 35 |
| No-Sticker | 7 | 31 | 38 |
| Total | 16 | 94 | 110 |

### Self-report questions

At the end of Experiment 2, we asked participants whether a) they associated list words with body parts when studying them (Table 4.7). b) connecting words with body parts made remembering the words easier, (Table 4.8), and whether c) they had used this technique before (Table 4.9). We were interested in whether the equivalence of the variants of the Body Scaffolds was also reflected in the responses of the self report questions. This was the case, as responses did not differ significantly ($p > 0.05$) across groups for any of the questions; Question 1: $\chi^2(6)= 0.656$, Question 2: $\chi^2(4)= 0.817$, Question 3: $\chi^2(2)= 0.461$, suggesting that the variants of the Body Scaffold do not differ among themselves in terms of self-reported usefulness or prior-usage.

### Serial position curve for lenient scoring

The serial position curve for lenient scoring of Experiment 2 is reported below in Figure 4.16. There was no correlation between neither BRQ and VVIQ scores and recall accuracy, and a significant ($p > 0.05$) correlation between PFT scores and recall accuracy.

### Scatter plots from individual difference questionnaires

Below are the scatter plot form the individual difference questionnaires (Figures 4.17, 4.18, 4.19).

## 4.6 Example Scaffolds

### Body Scaffold example 1

- foot
- shin
- knee
- thigh

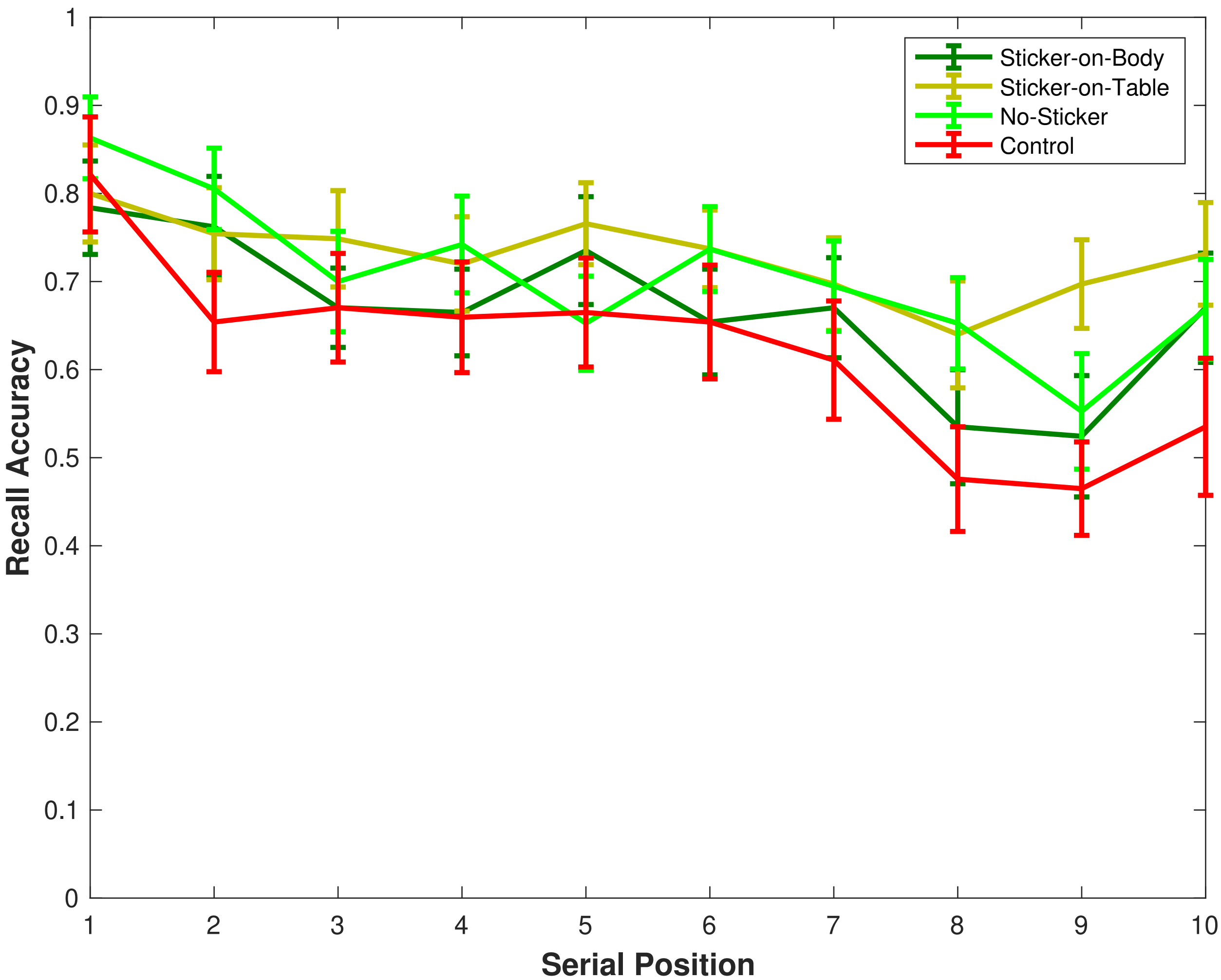



Figure 4.16: Experiment 2 - item memory (lenient scoring): Post-instruction recall accuracy for each group. Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994).

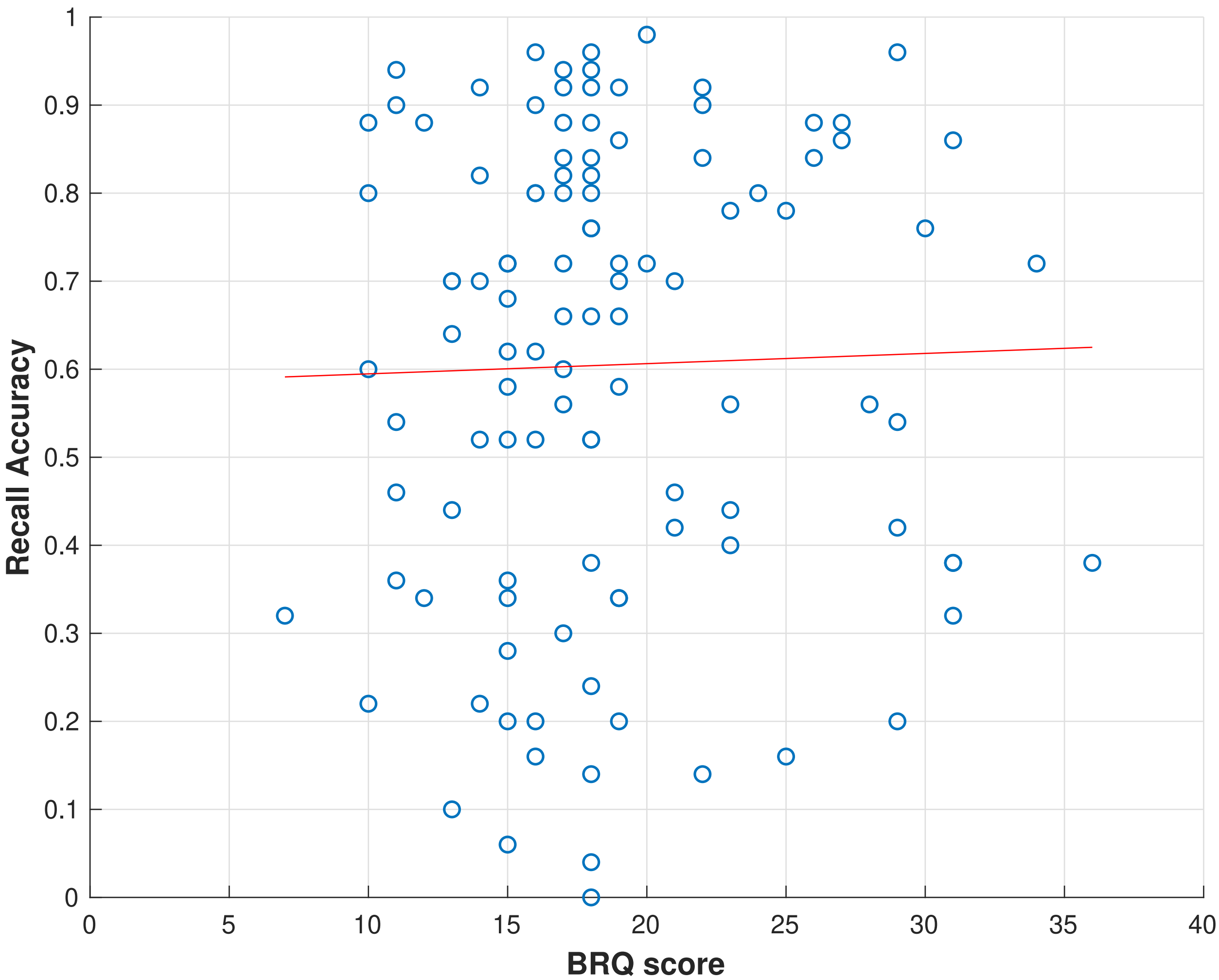



Figure 4.17: Experiment 2: Post-instruction recall accuracy by BRQ scores for all variants of the Body Scaffold combined

