Mnemonic scaffolds vary in effectiveness for serial recall

Felicitas E. Kluger, Debby M. Oladimeji, Yuwei Tan, Norman R. Brown and Jeremy B. Caplan

Department of Psychology and Neuroscience and Mental Health Institute, University of Alberta, Edmonton, AB, Canada.

Author Note

Corresponding Author: Felicitas Kluger (kluger@ualberta.ca). Supported in part by the Natural Sciences and Engineering Research Council of Canada. The dataset generated and analyzed during the current study is available from: https://osf.io/ktnpj/?view_only=d81700ffe06c4c15a2c5b2d767002ff0. The authors report no conflict of interest. We thank James Paterson and Boris Nikolai Konrad for helpful discussions that influenced the experimental design and research questions.

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.

Keywords: mnemonic techniques, mnemonic scaffolds, serial recall, Method of Loci, visual imagery, embodiment

Mnemonic scaffolds vary in effectiveness for serial recall

Introduction

Arguably the most effective mnemonic techniques are those that leverage previously learned material (Hu et al., 2009; 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). This does not require superior cognitive aptitudes or extraordinary brain anatomy (Chase & Ericsson, 1981; Ericsson & Staszewski, 1989; Ericsson & Kintsch, 1995; Maguire et al., 2003; 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.

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 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

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

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.

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, 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 dismissed. 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 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

[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).

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.

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, 1972a). 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.

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.

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.

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. 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.

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 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 1. The first phase measured serial recall ability prior to instruction and was thus

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

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 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.

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 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. As shown in previous studies, proactive interference is not to be expected (Massen &

Vaterrodt-Plünnecke, 2006; Bass & Oswald, 2014). 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 a special issue on the production effect (see the editorial 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 pre-instruction 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; $BF_{10} = p(H1)/p(H0)$. 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. 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 it 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.

Results and Discussion

First, we verified the absence of a subject sampling bias across groups (Table 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 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 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 5 and Figure 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 3a–c, respectively), but not the Activity Group, with the Bayes factor providing evidence for a null effect (Figure 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 0.1). 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) and lenient- (Figures 0.4) scored post-instruction recall accuracy (Table 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 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 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, 1972a); 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 classical 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 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 classical linear and Bayesian linear regressions to predict post-instruction recall accuracy (strict scoring) based on PFT scores in the Loci Group (Table 2 and Figure 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 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 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. Classical 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. Classical linear and Bayesian linear regressions and pairwise Bayesian correlations revealed that this was indeed the case for all Groups (Table 2). 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 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 actually be a measure of motivation and engagement of the participants throughout the whole experiment.

BRQ. Following our planned comparison, we conducted classical linear 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.

Interim summary

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.

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.

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.

Materials. We used the same materials as for Experiment 1.

Procedures. The basic paradigm was as in Experiment 1 (Figure 1). There were four groups: a Control Group, identical to the Control Group in Experiment 2, 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).

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, and Figure 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).

Comparison of the effectiveness of the Body Scaffold variants between groups. While reporting post-instruction serial recall data (strict scoring: Figure 8; lenient scoring: Figure 0.8), 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, Figure 8). 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). 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). 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). 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, Figure 0.8). 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 linear 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 linear 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 PFT scores for a follow-up, exploratory analysis.

BRQ. Following our planned comparison, we conducted classical linear 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 linear 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 0.10. 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 adds to the evidence that our hypothesis that high-imageryparticipants 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 linear 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 0.11. 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.

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 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 5e). Finally, individual-differences analyses showed no reliable relationship between visual imagery skill and body responsiveness and the usefulness of our four mnemonic scaffolds.

The prior knowledge effect does not apply to all mnemonic scaffolds

The failure of the 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 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 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.

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.

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).

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.

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.

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).

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 in 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, 1972a, 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).

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).

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.

References

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

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

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. (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

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

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.

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

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*. 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

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*(1), 61–64.

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.

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.

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

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.

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

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 (New York, N.Y.)*, *208*(4448), 1181–2. doi: 10.1126/SCIENCE.7375930

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.

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), 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*. 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

Hirst, W., & Echterhoff, G. (2011). Remembering in Conversations: The Social Sharing and Reshaping of Memories. *Annual Review of Psychology*, *63*, 55–79. doi:

10.1146/ANNUREV-PSYCH-120710-100340

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.

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*(DEC). 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

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

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.

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*, 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.

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

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

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*, 1–106.

Meeter, M., & Murre, J. M. (2004). *Consolidation of long-term memory: Evidence and alternatives* (Vol. 130) (No. 6). Psycholgical Bulletin. doi: 10.1037/0033-2909.130.6.843

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.

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

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* (Vol. 6) (No. 7). Nature Publishing Group. 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.

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).

Roediger, H. L. (1980). The Effectiveness of Four Mnemonics in Ordering Recall. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(5), 558–567.

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*, 104164. doi:

10.1016/j.cognition.2019.104164

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

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*(MAR), 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*, 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

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).

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.

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* (Vol. 17) (No. 5). Cambridge University Press. 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.