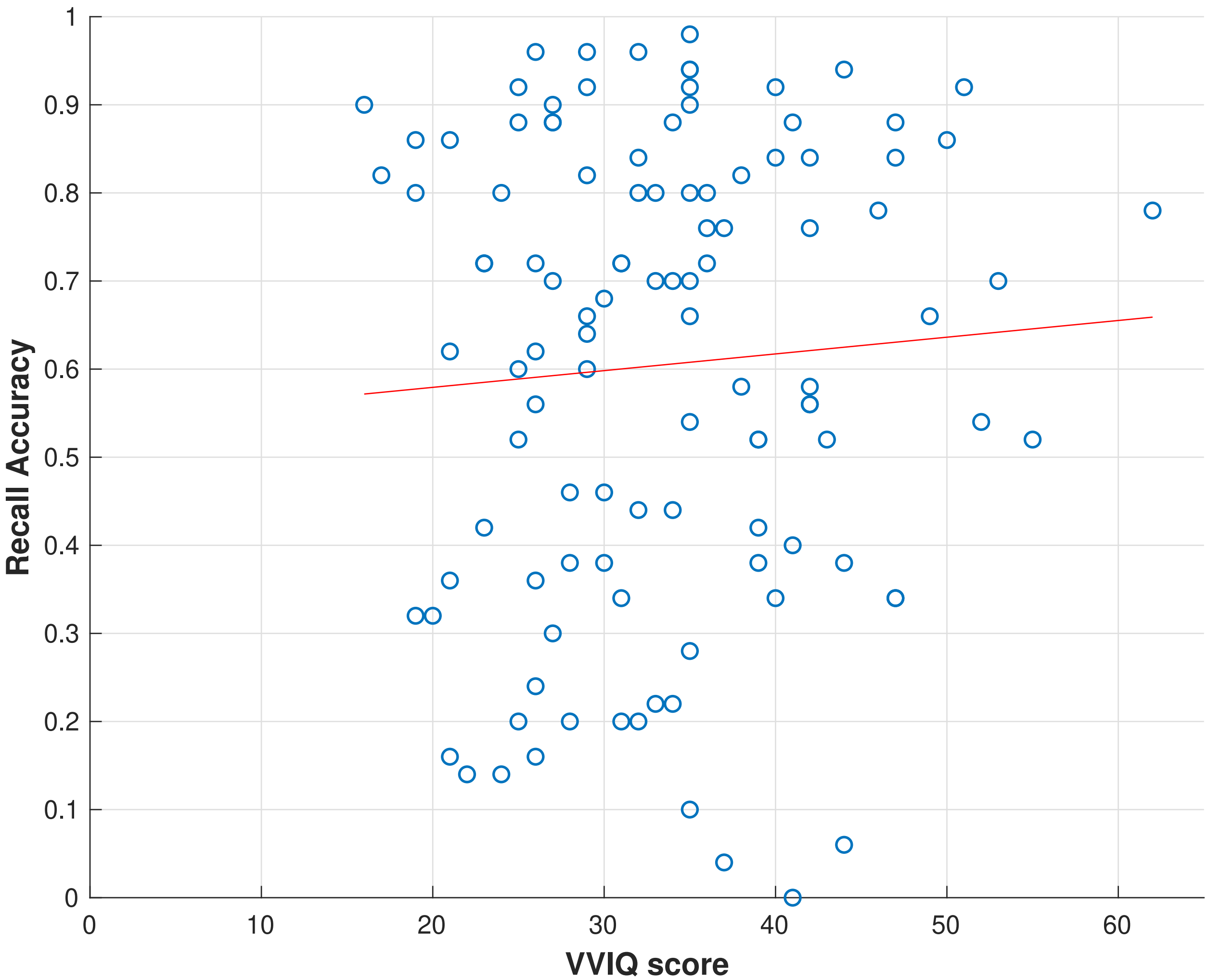



Figure 4.18: Experiment 2: Post-instruction recall accuracy by VVIQ scores for all variants of the Body Scaffold combined

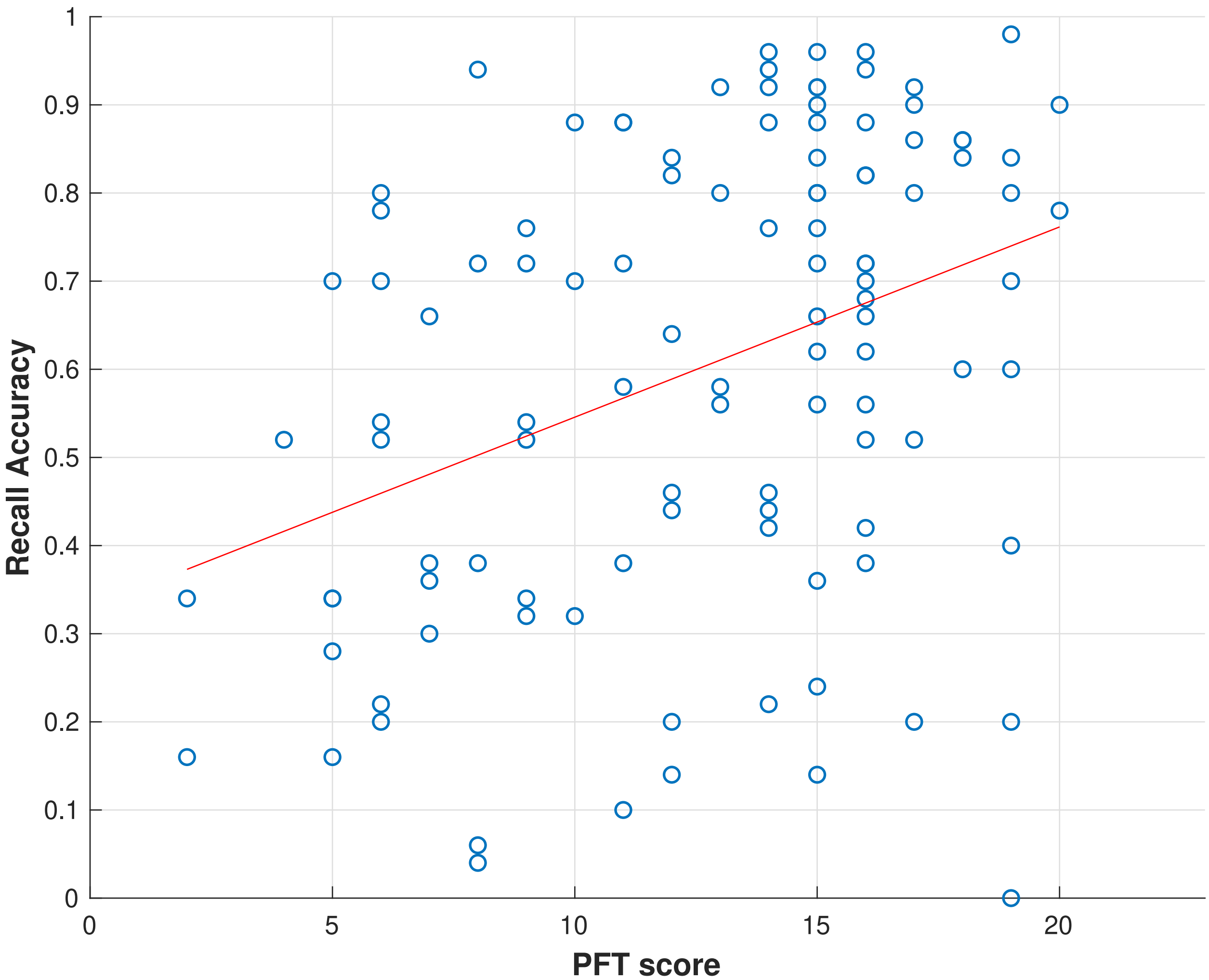



Figure 4.19: Experiment 2: Post-instruction recall accuracy by PFT scores for all variants of the Body Scaffold combined

- hipbone
- stomach
- chest
- neck
- head
- hair

## Body Scaffold example 2

- ankle
- calf
- knee
- stomach
- elbow
- shoulder
- neck
- chin
- eyes

## Loci Scaffold example 1

- street
- park
- bridge

- mall
- school
- church
- gym
- ball court
- track
- restaurant

## Loci Scaffold example 2

- coat closet
- indoor plant
- couch
- coffee table
- picture frames
- candle holders
- fuzzy carpet
- dining room table
- chairs
- curtains

## Autobiographical Story Scaffold example 1

- This weekend I went to my friend's cabin.
- I left for Pidgeon Lake Friday afternoon.
- Once we arrived, we begin playing games.
- Saturday morning started off with eggs Benedict for breakfast.
- We then did a Chinese gift exchange
- For dinner, I ate homemade pizza.
- We continued playing games for the remainder of the night.
- Before we left, we cleaned up the entire cabin.
- We then drove home
- Once I arrived home, I did homework then went to bed.

## Autobiographical Story Scaffold example 2

- I decided to meet my friend at the park
- I tied my shoes on my front step then started walking
- I met her at the corner and we hugged
- We decided that we were craving Slurpee as it was a hot summer day
- We walked past the park and field to the corner store
- We each got large Coca-Cola Slurpee
- We walked back towards the park making our way past the field
- A huge brown dog was playing fetch in the field

- He got an eye of my friends Slurpee dripping down the side of her cup
- The dog started madly running towards us and jumped on my friend taking the Slurpee

## Routine Routine Activity Scaffold example 1

- wake up
- check phone
- take a shower
- change my clothes
- do makeup
- do hair
- pack bag
- eat breakfast
- pack lunch
- put my shoes on

## Routine Routine Activity Scaffold example 2

- pick up toothbrush
- wet toothbrush with water
- put toothpaste on toothbrush
- wet toothbrush with water again
- brush your teeth for two minutes
- rinse off toothbrush with water

- spit out toothpaste in mouth into sink
- gargle water in mouth
- spit out water into sink
- put away toothbrush

# Chapter 5

# Mnemonic scaffolds: synthesis and future directions

The work presented in this dissertation helps us better understand under which conditions mnemonic scaffolds can be used to enhance memory and what driving factors behind their effectiveness may be. In this final chapter, we discuss how our findings challenge or expand previous theories on memory enhancement and how they may shape future research on memory training techniques.

## 5.1 The effectiveness of mnemonic scaffolds in enhancing memory: summary of key insights

Together, the findings from our experiments have shown that mnemonic scaffolds can enhance memory in novice learners after brief instructions within only one session (see also, e.g., Bouffard et al., 2017; Caplan & Madan, 2016; Legge et al., 2012). Our detailed analyses have revealed that the effectiveness of mnemonic scaffolds is influenced by some factors and not by others. Some of these factors challenge previous theories on memory enhancement and memory skill development. We first summarize the main answers to our research questions in the order we reported them in this dissertation before we discuss them in detail below.

Our first research question was *Under which conditions can story-based mnemonic scaffolds enhance memory for lists of words?* This research question was motivated by real-world teaching experience using autobiographical stories as mnemonic scaffolds. I was encouraged to investigate the Autobiographical Story Scaffold experimentally after teaching it to high school students and observing them use this technique successfully when studying for exams. Research on mnemonic strategies for high school students is important, as previous studies have found low levels of success in this population (Campos, Amor, & González, 2004). After reviewing the literature on the Method of Loci, we noticed that the Autobiographical Story Scaffold may be particularly well suited for older adults who struggle with the Method of Loci, as explained in Chapter 2. The main answers to our first research question are:

- for novice learners, Autobiographical Story Scaffolds benefit memory in recency positions in lists of ten words when one study item together with one sentence are shown at a time and when participants retype their sentences with the study items integrated
- locations mentioned in those stories can increase their effectiveness in serial recall for lists of ten words when one study item is shown together with one autobiographical sentence
- participants who form scaffold-word associations as instructed can benefit from those during serial recall for lists of ten and fifteen words when using either autobiographical or fictional scaffolds
- for lists of fifteen words, story-based mnemonic scaffolds do not enhance memory in novice learners when averaged across participants suggesting that list length may be a boundary condition for the effectiveness of story-based mnemonic scaffolds

Our second research question was *What are alternatives to the Method of Loci?* The Method of Loci is the most famous and often thought to be the best mnemonic strategy (Foer, 2011). We were interested in whether its apparent supremacy is justified or whether other scaffolds may be promising alternatives. Finding alternatives to the Method of Loci is important because previous research has suggested that this method is not useful for learners with low visuospatial aptitude (Sanchez, 2019) and is not successfully applied by older adults (e.g., Kliegl et al., 1990).

The key answers to our second research show that:

- for novice learners studying lists of ten words, the Body Scaffold is equally effective as the Method of Loci and a promising alternative
- the Autobiographical Story Scaffold, even though not as effective as the Method of Loci and the Body Scaffold, can enhance memory in novice learners
- routine activities are not effective mnemonic scaffolds for novice learners and not an alternative to the Method of Loci or the Body Scaffold

Our third research question was *Do individual differences in visual ability and body responsiveness contribute to the effectiveness of mnemonic scaffolds?* This question was motivated by the assumption that mnemonic scaffolds might differ from one another because they depend on skills that vary across individuals, or because they are influenced by individual traits or affinities. We have learned that:

- individual differences in visual ability measured with the VVIQ (Marks, 1973) and the PFT (French et al., 1963) and body responsiveness measured with the BRQ (Daubenmier, 2005) do not contribute to the effectiveness of mnemonic scaffolds
- engagement of the body does not change the effectiveness of the Body Scaffold

In the following paragraphs, we examine these answers to our research questions in depth, discussing how they challenge previous theories on mnemonic strategies.