Group	R^2	β	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 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.*

Group	R^2	β	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 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.*

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

Table 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

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

Table 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

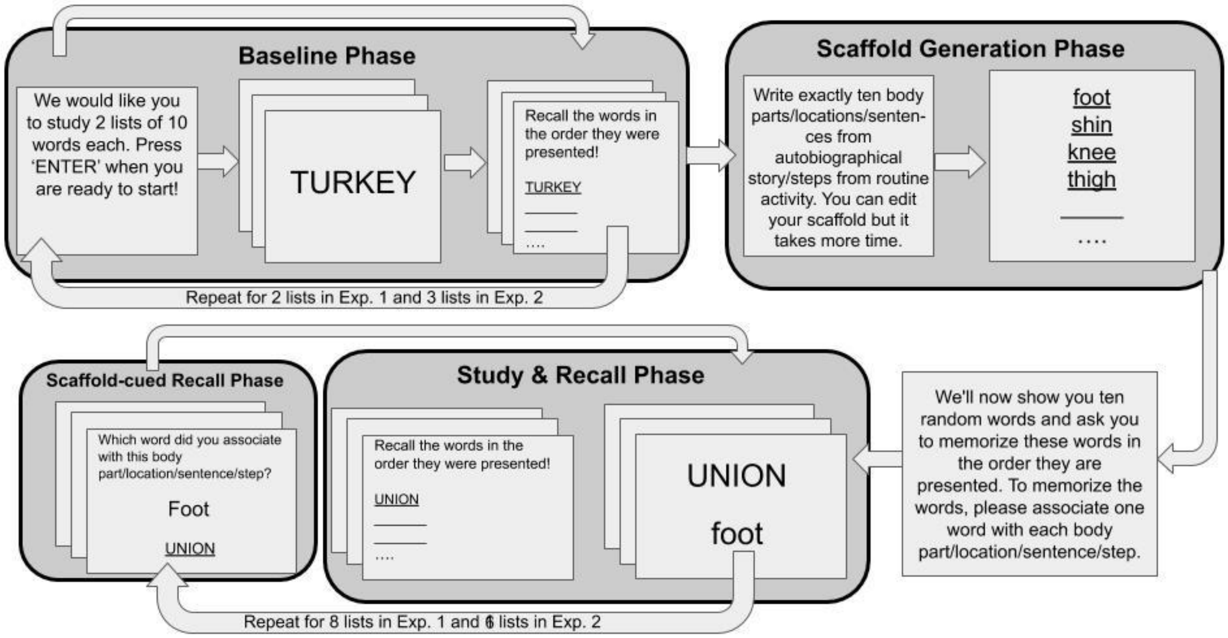


Figure 1. Illustration of the experimental design.

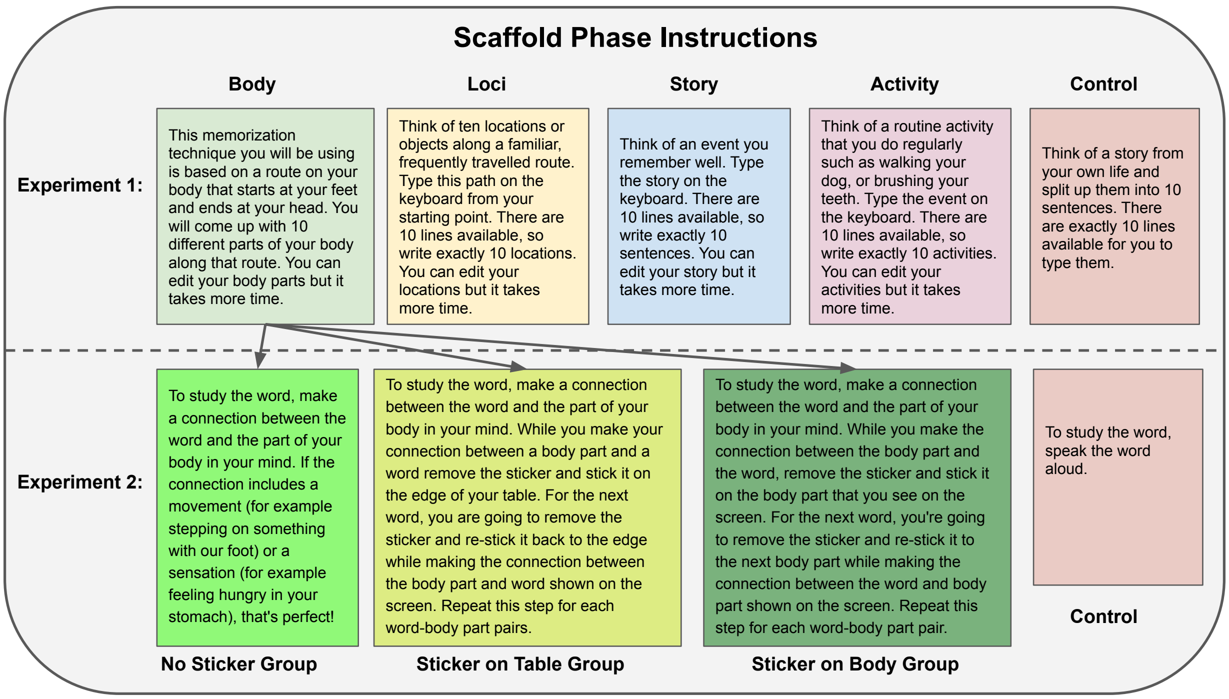


Figure 2. Illustration of the experimental groups in both experiments.

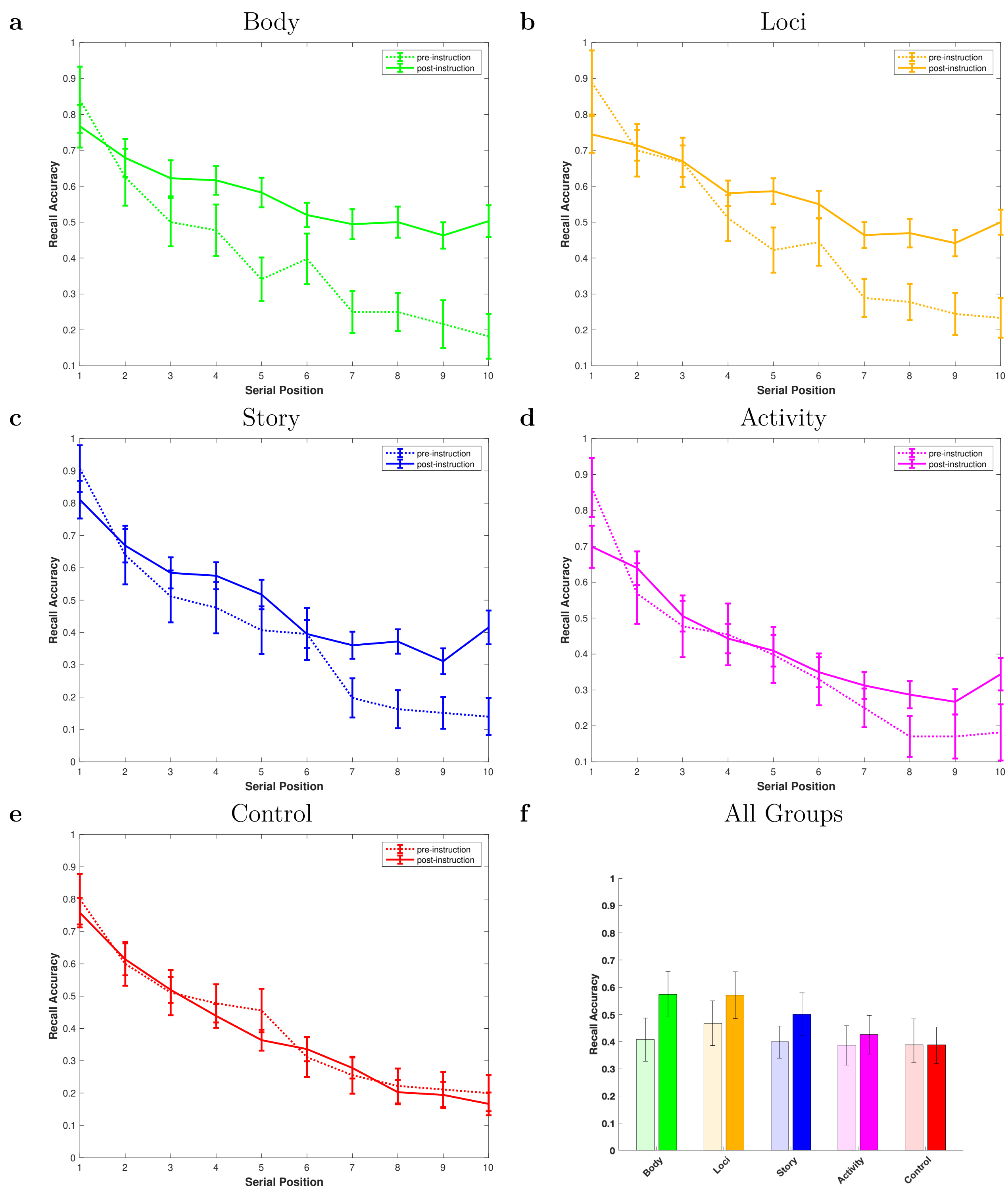


Figure 3. Experiment 1: Serial Position Curves for individual groups and bar graph for all groups comparing pre- and post-instruction serial recall accuracy (strict scoring). Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994).