## 5.2 Story-based mnemonic scaffolds may be effective under some conditions

The findings of Chapters 2, 3, and 4 have shown that story-based mnemonic scaffolds can enhance memory. Importantly, we were able to experimentally validate the anecdotal evidence of story scaffolds enhancing memory I observed in my students. In addition, we have shown that learners are able to rely on word-scaffold associations for serial recall of the words (we explain this in detail in the previous chapters). We also found that, for novice learners, list length may be a boundary condition.

Although participants benefited from their Autobiographical Story Scaffolds in recency positions of lists of ten words when shown one sentence from their autobiographical stories and one list word at a time, we did not replicate this for lists of fifteen words. Since we cannot know exactly how participants used their stories, we can only speculate what reasons for this lack of a mnemonic benefit may be. First, this might be due to the challenging nature of the task, if coming up with stories of fifteen words and studying multiple word lists overtaxed participants' attention and decreased their self-efficacy and motivation. Second, it may be due to the nature of the stories, if the quality of the stories decreases (by lacking detail in some sentences, for example) when participants struggle to come up with fifteen sentences and find themselves inserting "filler sentences" to reach the required number of sentences. Third, the lack of a mnemonic benefit with list length fifteen may be due to increased working memory load, if participants attempt to remember their whole fifteen sentences long story during the study phase. This detrimental effect of increased working memory load when participants think about their stories while recalling the study items is likely to also decrease the effectiveness of story scaffolds for shorter word lists. Similarly, as we observed in the experiment in Chapter 2, showing participants their whole story during study is not beneficial. This may be due to increased working memory load or diverted attention away from the study items to the stories.

The potentially detrimental effect of participants engaging with their stories during study rather than focusing on integrating the words into individual sentences also seems plausible given two important findings our experiments: First, contrary to our hypothesis, recall accuracy of the story scaffolds themselves did not predict serial recall accuracy for list length fifteen. In other words, memory for the stories as a whole did not contribute to the effectiveness of the strategy. However, across both experiments, we found that participants benefited from individual scaffold-word associations. Even though, on average, participants did not benefit from the scaffold instructions when studying lists of fifteen words, those participants who did form sentence-word associations as instructed and objectively measured by the sentence-cued recall task, saw a memory increase. Because there was no net effect of the scaffold, this implies not only that those participants who applied the scaffold benefited from it, but in fact, also that other participants who were either unable or unwilling to adopt it were harmed by the instructions in that their accuracy was lower than without the strategy. See Sanchez (2019) for a similar kind of cancellation effect.

While further research is needed to determine whether memory for the story scaffolds may matter for shorter lists, our findings from both experiments provide evidence that scaffold-word associations rather than memory for the scaffold itself are a driving factor behind the success of story-based mnemonic scaffolds.

Given our supported null effects from the experiment in Chapter 3, it remains unclear whether autobiographical or fictional stories make better mnemonic scaffolds, and whether the self-reference effect (Symons & Johnson, 1997) plays a role in mnemonic scaffolds. The findings from our experiments described in the previous chapters could be interpreted as hints both for and against a self-reference effect in mnemonic scaffolds. On one hand, the failure of routine activities in providing a mnemonic benefit in Chapter 4 could suggest that self-relevant information might be important for mnemonic scaffolds to be effective. This is because steps of routine activities, even though constructed by the participants, do not contain self-relevant information. In contrast, one may argue that Body Scaffolds, comprised of the learners' own bodies and the Method of Loci, comprised of familiar locations that learners have visited, are intrinsically self-relevant. Accordingly, one may say that the lack of a mnemonic benefit of routine activities (the only scaffold that failed in enhancing memory in our experiments) is due to a lack of self-relevance. On the other hand, the findings that participants who formed scaffold-word associations with either autobiographical or fictional stories saw a memory increase suggests that self-relevant information may not be important, but it is the individual scaffold-word associations that matter, as explained above.

## 5.3 Potential applications for healthy older adults

Aside from a possible self-reference effect in the Autobiographical Story Scaffold, it is a compelling idea that older adults, in particular, might benefit from this method. This is important because memory training in older adults has hit a roadblock since extensive research has shown that older adults are at a disadvantage when using the Method of Loci (Anschutz et al., 1987; Baltes & Kliegl, 1992; Gross & Rebok, 2011; Kliegl et al., 1990; Karbach & Verhaeghen, 2014; Yesavage et al., 1989). There are motivational and neurocognitive factors suggesting that the Autobiographical Story Scaffold might be particularly well suited for older adults. From a motivational perspective, we suggest that the navigational metaphor of the Method of Loci might induce a stereotype effect related to the fear of getting lost with increasing age (Levy, 2003). We hypothesize that this may potentially decrease seniors' self-efficacy when using the Method of Loci. In other words, there is the possibility that classic memory-training studies involving the Method of Loci may have overlooked the power of self-perception of navigation abilities by activating a stereotype threat of an age-related decrease in wayfinding ability (Oladimeji, Kluger, and Caplan, experiment in progress). This may be overcome by autobiographical stories, which may increase motivation and self-efficacy if older adults enjoy and feel competent recalling stories from their own lives.

From a neurocognitive perspective, the engagement of the hippocampus in the Method of Loci (e.g., Fellner et al., 2016) (even though hippocampal engagement seems not to be the reason for the memory-enhancing effect of the Method of Loci, as explained in earlier chapters) may pose a burden to older adults. This is because the hippocampus is one of the first areas affected by age-related memory decline (Galton et al., 2001; den Heijer et al., 2010; Apostolova et al., 2006). If older adults engage their hippocampi in a memory task when using imagined navigation even though it may not support memory, they likely get distracted by an unnecessary secondary task. If this secondary task decreases their self-efficacy by activating a stereotype effect related to the fear of getting lost and lowers their belief in their own navigational abilities, the memory benefit from the Method of Loci may even be worse. In addition to not posing a stereotype thread and potentially increasing motivation, remote autobiographical memories are among the longest preserved in aging, even in age-related cognitive decline and early-stage Alzheimer's Disease as those memories are retrieved by the neocortex rather than the hippocampus (Cabeza & St Jacques, 2007). Given these factors, further research on how to use the Autobiographical Story Scaffold in memory training for older adults seems promising.

In sum, if certain conditions are met, story-based mnemonic scaffolds can enhance memory and show potential for being particularly promising for older adults.

## 5.4 Not all types of prior knowledge make effective mnemonic scaffolds

Given theories and converging empirical evidence that new information is particularly fast and reliably learned when anchored onto prior knowledge (e.g., Skotko et al., 2004; Sommer, 2017; Tse et al., 2007, 2011), we had hypothesized that all scaffolds comprised of prior knowledge should provide a mnemonic benefit. Our experiments have shown that this is not the case: prior knowledge comprised of steps of routine activities does not enhance memory. When list length exceeds a certain level, the overall mnemonic benefit of narrative prior knowledge for novice learners is cancelled out, even though participants who form scaffold-word associations benefit from those. This suggests that the level to which prior knowledge facilitates memory in first-time users of mnemonic scaffolds is dependent on characteristics of both the scaffold and the learner. At face value, this result seems to contradict Script Theory because this theory poses that our knowledge is structured around stereotypical situations that serve as knowledge stores for new memories (Bartlett, 1932; Schank & Abelson, 1977; Abelson, 1981). When taking a closer look, however, there may not be a real contradiction. In Script Theory (Bartlett, 1932; Schank & Abelson, 1977; Abelson, 1981), in which routine activities and scripts are proposed as memory stores forming the basis for new memories surrounding a related topic (Bower, 1970), there is a key difference from routine activities used as mnemonic scaffolds: In Script Theory, routine activities are seen as the basis for information that is semantically or contextually related (Bartlett, 1932; Schank & Abelson, 1977; Abelson, 1981). In mnemonic scaffolds, in contrast, routine activities are used as anchors for unrelated new knowledge. Thus, our findings do not contradict Script Theory but indicate that routine activities do not enhance memory for *unrelated* information. This might be due to routine activities consisting of actions, which in contrast to locations, body parts or objects are abstract and dynamic and thus difficult to associate new information with.

As we only had first-time scaffold users as participants, we do not know whether more experienced memorizers would be able to use Routine Activity Scaffolds successfully. If more experienced learners have figured out which features of the study items and the anchors within the scaffold to attend to in order to form reliable scaffold-word associations, they might see a mnemonic benefit. We did not observe an improvement across lists within one experimental session but expect that mastery would take place over a longer time frame.

In sum, while the prior knowledge effect likely contributes effectiveness of mnemonic scaffolds, this effect cannot be generalized to all types of prior knowledge and may depend on particular characteristics of the scaffolds and the experience level of the learner.

## 5.5 The Body Scaffold is equally effective as the Method of Loci and does not benefit from embodied cognition

We have identified the Body Scaffold, which to our knowledge, has not been previously investigated, as an alternative to the renowned Method of Loci. This is not only exciting given that the Method of Loci is not suitable for older adults (Anschutz et al., 1987; Baltes & Kliegl, 1992; Gross & Rebok, 2011; Kliegl et al., 1990; Karbach & Verhaeghen, 2014; Yesavage et al., 1989) and possibly younger adults with low visuospatial aptitude (Sanchez, 2019). This is also relevant because the Body Scaffold is a lot more practical than the Numerical Peg System that has been directly compared to the Method of Loci and resulted in similar levels of memory increase (Roediger & Crowder, 1976). Whereas the Numerical

Peg System requires the learner to learn the pegs (usually objects rhyming with or having a similar shape as numbers) first, before being able to use them to learn new information, body part anchors do not need to be pre-learned. This makes it less practical. Moreover, the Numerical Peg System is restricted to the number of pegs that have been pre-learned, while the Body Scaffold can be adapted to fit longer lists by selecting more body parts. These practical reasons strongly favour the Body Scaffold over the Numerical Peg System. Additionally, the fact that the order of body parts is fixed and can simply be retrieved by looking at one's body while the order of locations in a route can be interchangeable favours the Body Scaffold over the Method of Loci. As explained in Chapter 4, a similarity between the Body Scaffold and the Method of Loci is that both scaffolds have a relatively fixed internal order. The order of anchors in the Body Scaffold can be retrieved by looking at one's body. The order of anchors in the Method of Loci can be retrieved following a fixed route without backtracking. This might be why those scaffolds outperform Autobiographical Story Scaffolds, as autobiographical events are not reliably retrieved in a chronological order (e.g., E. F. Loftus & Fathi, 1985). On the other hand, Chapter 3 found no effect of story-recall accuracy (strict scoring for order) on serial-recall.

The experiments in Chapter 4 have shown that the Body Scaffold is effective regardless of whether learners sit still or engage their bodies. It is difficult to relate our findings to previous research on embodied cognition and memory because previous studies did not involve anchoring study items to body parts but embodiment in the sense of moving one's whole body during study (Tuena et al., 2017; Plancher et al., 2013; Jebara et al., 2014; Sauzéon et al., 2011). Some of those studies, (e.g., Tuena et al., 2017) found that different levels of embodiment do not affect memory in, while others (Plancher et al., 2013; Jebara et al., 2014; Sauzéon et al., 2011) endorsed memory-related embodied cognition theories.