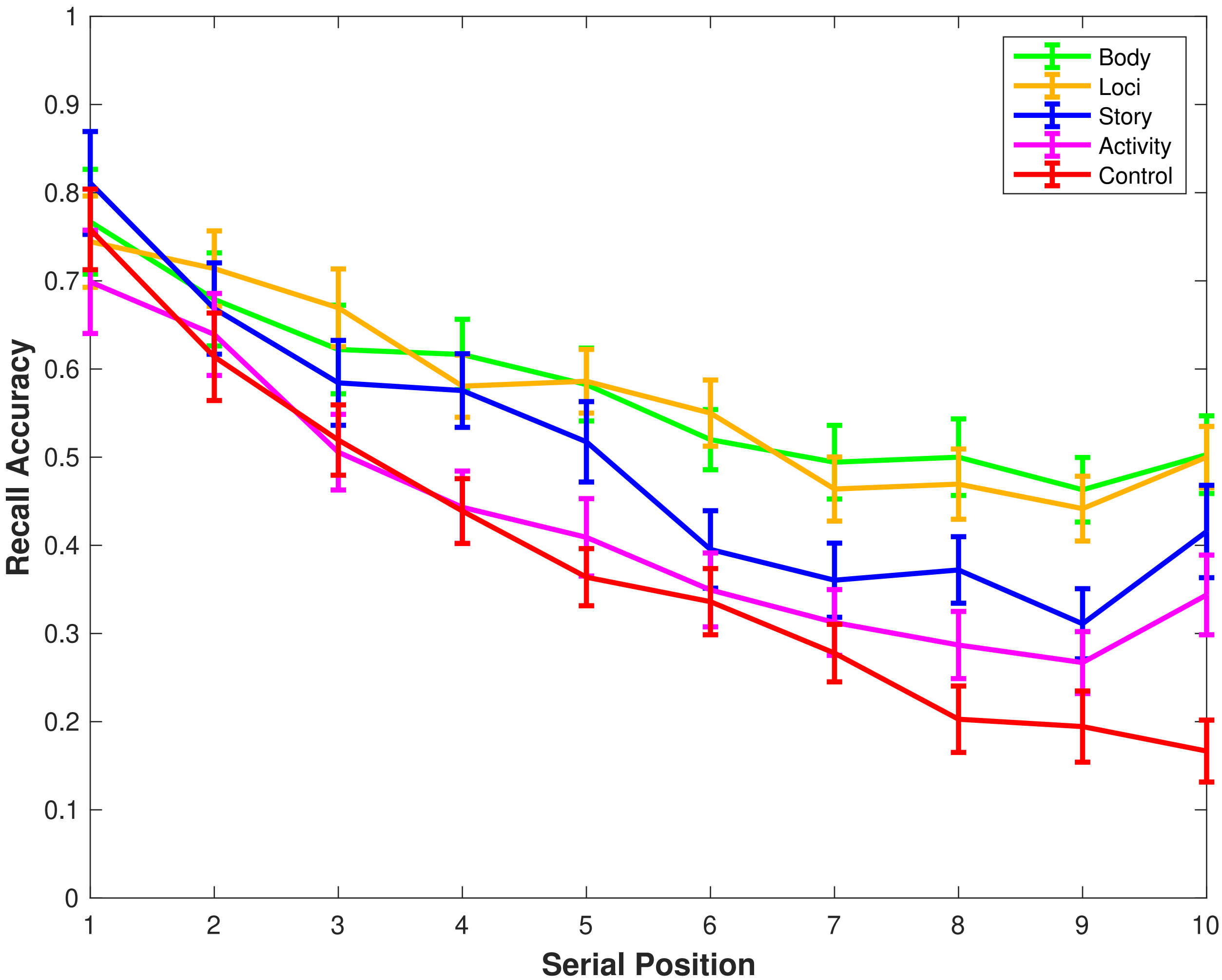


Figure 4. Experiment 1: Post-instruction accuracy as a function of serial position (strict scoring), as a function of group. Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994).

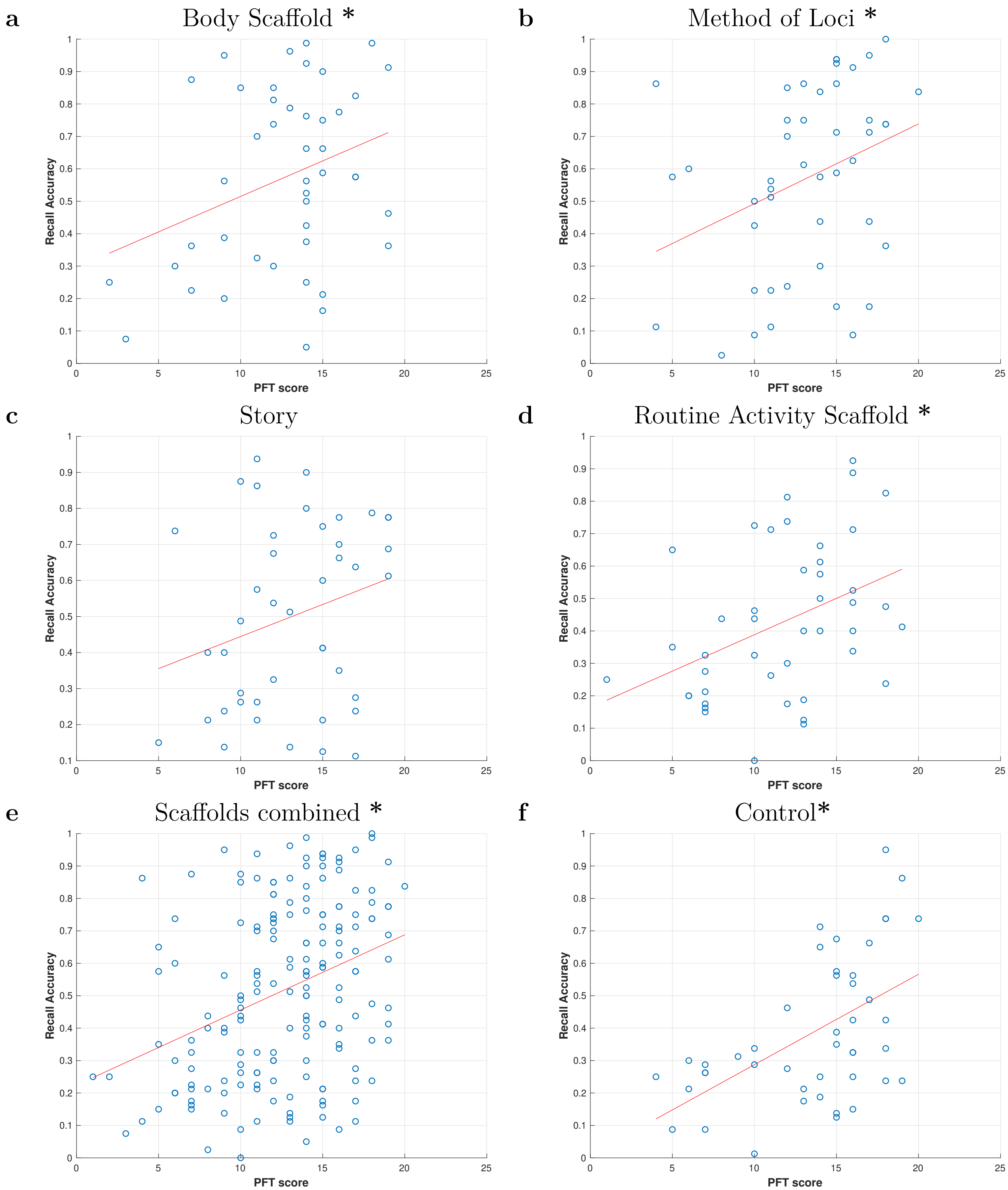


Figure 5. Experiment 1: Average serial recall by PFT scores. Regressions in groups marked with * are statistically significant ($p<0.05$).

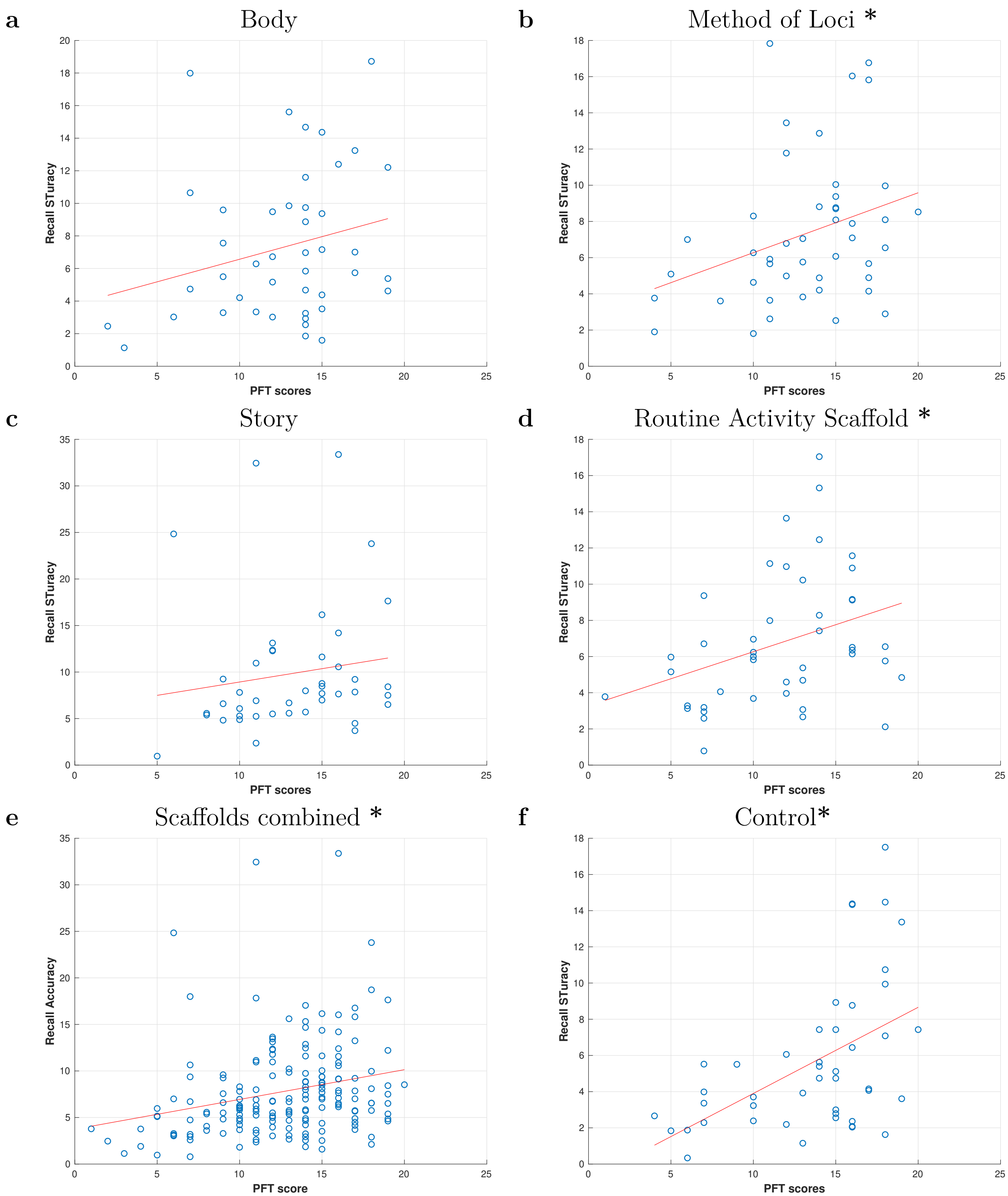


Figure 6. Experiment 1: Post-instruction study times by PFT scores. Regressions in groups marked with * are statistically significant.