From a neurological perspective, it has been shown that verbal processing of concrete stimuli involves the same sensorimotor neural correlates that are active during physical interaction with the object or entity itself (Zwaan, 2004; Pulvermüller, 2005; Fischer & Zwaan, 2008; Toni et al., 2008; Sakreida et al., 2013). Our findings from Chapter 4 have shown that adding sensorimotor perception to verbal processing of the study items does not enhance memory. This supports the notion that the internal representation of a verbal stimulus through simulation (Gallese, 2008; Zwaan, 2004) is an intrinsic part of cognition because additional sensorimotor perception is not required for a stimulus to be mentally processed. In the case of our second experiment in Chapter 4, adding sensorimotor perception to the verbal processing of the scaffold-word associations was not necessary for the same process to occur when the task is performed in one's mind. Despite this, there remains the possibility of a trade-off effect of directing attention to the body parts and away from the study items shown on the screen. That is, more time spent interacting with the body parts implies less time spent looking at the study item. In other words, if there was a benefit of touching the body parts during study, it might be cancelled out by diverted attention.

Together, our experiments on the Body Scaffold have shown that body parts are equally effective mnemonic scaffolds as locations along a familiar path. Embodied cognition may be irrelevant to the effectiveness of the Body Scaffold in much the same way that imagined navigation might be irrelevant to the effectiveness of the Method of Loci (Caplan et al., 2019; Bower, 1970). We discuss this next.

## 5.6 The effectiveness of the Method of Loci may not be based on imagined navigation

Regarding the question of whether the effectiveness of the Method of Loci is based on imagined navigation, the experiment in Chapter 2 and the second experiment in Chapter 4 have yielded conflicting findings. In the experiments in Chapter 2 (list length ten) and Chapter 3 (list length fifteen), we counted the number of locations in the stories and asked whether those render the story scaffolds more effective. Our motivation behind this was twofold. We were interested whether a) spatial cognition, which according to some researchers (e.g., Dresler et al., 2017; Fellner et al., 2016) underlies the effectiveness of the Method of Loci, contributes to the effectiveness of the story-based scaffolds, and b) whether instructions for story scaffolds can be improved by asking learners to make use of spatial features in their stories.

With lists of ten words, we found that participants benefited from spatial locations mentioned in their stories in the group that used the most effective presentation mode of viewing one word and one sentence at a time. This cautiously suggests that spatial features may contribute to the effectiveness of mnemonic scaffolds. We need to interpret this relationship with caution because no correlation between the number of locations mentioned and memory success was found in the remaining groups (nor in the experiment with list length fifteen in Chapter 3, in which list length was likely the limiting factor in the effectiveness of the story scaffolds).

Our main finding that the Body Scaffold is as effective as the Method of Loci from Chapter 4 challenges the assumption that spatial cognition is a driving factor behind the effectiveness of the Method of Loci. This converges with evidence that although spatial cognition and the engagement of brain areas associated with navigation are also active when using the Method of Loci (e.g., Fellner et al., 2016), these are not causally related to its mnemonic benefit (e.g., Bower, 1970; Caplan et al., 2019). From this perspective, it is important to note that it may not be the spatial features of locations invoking spatial cognition that render stories with more locations more effective scaffolds than stories with fewer locations. Instead, the fact that locations or objects along a familiar route are, like body parts high-imagery, concrete entities, may explain why stories with more such entities make better scaffolds than stories with fewer concrete entities. This assumption is supported by findings that high-imagery words are generally better remembered than low-imagery words (e.g., Caplan & Madan, 2016; Paivio, 1971). In other words, it seems plausible that this imagery effect may translate to word-anchor associations in that concrete, high-imagery anchors are more effective than low-imagery anchors.

Even though we found some evidence that spatial locations increase the effectiveness of the Autobiographical Story Scaffold, our findings overall challenge the notion that the effectiveness of the Method of Loci is based on imagined navigation. Instead, features of scaffolds consisting of locations as in the Method of Loci that are shared with non-navigational mnemonic scaffolds as the Body Scaffold may drive their effectiveness.

## 5.7 Individual differences and the effectiveness of mnemonic scaffolds

In Chapter 4, we administered three individual differences questionnaires to test whether individual differences in visual ability and body responsiveness contribute to the effectiveness of mnemonic scaffolds. In both experiments of Chapter 4, individual differences did not differentiate levels of recall accuracy. This leaves open the possibility that embodiment, vividness of visual imagery, and visuospatial aptitude are all necessary, but at such a minimal level that the corresponding domain-affinity or skill makes little difference. Below, we discuss our findings regarding our individual differences measures individually.

### 5.7.1 Individual differences in body responsiveness may not contribute to the effectiveness of the Body Scaffold

Given the embodiment component of the Body Scaffold, we hypothesized that participants with high scores in the BRQ would benefit more from this method than participants with low BRQ scores. This was not confirmed by our experiments in Chapter 4. BRQ scores did not predict any measure of scaffold-dependent recall accuracy in any group neither in Experiment 1 nor in Experiment 2 of Chapter 4. Similarly, as shown in Experiment 2 of Chapter 4 and explained earlier in the present Chapter, the Body Scaffold is robust to variations in bodily engagement. There remains the possibility that body responsiveness may influence the effectiveness with which participants use the Body Scaffold, but the BRQ might not be sensitive enough to capture the possibly relevant factors.

### 5.7.2 Individual differences in vividness of visual imagery may not contribute to the effectiveness of mnemonic scaffolds

Our findings do not only challenge the notion that imagined navigation underlies the effectiveness of the Method of Loci (Dresler et al., 2017; Fellner et al., 2016; Foer, 2011), they also challenge the widely held assumption that visual imagery is key to effective use of mnemonic scaffolds (Foer, 2011; Konrad, 2013). This assumption has been popularized in both historical and contemporary memory training literature (Foer, 2011; Yates, 1966). During the Method of Loci, for example, memory athletes reportedly transform the to-be-remembered information into a vivid image involving the location or object along a familiar route. If the assumption that vivid visual imagery is essential to use mnemonic scaffolds was true, one would expect that learners with high vividness of visual imagery would be more adept at using mnemonic scaffolds than learners with low vividness of visual imagery. Contrary to this, we found a supported null effect of self-reported vividness of visual imagery and the success with which participants use mnemonic scaffolds. All self-report measures need to be interpreted with caution, but since vividness of visual imagery can to date not be objectively measured we had no other choice than to use the best-investigated self-report tool, the VVIQ (Marks, 1973) even though its construct validity has been challenged (McKelvey, 1995) and reliance on self-report itself is problematic. Two previous studies on the Method of Loci in younger (McKellar, Marks and Barron reported by Marks, 1972b) and older adults (Kliegl et al., 1990) have also found no relationship between VVIQ scores and memory enhancement.

Aside from the VVIQ, findings that congenitally blind participants can perform well with the Method of Loci (De Beni & Cornoldi, 1985) compellingly support the notion that the effectiveness of mnemonic scaffolds does not rely on forming vivid mental images. This leaves open two possibilities. First, visual imagery might not be detrimental to memory, and second, it might actually be beneficial, but we were not able to measure its benefit with the questionnaires we used.

In addition Thompson et al. (1991) have shown that a person with superior memory for the digits of pi uses scaffolds of prior knowledge to memorize the digits of pi without evidence of using visual imagery. This further suggests that visual imagery may not be required for successful use of mnemonic scaffolds.

In sum, the claim of memory experts that forming vivid mental images is necessary when using mnemonic scaffolds should not be taken at face value. Converging with evidence from congenitally blind individuals and some memory champions, our findings suggest that vivid imagery does not contribute to the effectiveness of mnemonic scaffolds.

### 5.7.3 Individual differences in visuospatial aptitude may not contribute to the effectiveness of mnemonic scaffolds

In addition to the VVIQ, we used the PFT as an objective measure of spatial visual imagery ability. We replicated Sanchez' (2019) findings that PFT scores predicted recall accuracy when using the Method of Loci. Importantly, however, as described in Chapter 4, our interpretation of this correlation diverges Sanchez' (2019). This is because we did not only find a positive relationship with PFT scores and post-instruction serial recall accuracy, but also with pre-instruction serial recall accuracy and recall accuracy in the Control. As the PFT consists of problems that get progressively more difficult, it likely measures task engagement. Participants who put effort into solving the PFT problems likely put a similar level of engagement into studying the words with and without the strategy. We, therefore, suppose that PFT scores reflect motivational factors and compliance levels rather than a relationship between visuospatial aptitude and mnemonic benefit. Consequently, to test a possible relationship between visuospatial aptitude and effectiveness of the Method of Loci, future research needs to incorporate measures of visuospatial aptitude that are not confounded with levels of engagement and motivation.

Together, our experiments suggest that participants' motivation and engagement are important to consider when testing the effect of individual differences in the effectiveness of mnemonic scaffolds.

## 5.8 Effective use of mnemonic scaffolds as a skill

Effective use of mnemonic techniques can be considered a skill that expert memorizers have mastered. In Skilled Memory Theory Ericsson and colleagues (1981; 1989; 2003) sought to provide a comprehensive account of how effective use of mnemonic techniques is a skill acquired through training. Even though our experiments were not focused on the training aspect as our participants were novice learners who were introduced to mnemonic scaffolds for the first time, our findings challenge multiple aspects of Skilled Memory Theory (Chase & Ericsson, 1981; Ericsson & Staszewski, 1989; Ericsson, 2003). Ericsson and colleagues identified three principles that explain superior memory performance in a broad range of expertise contexts (Chase & Ericsson, 1981; Ericsson & Staszewski, 1989; Ericsson, 2003): (a) retrieval structures, (b) meaningful encoding, and (c) speed-up by practice. Below, we will discuss in how far our previous chapters call for a modification of these principles.

### 5.8.1 Mnemonic scaffolds rather than retrieval structures

Skilled Memory Theory (Chase & Ericsson, 1981; Ericsson & Staszewski, 1989; Ericsson, 2003) is strongly influenced by research on expert performance. The direct application of some concepts of expert performance to memory skill acquisition glosses over some important aspects of memory training. This becomes evident in how Ericsson and colleagues (1981; 1989; 2003) have described retrieval structures as one of their three principles of skilled memory. Ericsson and colleagues' (1981; 1989; 2003) concept of retrieval structures is partially derived from their observation that experts with ample knowledge in their domain of expertise acquire new information within this domain much faster than information outside their domain of expertise (e.g., Chase & Ericsson, 1982, 1981; Staszewski, 1990).

According to Skilled Memory Theory (Chase & Ericsson, 1981; Ericsson & Staszewski, 1989; Ericsson, 2003), the breadth and depth of knowledge within a specific domain of expertise promotes meaningful encoding of new information. The semantic and conceptual overlap between learned and new information within a specific domain of expertise facilitates those associations (e.g., Chase & Ericsson, 1982, 1981; Staszewski, 1990). It is important to note that Ericsson and colleagues (1981; 1989; 2003) based their theory on observations in experts with domain-specific knowledge and memory expertise acquired through massive practice. It stands to reason that different mechanisms may contribute to superior memory in those two groups. The relevance of a distinction between "retrieval structures" in expert knowledge acquisition and "mnemonic scaffolds" in memory skill development becomes clear when comparing skilled memorizers to experts expanding their domain-specific knowledge.