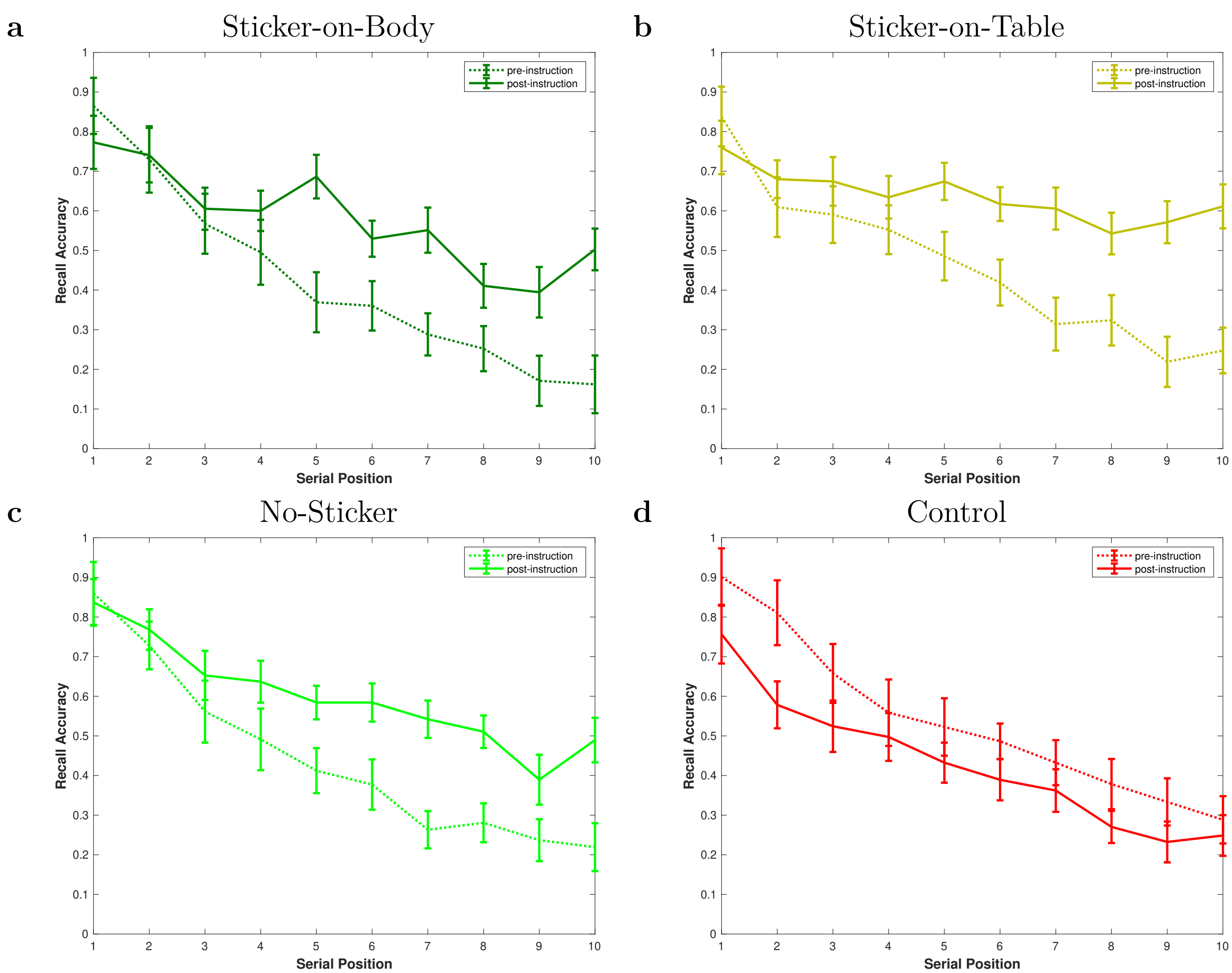


Figure 7. Experiment 2: Serial Position Curves comparing pre- and post-instruction recall accuracy (strict scoring). Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994)

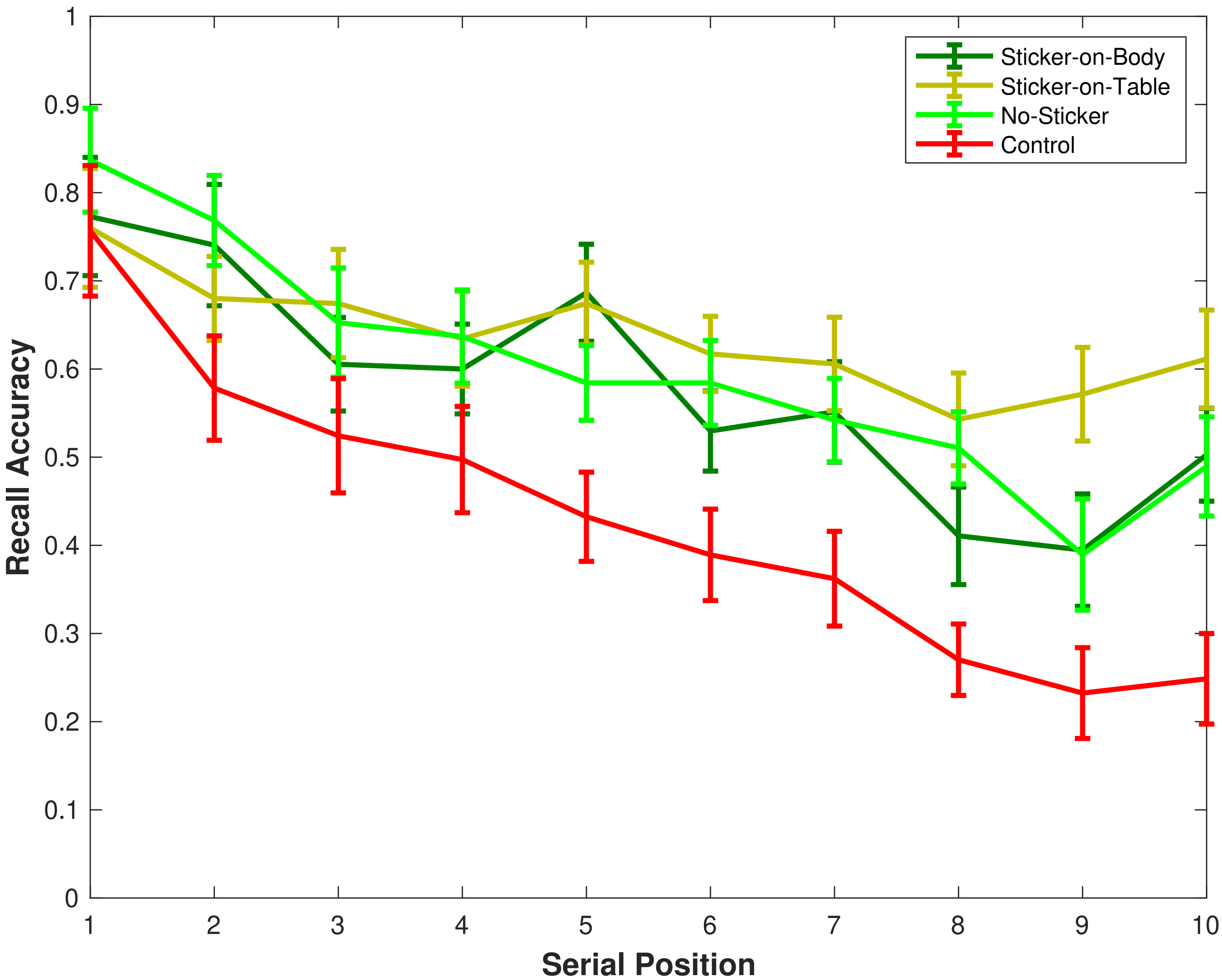


Figure 8. Experiment 2 - order memory (strict scoring): Post-instruction recall accuracy for each group. Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994).