Experts who show superior memory performance within their domain of expertise use existing knowledge to unintentionally associate semantically related information. In contrast, skilled memorizers who show superior memory performance when learning arbitrary information intentionally use pre-learned mnemonic scaffolds to associate semantically unrelated information. During study, mnemonic scaffolds provide a set of anchors or pegs for associations between existing memory representations and semantically unrelated study items. During recall, these serve as a system of retrieval cues making the encoded information rapidly and reliably accessible (Roediger, 1980). The following example elucidates the importance of differentiating between mnemonic scaffolds and retrieval structures: Chess-related knowledge providing a memory advantage for expert chess players (Chase & Simon, 1973; Lane & Chang, 2018) is quite different from a set of objects and locations in an environment providing the memory substrate for the Method of Loci.

To disentangle domain-specific knowledge acquisition from skilled memory, we use the term mnemonic scaffolds for six reasons: (a) to denote pre-learned scaffolds of memories rather than existing knowledge within a specific domain, (b) to describe that mnemonic scaffolds are artificially built or intentionally modified memory structures rather than naturally acquired knowledge, (c) to denote that mnemonic scaffolds are used intentionally while retrieval structures provide a memory benefit even when used unintentionally, (d) to differentiate that mnemonic scaffolds are used to encode semantically unrelated information, while retrieval structures when unintentionally used within a knowledge domain only provide a memory benefit for semantically related information, (e) to denote the distinction between serial encoding and recall in mnemonic strategies and incidental encoding and free recall in knowledge structures, and finally (e) to avoid terminological confusion with the word retrieval: mnemonic scaffolds play an equally important role both during study (providing anchors for associations) and during retrieval (providing retrieval cues).

### 5.8.2 Individual scaffold-word associations rather than meaningful encoding

Ericsson and colleagues (1981; 1989; 2003) derived the concept of meaningful encoding from Craik and Lockhart's (1972) Levels of Processing Framework. According to this framework, meaningful (also called "elaborative" or "deep," Craik & Lockhart, 1972) encoding tasks facilitate memory by guiding the learner in studying new information in a cognitively engaging way. In Skilled Memory Theory (Chase & Ericsson, 1981; Ericsson & Staszewski, 1989; Ericsson, 2003), the mnemonic advantage is understood in terms of the assumption that retention is an increasing function of the amount of elaboration that study items receive during encoding (Chase & Ericsson, 1981; Ericsson & Staszewski, 1989; Ericsson, 2003; Klein & Kihlstrom, 1986). Encoding tasks that are thought to promote deeper processing of the study items can be viewed as encouraging the learner to engage in greater trace elaboration (Anderson et al., 1976). In their Skilled Memory Theory, Ericcson and colleagues (1981; 1989; 2003) refer to the Levels of Processing framework without further conceptualizing meaningful encoding. While Levels of Processing has had a profound influence on modern memory research, it is important to note that the link between "deep", "elaborative", or "meaningful" processing and retention may be correlational rather than causal (Nairne, 2002). In addition, the fact that those terms are used interchangeably and cannot be operationalized in a way that can be experimentally tested highlights the fact that Levels of Processing Theory (Craik & Lockhart, 1972) and its application to Skilled Memory Theory (Chase & Ericsson, 1981; Ericsson & Staszewski, 1989; Ericsson, 2003) is circular and not useful to pinpoint driving factors of skilled memory.

In our experiments, we have identified individual scaffold-word associations to be a driving factor behind the effectiveness of mnemonic scaffolds. In all experiments, scaffold-cued recall was a strong predictor of serial recall. This suggests that participants who have formed individual associations between parts of the scaffold and study items benefited from those during serial recall. Participants who did not associate study items with parts of their scaffolds did not receive a mnemonic benefit. This effect was observed across different types of prior knowledge the scaffolds were comprised of. Therefore, individual scaffold-word associations are likely a driving force behind the mnemonic benefit. Given that, additional training on how to associate study items with anchors of the scaffolds may be a promising next step to increase the effectiveness in mnemonic scaffolds in novice learners. Note that our participants had no prior experience with the use of mnemonic techniques. With training, learners will likely become more skillful at forming scaffold-word associations, suggesting that this is, indeed, a principle behind successful use of mnemonic scaffolds as a memory skill.

Importantly, our scaffold-cued recall task has allowed us to test the importance of scaffold-word associations for memory success empirically. This is an important advancement from the circular and unoperationalized concept of meaningful encoding in Skilled Memory Theory (Chase & Ericsson, 1981; Ericsson & Staszewski, 1989; Ericsson, 2003). Consequently, contemporary theories of skilled memory need to move away from meaningful encoding as a principle of a memory skill and instead account for individual scaffold-word associations as a driving force.

### 5.8.3 Interdependent components of skilled memory rather than speed up by practice

The principle of speed up by practice in Skilled Memory Theory (Chase & Ericsson, 1981; Ericsson & Staszewski, 1989; Ericsson, 2003) is derived from the observation that speed of encoding and retrieval processes is increased with extended practice. According to Ericsson and colleagues (1981; 1989; 2003) it is practice that allows trained memorizers to rapidly store long lists of items of a certain type of material in memory for subsequent recall. Given that we neither know which aspects of practice are important for memory success, nor what trained memorizers do when they practice, attributing memory skill to getting faster at memorizing through practice draws an incomplete picture. In other words, describing speed up by practice as a principle of skilled memory seems circular and completely misses the point that the skill to successfully use mnemonic scaffolds may consist of interdependent components that could be practiced in different stages throughout the training process. Even though we cannot draw direct inferences from our single-session experiments with novice learners, it stands to reason that the skill to form individual scaffold-word associations may be one such component of the memory skill. A second component may be the skill to keep track of order within the scaffold. This is supported by our finding that the Method of Loci and the Body Scaffold (both of which have a clear, prescribed order) are superior to story-based and Routine Activity Scaffolds (the order of which is more variable).

In sum, contemporary accounts of skilled memory and empirically grounded memory training protocols should account for the possibility that the memory skill consists of interdependent aspects rather than generalizing that practice per se is a principle of skilled memory.

## 5.9 Limitations

Even though our findings have expanded our understanding of how mnemonic scaffolds can enhance memory, the general inferences we can draw from our experiments to real-world learning scenarios are limited by several factors. First, the study material — standardized word lists — is very different from real-word study material and largely dependent on the learning topic and context. This might affect learning in various ways and make the scaffolds either harder or easier to use, depending on the context. Second, the motivation of research participants who use mnemonic scaffolds to receive course credit or a payment for their participation in the experiment is different from the motivation of real-world learners who use mnemonic scaffolds because they need or want to learn new material. If motivation to learn is high because learners actually want to learn the study material they are likely to use mnemonic scaffolds more effectively than test participants in experiments. Third, it is important to note that we used the presentation mode that worked best for Autobiographical Story Scaffold for the other scaffolds we investigated. There is the possibility that different types of scaffolds require different presentation modes in order to be effective. Fourth, our individual differences measures were either based on self-report or likely confounded with participants' individual levels of engagement and motivation. Fifth, and perhaps most importantly when it comes to drawing inferences from our findings to the development of memory training protocols is the fact that we conducted single-session experiments with first-time users. This is why we need to integrate the training aspect into future studies on mnemonic scaffolds, as explained below.

In addition, only a few studies (for a meta analysis, see Twomey & Kroneisen, 2021) have used randomized control trials or checked participants' compliance. Therefore, a standardized way to instruct the Method of Loci and to check compliance levels is important.

## 5.10 Future research

In order to understand mechanisms behind successful use of mnemonic scaffolds and in order to develop empirically grounded memory training protocols, future research should address several factors, as listed below.

First, researchers need to move away from paid participants to real-world learners and account for differences in their motivation to use mnemonic scaffolds. Second, researchers also need to include real-world learning material and investigate dependencies with the type of study material. Third, in order to develop memory training protocols and to understand the mechanisms behind successful use of mnemonic scaffolds, it is necessary to disentangle various aspects of training, such as individual scaffold-word associations and order within the scaffold, and investigating co-dependencies. Fourth, to understand the mechanisms behind successful use of mnemonic scaffolds from a neurological perspective and relate those to training protocols, researchers need to track learning curves and pinpoint behavioural and neurological changes that occur over the course of the training and result in an increase in memory performance. Fourth, to address the caveats with the individual differences measures we described earlier, future research should investigate individual differences with objective measures with high construct validity rather than relying on self-report. Fifth, future research should test boundary conditions with regard to different types of scaffolds, characteristics of the task, and characteristics of the learning environment.

## 5.11 Conclusion

Together, we have learned from our studies that mnemonic scaffolds can enhance memory in novice learners in only one single training session. Not all scaffolds are equally effective. Importantly, we have identified the Body Scaffold as an alternative for learners who struggle with the Method of Loci. Autobiographical Story Scaffolds also provide a mnemonic benefit. The ability to form individual scaffold-word associations is a driving factor behind successful use of mnemonic scaffolds. Embodied cognition, imagined navigation, and visual imagery aptitude may not contribute to their effectiveness. With further fine-tuning of the scaffolds and they way they are instructed, mnemonic scaffolds can be used to greatly boost learning performance and enhance memory in education and ageing populations. To exploit the full potential of mnemonic scaffolds, future research needs to integrate experiments of mnemonic scaffolds into daily life of real-world learners, track their learning curves and pinpoint changes in learning behaviour that predict increase in memory performance.

# Bibliography


Abelson, R. P. (1981). Psychological status of the script concept. *American Psychologist*, *36*(7), 715–729. doi: 10.1037/0003-066X.36.7.715

Alba, J. W., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin*, *93*(2), 203–231.

Anderson, M. C., Pichert, J. W., Goetz, E. T., Schallert, D. L., Stevens, K. V., & Trollip, S. R. (1976). Instantiation of general terms. *Journal of Verbal Learning and Verbal Behavior*, *15*(6), 667–679.

Anschutz, L., Camp, C. J., Markley, R. P., & Kramer, J. J. (1987). Remembering mnemonics: A three-year follow-up on the effects of mnemonics training in elderly adults. *Experimental Aging Research*, *13*(3), 141–143. doi: 10.1080/03610738708259315

Apostolova, L. G., Dutton, R. A., Dinov, I. D., Hayashi, K. M., Toga, A. W., Cummings, J. L., & Thompson, P. M. (2006). Conversion of Mild Cognitive Impairment to Alzheimer Disease Predicted by Hippocampal Atrophy Maps. *Archives of Neurology*, *63*(5), 693. doi: 10.1001/archneur.63.5.693

Arbuckle, T. Y., Vanderleck, V. F., Harsany, M., & Lapidus, S. (1990). Adult age differences in memory in relation to availability and accessibility of knowledge-based schemas. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(2), 305–315.

Baltes, P. B., & Kliegl, R. (1992). Further Testing of Limits of Cognitive Plasticity: Negative Age Differences in a Mnemonic Skill Are Robust. *Developmental Psychology*, *28*(1), 121–125. doi: 10.1037/0012-1649.28.1.121

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4), 577–609. doi: 10.1017/S0140525X99002149

Barsalou, L. W. (2008). Grounded Cognition. *Annual Review of Psychology*, *59*(1), 617–645. doi: 10.1146/annurev.psych.59.103006.093639

Bartlett, F. C. (1932). *Remembering.* Cambridge: Cambridge University Press.

Bass, W., & Oswald, K. (2014). Proactive control of proactive interference using the method of loci. *Advances in cognitive psychology*, *10*(2), 49–58. doi: 10.5709/ACP-0156-3

Becchetti, A. (2010, 6). Hippocampal formation and the classical art of memory. *Proceedings of the National Academy of Sciences*, *107*(25), E104. doi: 10.1073/pnas.1004950107

Bellezza, F. S. (1981). Mnemonic Devices: Classification, Characteristics, and Criteria. *Review of Educational Research*, *51*(2), 247–275.

Bellezza, F. S. (1983). The spatial-arrangement mnemonic. *Journal of Educational Psychology*, *75*, 830–837.

Bellezza, F. S. (1984). The Self as a Mnemonic Device: The Role of Internal Cues. *Journal of Personality and Social Psychology*, *47*(3), 506–516. doi: 10.1037/0022-3514.47.3.506

Bellezza, F. S. (1986). A mnemonic based on arranging words on visual patterns. *Journal of Educational Psychology*, *78*(2), 217–224.

Bellezza, F. S., & Bower, G. H. (1981). The representational and processing characteristics of scripts. *Bulletin of the Psychonomic Society*, *18*(1), 1–4. doi: 10.3758/BF03333553

Bellezza, F. S., & Buck, D. K. (1988). Expert knowledge as mnemonic cues. *Applied Cognitive Psychology*, *2*(2), 147–162. doi: 10.1002/acp.2350020206

Bellmund, J. L., Deuker, L., Schröder, T. N., & Doeller, C. F. (2016). Grid-cell representations in mental simulation. *eLife*, *5*. doi: 10.7554/eLife.17089

Blajenkova, O., Motes, M. A., & Kozhevnikov, M. (2005). Individual differences in the representations of novel environments. *Journal of Environmental Psychology*, *25*(1), 97–109. doi: 10.1016/j.jenvp.2004.12.003

Bodner, G. E., & MacLeod, C. M. (2016). The Benefits of Studying by Production . . . and of Studying Production: Introduction to the Special Issue on the Production Effect in Memory. *Canadian Journal of Experimental Psychology*, *70*(2), 89–92. doi: 10.1037/cep0000094

Bodner, G. E., & Taikh, A. (2012). Reassessing the basis of the production effect in memory. *Journal of Experimental Psychology: Learning Memory and Cognition*, *38*(6), 1711–1719. doi: 10.1037/a0028466

Bouffard, N., Stokes, J., Kramer, H. J., & Ekstrom, A. D. (2017). Temporal encoding strategies result in boosts to final free recall performance comparable to spatial ones. *Memory and Cognition*, *46*(1), 17–31. doi: 10.3758/s13421-017-0742-z

Bower, G. H. (1970). Analysis of a mnemonic device: modern psychology uncovers the powerful components of an ancient system for improving memory. *American Scientist*, *58*(5), 496–510.

Bower, G. H., & Clark, M. C. (1969). Narrative stories as mediators for serial learning. *Psychonomic Science*, *14*, 181–182.

Bower, G. H., & Gilligan, S. G. (1979). Remembering information related to one's self. *Journal of Research in Personality*, *13*(4), 420–432. doi: 10.1016/0092-6566(79)90005-9

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436.

Brandt, K. R., Cooper, L. M., & Dewhurst, S. A. (2005). Expertise and recollective experience: recognition memory for familiar and unfamiliar academic subjects. *Applied Cognitive Psychology*, *19*(9), 1113–1125. doi: 10.1002/acp.1163

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated Cognition and the Culture of Learning. *Educational Researcher*, *18*(1), 32–42. doi: 10.3102/0013189X018001032

Bruett, H., Fang, X., Kamaraj, D. C., Haley, E., & Coutanche, M. N. (2018). Expertise moderates incidentally learned associations between words and images. *Frontiers in Psychology*, *9*(2085). doi: 10.3389/fpsyg.2018.02085

Cabeza, R., & St Jacques, P. (2007). Functional neuroimaging of autobiographical memory. *Trends in Cognitive Sciences*, *11*(5), 219–227. doi: 10.1016/J.TICS.2007.02.005

Campos, A., Amor, A., & González, M. A. (2004). The Importance of the Keyword-Generation Method in Keyword Mnemonics. *Experimental Psychology*, *51*(2), 125–131. doi: 10.1027/1618-3169.51.2.125

Caplan, J. B., Legge, E. L. G., Cheng, B., & Madan, C. R. (2019). Effectiveness of the method of loci is only minimally related to factors that should influence imagined navigation. *Quarterly Journal of Experimental Psychology*, *72*(10), 61–64.

Caplan, J. B., & Madan, C. R. (2016). Word Imageability Enhances Association-memory by Increasing Hippocampal Engagement. *Journal of Cognitive Neuroscience*, *20*(10), 1522–1538. doi: 10.1162/jocn

Carey, B. (2014). Remembering, as an Extreme Sport. *The New York Times*, D1.

Carr, T. H., Dagenbach, D., VanWieren, D., Carlson-Radvansky, L. A., A, A., & Brown, J. (1994). Acquiring general knowledge from specific episodes of experience. In:. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance xv: Conscious and nonconscious information* (p. 697–724). The MIT Press.

Chai, X. J., & Jacobs, L. F. (2010). Effects of cue types on sex differences in human spatial memory. *Behavioural Brain Research*, *208*(2), 336–342. doi: 10.1016/j.bbr.2009.11.039

Chase, W. G., & Ericsson, K. A. (1981). Skilled memory. *Cognitive skills and their acquisition*, 141–189.

Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (pp. 1–58). San Diego: Academic Press.

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (p. 215–281). New York, NY: Academic Press.

Cornoldi, C., & De Beni, R. (1991). Memory for discourse: Loci mnemonics and the oral presentation effect. *Applied Cognitive Psychology*, *5*(6), 511–518. doi: https://doi.org/10.1002/acp.2350050606

Coutanche, M. N., Thompson-Schill, S. L., Sharon, T., Moscovitch, M., & Gilboa, A. (2014). Fast mapping rapidly integrates information into existing memory networks. *Journal of Experimental Psychology: General*, *108*(3), 2296–2303. doi: 10.1073/pnas.1005238108

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684. doi: 10.1016/S0022-5371(72)80001-X

Craik, F. I., & Tulving, E. (1975, 9). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*(3), 268–294. doi: 10.1037/0096-3445.104.3.268

Daubenmier, J. J. (2005). The relationship of yoga, body awareness, and body responsiveness to self-objectification and disordered eating. *Psychology of Women Quarterly*, *29*(2), 207–219.

Daubenmier, J. J., Sze, J., Kerr, C. E., Kemeny, M. E., & Mehling, W. (2013). Follow your breath: respiratory interoceptive accuracy in experienced meditators. *Psychophysiology*, *50*(8), 777–789.

De Beni, R., & Cornoldi, C. (1985). The effects of imaginal mnemonics on congenitally total blind and on normal subjects. In D. F. Marks & D. G. Russell (Eds.), *Imagery 1.* Dunedin, N.Z.: Imagery.

den Heijer, T., van der Lijn, F., Koudstaal, P. J., Hofman, A., van der Lugt, A., Krestin, G. P., ... Breteler, M. M. B. (2010). A 10-year follow-up of hippocampal volume on magnetic resonance imaging in early dementia and cognitive decline. *Brain*, *133*(4), 1163–1172. doi: 10.1093/brain/awq048

Dresler, M., Shirer, W. R., Konrad, B. N., Müller, N. C., Wagner, I. C., Fernández, G., ... Greicius, M. D. (2017). Mnemonic Training Reshapes Brain Networks to Support Superior Memory. *Neuron*, *93*(5), 1227–1235. doi: 10.1016/j.neuron.2017.02.003

Drewnowski, A., & Murdock, B. B. (1980). The role of auditory features in memory span for words. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(3), 319–332. doi: 10.1037/0278-7393.6.3.319

Eichenbaum, H. (2014). Time cells in the hippocampus: a new dimension for mapping memories. *Nature reviews. Neuroscience*, *15*(11), 732. doi: 10.1038/NRN3827

Ekstrom, A. D., Caplan, J. B., Ho, E., Shattuck, K., Fried, I., & Kahana, M. J. (2005). Human hippocampal theta activity during virtual navigation. *Hippocampus*, *15*(7), 881–889. doi: 10.1002/HIPO.20109

Ericsson, K. A. (2003). Exceptional memorizers: made, not born. *TRENDS in Cognitive Sciences*, *7*(6), 233–235. doi: 10.1016/S1364-6613(03)00103-7

Ericsson, K. A., Chase, W. G., Faloon, S., Ericcson, K. A., Chase, W. G., & Faloon, S. (1980). Acquisition of a memory skill. *Science*, *208*(4448), 1181–2. doi: 10.1126/SCIENCE.7375930

Ericsson, K. A., Delaney, P. F., Weaver, G., & Mahadevan, R. (2004). Uncovering the structure of a memorist's superior "basic" memory capacity. *Cognitive Psychology*, *49*(3), 191–237. doi: 10.1016/j.cogpsych.2004.02.001

Ericsson, K. A., & Kintsch, W. (1995). Long-Term Working Memory. *Psychological Review*, *102*(2), 211–245.

Ericsson, K. A., & Staszewski, J. J. (1989). Skilled memory and expertise: Mechanisms of exceptional performance. In D. Klahr & D. Kotovsky (Eds.), *Complex information processing: The impact of herbert a. siomon* (p. 235–267). Lawrence Erlbaum Associates, Inc.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. doi: 10.3758/BF03193146

Fellner, M.-C., Volberg, G., Wimber, M., Goldhacker, M., Greenlee, M. W., & Hanslmayr, S. (2016). Spatial Mnemonic Encoding: Theta Power Decreases and Medial Temporal Lobe BOLD Increases Co-Occur during the Usage of the Method of Loci. *eNeuro*, *3*(6), 0184–16. doi: 10.1523/ENEURO.0184-16.2016

Fischer, M. H., & Zwaan, R. A. (2008). Embodied language: A review of the role of the motor system in language comprehension. *Quarterly Journal of Experimental Psychology*, *61*(6), 825–850. doi: 10.1080/17470210701623605

Foer, J. (2011). *Moonwalking with Einstein: The art and science ofremembering everything.* New York, NY: Penguin Press.

Fraundorf, S. H., Diaz, M. I., Finley, J. R., Lewis, M. L., Tooley, K. M., Isaacs, A. M., ... Brehm, L. (2014). *CogToolbox for MATLAB.*

French, J. W., Ekstrom, R. B., & Price, L. A. (1963). *Kit of reference tests for cognitive factors.* Princeton, N.J: Educational Testing Service.

Gallese, V. (2008). Mirror neurons and the social nature of language: the neural exploitation hypothesis. *Social Neuroscience*, *3*(3-4), 317–333.