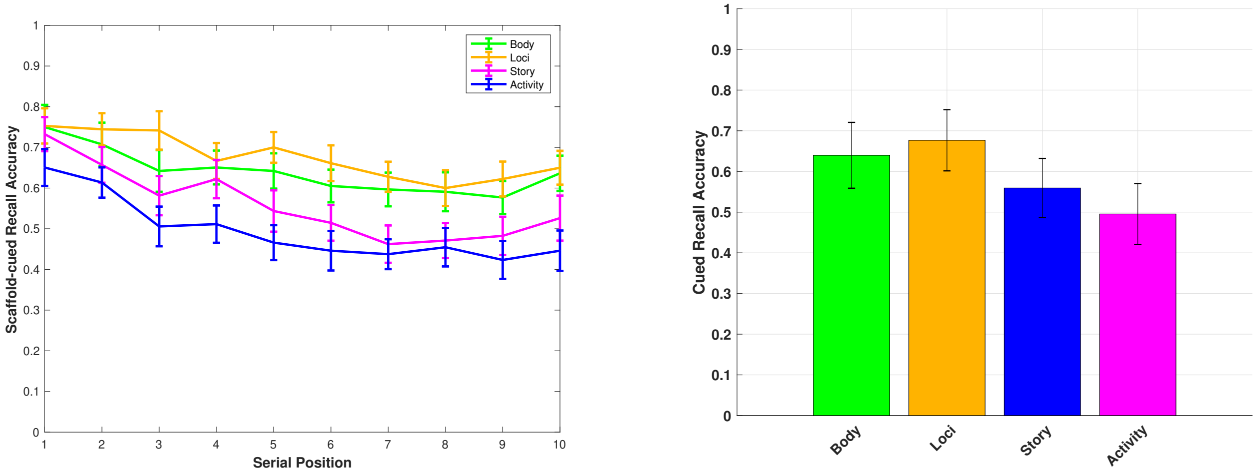
Supplementary Materials A

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) × 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 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 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.

Scaffold-cued recall. To evaluate whether participants were successful in forming word–scaffold associations (as instructed) we conducted a linear regression to predict post



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

(b) Scaffold-cued recall collapsed across serial position

Figure 0.1. 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).

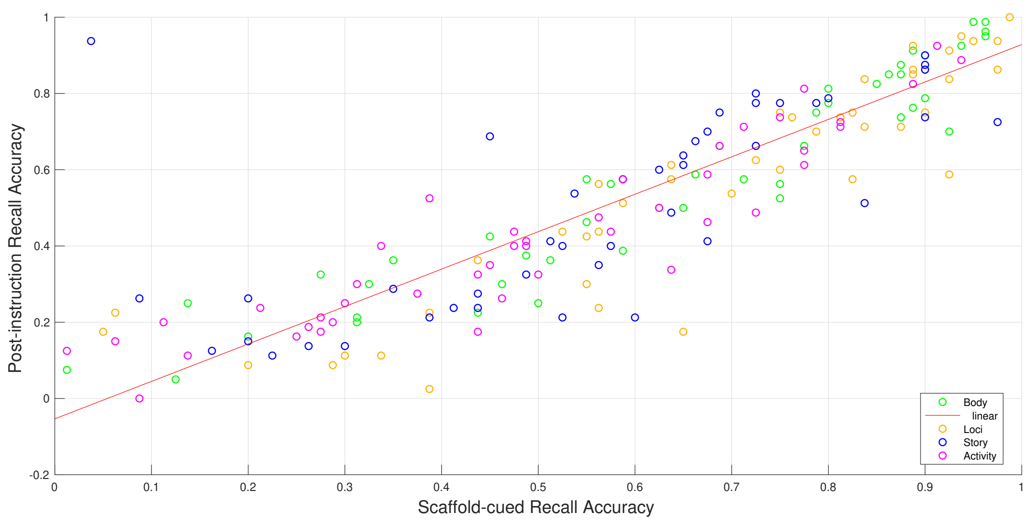


Figure 0.2. Experiment 1: Average post-instruction serial recall accuracy (strict scoring) by scaffold-cued recall accuracy

instruction serial recall accuracy (strict scoring) based on scaffold-cued recall accuracy (Figure 0.2. 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 0.1). 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.