Galton, C. J., Patterson, K., Graham, K., Lambon-Ralph, M. A., Williams, G., Antoun, N., ... Hodges, J. R. (2001). Differing patterns of temporal atrophy in Alzheimer's

disease and semantic dementia. *Neurology*, *57*(2), 216–25.
Gesualdo, F. (1592). *Plutosofia*. Padua.
Greenberg, D. L., & Rubin, D. C. (2003). The neuropsychology of autobiographical memory. *Cortex*, *39*(4-5), 687–728. doi: 10.1016/S0010-9452(08)70860-8
Groninger, L. D. (1971). Mnemonic imagery and forgetting. *Psychonomic Science*, *23*(2), 161–163. doi: https://doi.org/10.3758/BF03336056
Gross, A. L., & Rebok, G. W. (2011). Memory Training and Strategy Use in Older Adults: Results from the ACTIVE Study. *Psychology and aging*, *26*(3), 503–517. doi: 10.1037/a0022687
Habermas, T. (2018). *Emotion and narrative: Perspectives in autobiographical storytelling.* Cambridge: Cambridge University Press.
Handy, T. C., Grafton, S. T., Shroff, N. M., Ketay, S., & Gazzaniga, M. S. (2003). Graspable objects grab attention when the potential for action is recognized. *Nature Neuroscience*, *6*(4), 421–427. doi: 10.1038/nn1031
Hebscher, M., Levine, B., & Gilboa, A. (2018). The precuneus and hippocampus contribute to individual differences in the unfolding of spatial representations during episodic autobiographical memory. *Neuropsychologia*, *110*, 123–133. doi: 10.1016/j.neuropsychologia.2017.03.029
Hirst, W., & Echterhoff, G. (2011). Remembering in Conversations: The Social Sharing and Reshaping of Memories. *Annual Review of Psychology*, *63*(1), 55–79. doi: 10.1146/ANNUREV-PSYCH-120710-100340
Horner, A. J., Bisby, J. A., Wang, A., Bogus, K., & Burgess, N. (2016). The role of spatial boundaries in shaping long-term event representations. *Cognition*, *154*, 151–164. doi: 10.1016/j.cognition.2016.05.013
Hu, Y., Ericsson, K. A., Yang, D., & Lu, C. (2009). Superior self-paced memorization of digits in spite of a normal digit span: The structure of a memorist's skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(6), 1426–1442. doi: 10.1037/a0017395
Hunter, I. M. L. (1956). Mnemonic systems and decives. *Science News*, *39*, 75–97.
Ishikawa, T., & Montello, D. R. (2006). Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places. *Cognitive Psychology*, *52*(2), 93–129. doi: 10.1016/j.cogpsych.2005.08.003
Itskov, V., Curto, C., Pastalkova, E., & Buzsáki, G. (2011). Cell assembly sequences arising from spike threshold adaptation keep track of time in the hippocampus. *The Journal of Neuroscience*, *31*(8), 2828. doi: 10.1523/JNEUROSCI.3773-10.2011
JASP Team. (2019). *JASP (Version 0.10.2).*
Jebara, N., Orriols, E., Zaoui, M., Berthoz, A., & Piolino, P. (2014). Effects of enactment in episodic memory: A pilot virtual reality study with young and elderly adults. *Frontiers in Aging Neuroscience*, *6*, 388. doi: 10.3389/fnagi.2014.00338
Kan, I. P., Alexander, M. P., & Verfaellie, M. (2009). Contribution of prior semantic knowledge to new episodic learning in amnesia. *Journal of Cognitive Neuroscience*, *21*(5), 938–944. doi: 10.1162/jocn.2009.21066
Karbach, J., & Verhaeghen, P. (2014). Making Working Memory Work: A Meta-Analysis of Executive-Control and Working Memory Training in Older Adults. *Psychological*

*Science*, *25*(11), 2027–2037. doi: 10.1177/0956797614548725

Keogh, R., & Pearson, J. (2018). The blind mind: No sensory visual imagery in aphantasia. *Cortex*, *105*(2015), 53–60. doi: 10.1016/j.cortex.2017.10.012

Klein, S. B., & Kihlstrom, J. F. (1986). Elaboration, Organization, and the Self-Reference Effect in Memory. *Journal of Experimental Psychology: General*, *115*(1), 26–38. doi: 10.1037/0096-3445.115.1.26

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in psychtoolbox-3. *Perception*, *36*(14), 1–16.

Kliegl, R., Smith, J., & Baltes, P. B. (1990). On the Locus and Process of Magnification of Age Differences During Mnemonic Training. *Developmental Psychology*, *26*(6), 894–904. doi: 10.1037/0012-1649.26.6.894

Kondo, Y., Suzuki, M., Mugikura, S., Abe, N., Takahashi, S., Iijima, T., & Fujii, T. (2005). Changes in brain activation associated with use of a memory strategy: a functional MRI study. *NeuroImage*, *24*(4), 1154–1163. doi: 10.1016/j.neuroimage.2004.10.033

Konrad, B. N. (2013). *Superhirn Gedächtnistraining mit einem Weltmeister*. Vienna: Goldegg Verlag.

Kucera, H., & Francis, W. (1967). *Computational analysis of present-day American English.* Providence, R.I: Brown University Press.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought.* New York, NY: Basic Books.

Lane, D. M., & Chang, Y. H. A. (2018). Chess knowledge predicts chess memory even after controlling for chess experience: Evidence for the role of high-level processes. *Memory and Cognition*, *46*(3), 337–348. doi: 10.3758/s13421-017-0768-2

Lashley, K. S. (1951). The problem of serial order in behavior. In *Cerebral mechanisms in behavior: the hixon symposium.* doi: 10.1016/j.humov.2007.04.001

Lea, G. (1975). Chronometric analysis of the method of loci. *Journal of Experimental Psychology: Human Perception and Performance*, *1*(2), 95–104. doi: https://doi.org/10.1037/0096-1523.1.2.95

Legge, E. L. G., Madan, C. R., Ng, E. T., & Caplan, J. B. (2012). Building a memory palace in minutes: Equivalent memory performance using virtual versus conventional environments with the Method of Loci. *Acta Psychologica*, *141*(3), 380–390. doi: 10.1016/j.actpsy.2012.09.002

Levy, B. R. (2003). Mind Matters : Cognitive and Physical Effects of Aging Self-Stereotypes. *The journal of gerontology*, *58*(4), 203–211.

Lewis, P. A., & Durrant, S. J. (2011). Overlapping memory replay during sleep builds cognitive schemata. *Trends in Cognitive Sciences*, *15*(8), 343–351. doi: 10.1016/j.tics.2011.06.004

Lieberman, D. A. (2011). *Human learning and memory.* Cambridge University Press.

Loftus, E. F., & Fathi, D. C. (1985). Retrieving Multiple Autobiographical Memories. *Social Cognition*, *3*(3), 280–295. doi: 10.1521/soco.1985.3.3.280

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, *1*(4), 1–15.

Long, D. L., & Prat, C. S. (2002). Memory for Star Trek: The role of prior knowledge in recognition revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(6), 1073–1082. doi: 10.1037//0278-7393.28.6.1073

MacLeod, C. M., & Bodner, G. E. (2017). The Production Effect in Memory. *Current Directions in Psychological Science*, *26*(4), 390–395. doi: 10.1177/0963721417691356

Maguire, E. A., Valentine, E. R., Wilding, J. M., & Kapur, N. (2003). Routes to remembering: The brains behind superior memory. *Nature Neuroscience*, *6*(1), 90–95. doi: 10.1038/nn988

Mallow, J., Bernarding, J., Luchtmann, M., Bethmann, A., & Brechmann, A. (2015). Superior memorizers employ different neural networks for encoding and recall. *Frontiers in Systems Neuroscience*, *9*, 128. doi: 10.3389/fnsys.2015.00128

Mankin, E. A., Sparks, F. T., Slayyeh, B., Sutherland, R. J., Leutgeb, S., & Leutgeb, J. K. (2012). Neuronal code for extended time in the hippocampus. *Proceedings of the National Academy of Sciences*, *109*(47), 19462–19467. doi: 10.1073/pnas.1214107109

Marks, D. F. (1972a). The Function and Nature of Imagery. In S. P.W. (Ed.), (chap. Individual). New York and London: Academic Press.

Marks, D. F. (1972b). Individual differences in the vividness of visual imagery and their effect on function. In P. W. Sheehan (Ed.), *The function and nature of imagery* (pp. 83–108). New York: Academic Press.

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, *64*(1), 17–24.

Massen, C., & Vaterrodt-Plünnecke, B. (2006). The role of proactive interference in mnemonic techniques. *Memory*, *14*(2), 189–196. doi: 10.1080/09658210544000042

McAdams, D. P., & McLean, K. C. (2013). Narrative Identity. *Current Directions in Psychological Science*, *22*(3), 233–238. doi: 10.1177/0963721413475622

McCabe, J. A. (2015). Location, location, location! Demonstrating the mnemonic benefit of the method of loci. *Teaching of Psychology*, *42*(2), 169–173. doi: https://doi.org/10.1177/0098628315573143

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419–457. doi: 10.1037/0033-295X.102.3.419

McKelvey, S. (1995). The VVIQ as a psychometric test of individual differences in visual imagery vividness: a critical quantitative review and plea for direction. *Journal of Mental Imagery*, *19*(3), 1–106.

McKenzie, S., & Eichenbaum, H. (2011). Consolidation and Reconsolidation: Two Lives of Memories? *Neuron*, *71*(2), 224–233. doi: 10.1016/j.neuron.2011.06.037

Meeter, M., & Murre, J. M. (2004). Consolidation of long-term memory: Evidence and alternatives. *Psychological Bulletin*, *130*(6), 843–857. doi: 10.1037/0033-2909.130.6.843

Miller, J. F., Neufang, M., Solway, A., Brandt, A., Trippel, M., Mader, I., ... Schulze-Bonhage, A. (2013). Neural activity in human hippocampal formation reveals the spatial context of retrieved memories. *Science*, *342*(6162), 1111–1114. doi: 10.1126/science.1244056

Monaco, J. D., Rao, G., Roth, E. D., & Knierim, J. J. (2014). Attentive scanning behavior drives one-trial potentiation of hippocampal place fields. *Nature Neuroscience*, *17*(5), 725–731. doi: 10.1038/nn.3687

Moser, M.-B., Rowland, D. C., & Moser, E. I. (2015). Place cells, grid cells, and memory. *Cold Spring Harbor perspectives in biology*, *7*(2), a021808. doi: 10.1101/cshperspect.a021808

Müller, N. C. J., Konrad, B. N., Kohn, N., Muñoz-López, M., Czisch, M., Fernández, G., & Dresler, M. (2018). Hippocampal–caudate nucleus interactions support exceptional memory performance. *Brain Structure and Function*, *223*(3), 1379–1389. doi: 10.1007/s00429-017-1556-2

Murdock, B. B. (1974). *Human memory: theory and data.* New York: John Wiley & Sons.

Nairne, J. S. (2002). The myth of the encoding-retrieval match. *Memory*, *10*(5-6), 389–395. doi: 10.1080/09658210244000216

Niedenthal, P. M. (2007). Embodying emotion. *Science*, *316*(5827), 1002–1005. doi: 10.1126/science.1136930

Nyberg, L., Sandblom, J., Jones, S., Neely, A. S., Petersson, K. M., Ingvar, M., & Bäckman, L. (2003, 11). Neural correlates of training-related memory improvement in adulthood and aging. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(23), 13728–33. doi: 10.1073/pnas.1735487100

Oladimeji, D. M., Kluger, F. E., & Caplan, J. B. (0). *Wayfinding Ability and the Effectiveness of the Method of Loci in Older Adults.*

O'Reilly, R. C., Bhattacharyya, R., Howard, M. D., & Ketz, N. (2014). Complementary learning systems. *Cognitive Science*, *38*(6), 1229–1248. doi: 10.1111/j.1551-6709.2011.01214.x

Paivio, A. (1971). Imagery and deep structure in the recall of English nominalizations. *Journal of Verbal Learning and Verbal Behavior*, *10*, 1–12.

Peirce, J., Gray, J. R., Simpson, S., Macaskill, M., Höchenberger, R., Sogo, H., . . . Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*, 195–203. doi: 10.3758/s13428-018-01193-y

Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Plancher, G., Barra, J., Orriols, E., & Piolino, P. (2013). The influence of action on episodic memory: A virtual reality study. *Quarterly Journal of Experimental Psychology*, *66*(5), 895–909. doi: 10.1080/17470218.2012.722657

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, *6*(7), 576–582. doi: 10.1038/nrn1706

Raaijmakers, J. (1993). The story ofthe two-store model: Past criticisms, current status, and future directions. In D. Meyer & S. Kornblum (Eds.), *Attention and performance xiv: synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (p. 467–488). Cambridge, MA: The MIT Press.

Raftery, A. E., & Kass, R. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, *90*(430), 773–795.

Raz, A., Packard, M. G., Alexander, G. M., Buhle, J. T., Zhu, H., Yu, S., & Peterson, B. S. (2009). A slice of $\pi$ : An exploratory neuroimaging study of digit encoding and retrieval in a superior memorist. *Neurocase*, *15*(5), 361–372. doi: 10.1080/13554790902776896

Richardson, D. C., Spivey, M. J., Cheung, J., & Cheung, J. (2001). Motor Representations in Memory and Mental Models: Embodiement in Cognition. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *23*(23).

Richardson, J. (1995). The efficacy of imagery mnemonics in memory remediation. *Neuropsychologia*, *33*(11), 1345–1357. doi: 10.1016/0028-3932(95)00068-E

Robin, J. (2018). Spatial scaffold effects in event memory and imagination. *Wiley Interdisciplinary Reviews: Cognitive Science*, *9*(4), e1462. doi: 10.1002/wcs.1462

Robin, J., Buchsbaum, B. R., & Moscovitch, M. (2018). The primacy of spatial context in the neural representation of events. *Journal of Neuroscience*, *38*(11), 2755–2765. doi: 10.1523/JNEUROSCI.1638-17.2018

Robin, J., Wynn, J., & Moscovitch, M. (2016). The spatial scaffold: The effects of spatial context on memory for events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(2), 308–315. doi: 10.1037/xlm0000167

Roediger, H. L. (1980). The effectiveness of four mnemonics in ordering recall. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(5), 558–567.

Roediger, H. L., & Crowder, R. G. (1976). *A serial position effect in recall of United States presidents* (Vol. 8) (No. 4). Springer-Verlag. doi: 10.3758/BF03335138

Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology*, *35*(9), 677–688. doi: 10.1037/0022-3514.35.9.677

Rolls, E. T. (2017). A scientific theory of Ars Memoriae : Spatial view cells in a continuous attractor network with linked items. *Hippocampus*, *27*(5), 570–579. doi: 10.1002/hipo.22713

Ross, J., & Lawrence, K. A. (1968). Some observations on memory artifice. *Psychonomic Science*, *13*(2), 107–108. doi: https://doi.org/10.3758/BF03342433

Rubin, D. C. (2020). The ability to recall scenes is a stable individual difference: Evidence from autobiographical remembering. *Cognition*, *197*(988), 104164. doi: 10.1016/j.cognition.2019.104164

Rubin, D. C. (2021). Properties of autobiographical memories are reliable and stable individual differences. *Cognition*, *210*(6), 104583. doi: 10.1016/J.COGNITION.2021.104583

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In (Eds.),. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence, and education* (p. pp. 33–58). Hillsdale, NJ: Erlbaum.

Sahadevan, S. S., Chen, Y. Y., & Caplan, J. B. (2021). Imagery-based strategies for memory for associations. *Memory*, *20*(10), 1275–1295.

Sakreida, K., Scorolli, C., Menz, M. M., Heim, S., Borghi, A. M., & Binkofski, F. (2013). Are abstract action words embodied? An fMRI investigation at the interface between language and motor cognition. *Frontiers in Human Neuroscience*, *7*, 125. doi: 10.3389/fnhum.2013.00125

Sanchez, C. A. (2019). The utility of Visuospatial Mnemonics is Dependent on Visuospatial Aptitudes. *Applied Cognitive Psychology*, *33*(4), 702–708.

Sauzéon, H., Pala, P., Florian, L., Wallet, G., Déjos, M., Zheng, X., . . . Bernard, N. (2011). The Use of Virtual Reality for Episodic Memory Assessment. *Experimental psychology*, *59*(2), 99–108. doi: 10.1027/1618-3169/a000131

Schank, R. C., & Abelson, R. P. (1977). *Scripts Plans Goals and Understanding An Inquiry into Human Knowledge Structures.* Oxford, England: Lawrence Erlbaum. doi: 10.2307/412850

Sharon, T., Moscovitch, M., & Gilboa, A. (2011). Rapid neocortical acquisition of long-term arbitrary associations independent of the hippocampus. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(3), 1146–1151. doi: 10.1073/pnas.1005238108

Skotko, B. G., Einstein, G., Rubin, D. C., Kensinger, E. A., Locascio, J. J., Tupler, L. A., ... Corkin, S. (2004). Puzzling thoughts for H.M.: Can new semantic information be anchored to old semantic memories? *Neuropsychology*, *18*(4), 756–769. doi: 10.1037/0894-4105.18.4.756

Smith, C. N., Urgolites, Z. J., Hopkins, R. O., & Squire, L. R. (2014). Comparison of explicit and incidental learning strategies in memory-impaired patients. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(1), 475–479. doi: 10.1073/pnas.1322263111

Sommer, T. (2017). The Emergence of Knowledge and How it Supports the Memory for Novel Related Information. *Cerebral cortex*, *27*(3), 1906–1921. doi: 10.1093/cercor/bhw031

Spence, J. D. (1984). *The memory palace of Matteo Ricci.* New York, NY: Viking Penguin.

Staszewski, J. J. (1990). Exceptional Memory: The Influence of Practice and Knowledge on the Development of Elaborative Encoding Strategies. In W. Schneider & F. Weinert (Eds.), *Interactions among aptitudes, strategies, and knowledge in cognitive performance* (Vol. 27). New York, NY: Springer. doi: 10.5860/choice.27-6027

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin*, *121*(3), 371–394. doi: 10.1037/0033-2909.121.3.371

Takahashi, M., Shimizu, H., Saito, S., & Tomoyori, H. (2006). One percent ability and ninety-nine percent perspiration: A study of a japanese memorist. *Journal of Experimental Psychology: Learning Memory and Cognition*, *32*(5), 1195–1200. doi: 10.1037/0278-7393.32.5.1195

Thompson, C. P., Cowan, T., Frieman, J., Mahadevan, R. S., Vogl, R. J., & Frieman, J. (1991). Rajan: A study of a memorist. *Journal of Memory and Language*, *30*(6), 702–724. doi: 10.1016/0749-596X(91)90033-G

Toni, I., de Lange, F. P., Noordzij, M. L., & Hagoort, P. (2008). Language beyond action. *Journal of Physiology Paris*, *102*(1-3), 71–79. doi: 10.1016/j.jphysparis.2008.03.005

Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., ... Morris, R. G. (2007). Schemas and memory consolidation. *Science*, *316*(5821), 76–82. doi: 10.1126/science.1135935

Tse, D., Takeuchi, T., Kakeyama, M., Kajii, Y., Okuno, H., Tohyama, C., ... Morris, R. G. (2011). Schema-dependent gene activation and memory encoding in neocortex. *Science*, *333*(6044), 891–895. doi: 10.1126/science.1205274

Tuena, C., Serino, S., Gaston-Bellegarde, A., Orriols, E., Makowski, D., Riva, G., & Piolino, P. (2017). How virtual embodiment affects episodic memory functioning: a proof-of-concept study. In B. K. Wiederhold, G. Riva, C. Fullwood, A. Attrill-Smith, & G. Kirwan (Eds.), *Annual review of cybertherapy and telemedicine 2017 : A healthy mind in a healthy virtual body: The future of virtual reality in health care* (pp. 98–104).

Tulving, E. (2002). Episodic Memory: From Mind to Brain. *Annual Review of Psychology*, *53*, 1–25.

Twomey, C., & Kroneisen, M. (2021). The effectiveness of the loci method as a mnemonic device: Meta-analysis. *Quarterly Journal of Experimental Psychology*, *74*(8), 1–10.

doi: 10.1177/1747021821993457
Van Kesteren, M. T., Ruiter, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences*, *35*(4), 211–219. doi: 10.1016/j.tins.2012.02.001
Varela, F., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.
Voyer, D., Saint Aubin, J., Altman, K., & Gallant, G. (2021). Sex differences in verbal working memory: A systematic review and meta-analysis. *Psychological Bulletin*, *147*(4), 352–398. Retrieved from `/record/2021-59065-002` doi: 10.1037/BUL0000320
Wang, A. Y., & Thomas, M. H. (2000). Looking for long-term mnemonic effects on serial recall: the legacy of Simonides. *The American journal of psychology*, *113*(3), 331–40.
Weisberg, S. M., Schinazi, V. R., Newcombe, N. S., Shipley, T. F., & Epstein, R. A. (2014). Variations in cognitive maps: Understanding individual differences in navigation. *Journal of Experimental Psychology: Learning Memory and Cognition*, *40*(3), 669–682. doi: 10.1037/a0035261
Wen, W., Ishikawa, T., & Sato, T. (2011). Working memory in spatial knowledge acquisition: Differences in encoding processes and sense of direction. *Applied Cognitive Psychology*, *25*(4), 654–662. doi: 10.1002/acp.1737
Wenger, M. J., & Payne, D. G. (1995). On the Acquisition of Mnemonic Skill: Application of Skilled Memory Theory. *Journal of Experimental Psychology: Applied*, *1*(3), 194–215. doi: 10.1037/1076-898X.1.3.194
Wilding, J., & Valentine, E. (2006). Exceptional memory. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (p. 539–552). Cambridge, UK: Cambridge University Press.
Winocur, G., & Moscovitch, M. (2011). Memory transformation and systems consolidation. *Journal of the International Neuropsychological Society*, *17*(5), 766–780. doi: 10.1017/S1355617711000683
Worthen, J. B., & Hunt, R. (2011). Mnemonology: Mnemonics for the 21st century. *Mnemonology: Mnemonics for the 21st Century*, 1–161. doi: 10.4324/9780203834107
Worthen, J. B., & Hunt, R. R. (2008). Mnemonics: Underlying processes and practical applications. In J. H. Byrne (Ed.), *Learning and memory: A comprehensive reference* (p. 145–153).
Yates, F. (1966). *The art of memory.* Lancet Publishing Group. doi: 10.1016/S1474-4422(18)30430-7
Yesavage, Lapp, & Sheikh. (1989). Mnemonics as modified for use by the elderly. In L. W. Poon, D. C. Rubin, & Barbara A. Wilson (Eds.), *Everyday cognition in adulthood and late life* (pp. 598–611). New York: Cambridge University Press.
Zimmer, H. D., & Cohen, L. R. (2001). *Memory for action: A distinct form of episodic memory?* Oxford, England: Oxford University Press.
Zwaan, R. A. (2004). The immersed experiencer: toward an embodied theory of language comprehension. In B. Ross (Ed.), *The psychology of learning and motivation* (Academic ed., p. 35–62). New York, NY: The psychology of learning and motivation: Advances in research and theory.