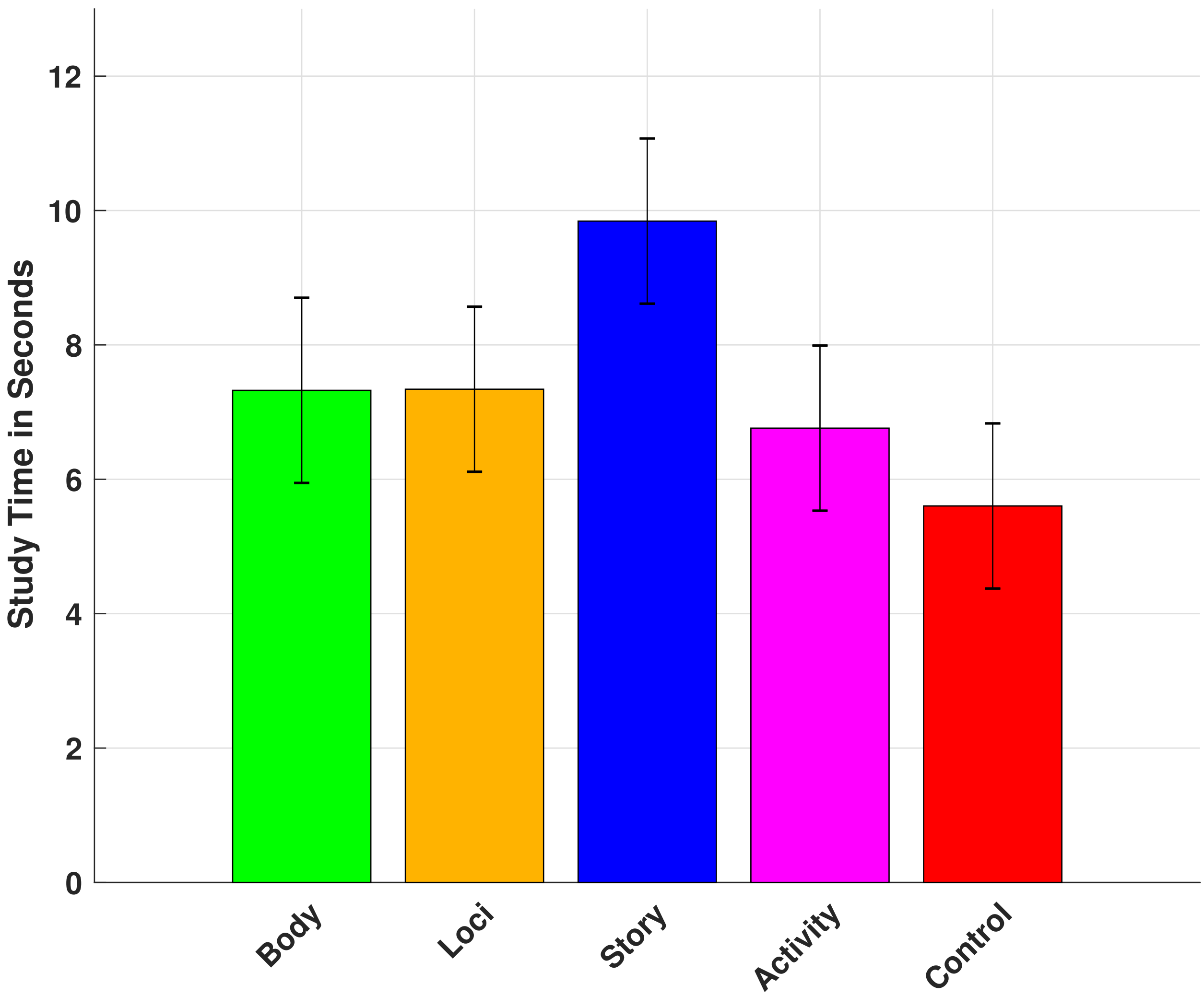


Figure 0.3. Experiment 1: Study times in seconds for each mnemonic scaffolds group. Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994).

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 0.3), 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 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 linear 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.

Serial position curve for lenient scoring. The serial position curve for lenient scoring of Experiment 1 is reported below in Figure 0.4

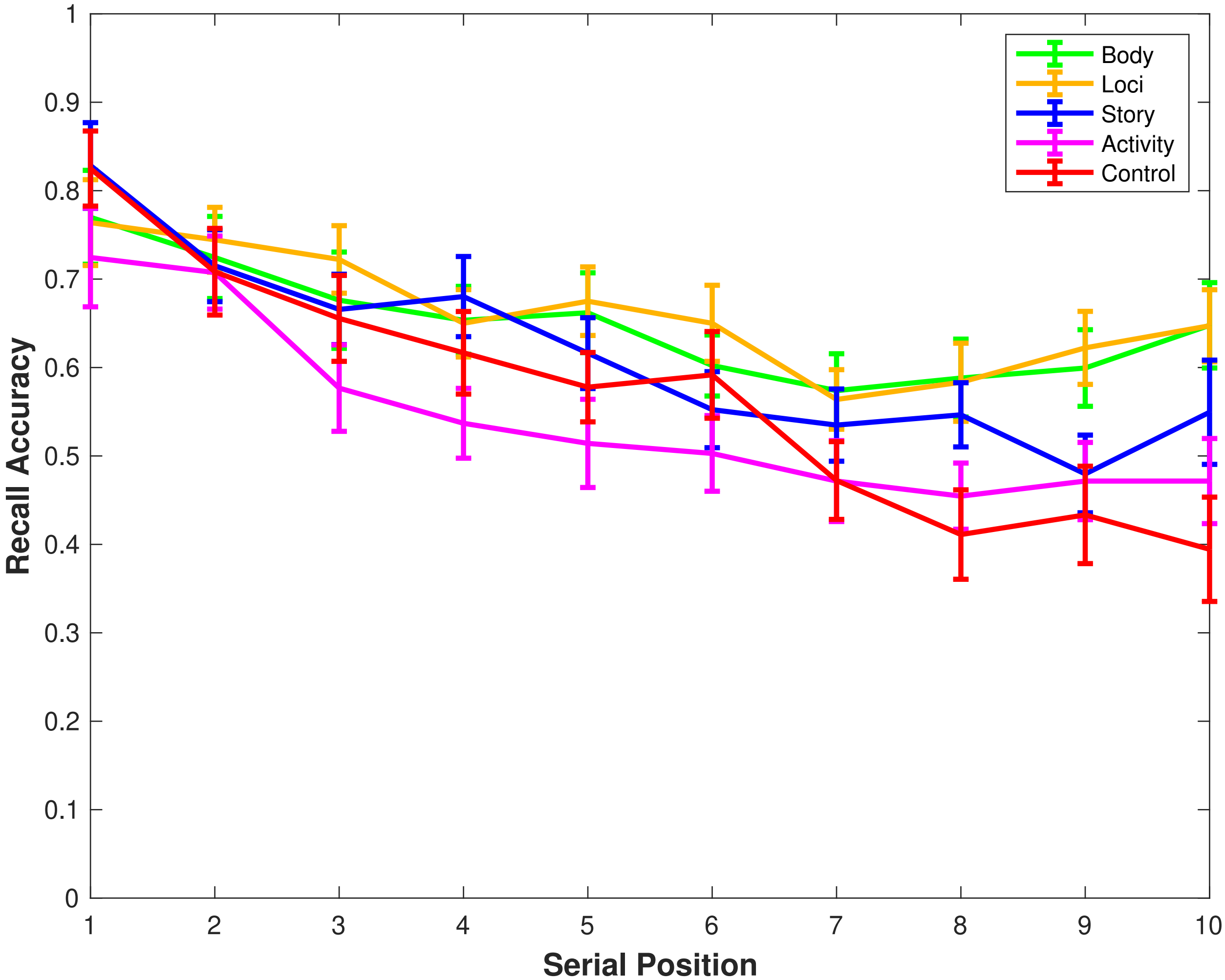


Figure 0.4. Experiment 1: Serial position curve for post-instruction recall accuracy (lenient scoring) as a function of group. Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994).

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 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 0.5 and 0.6. 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 success with which participants formed scaffold–word associations, we conducted a one-way ANOVA of Body Scaffold Variant on scaffold-cued recall accuracy (Figures 0.6 and 0.5). 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.

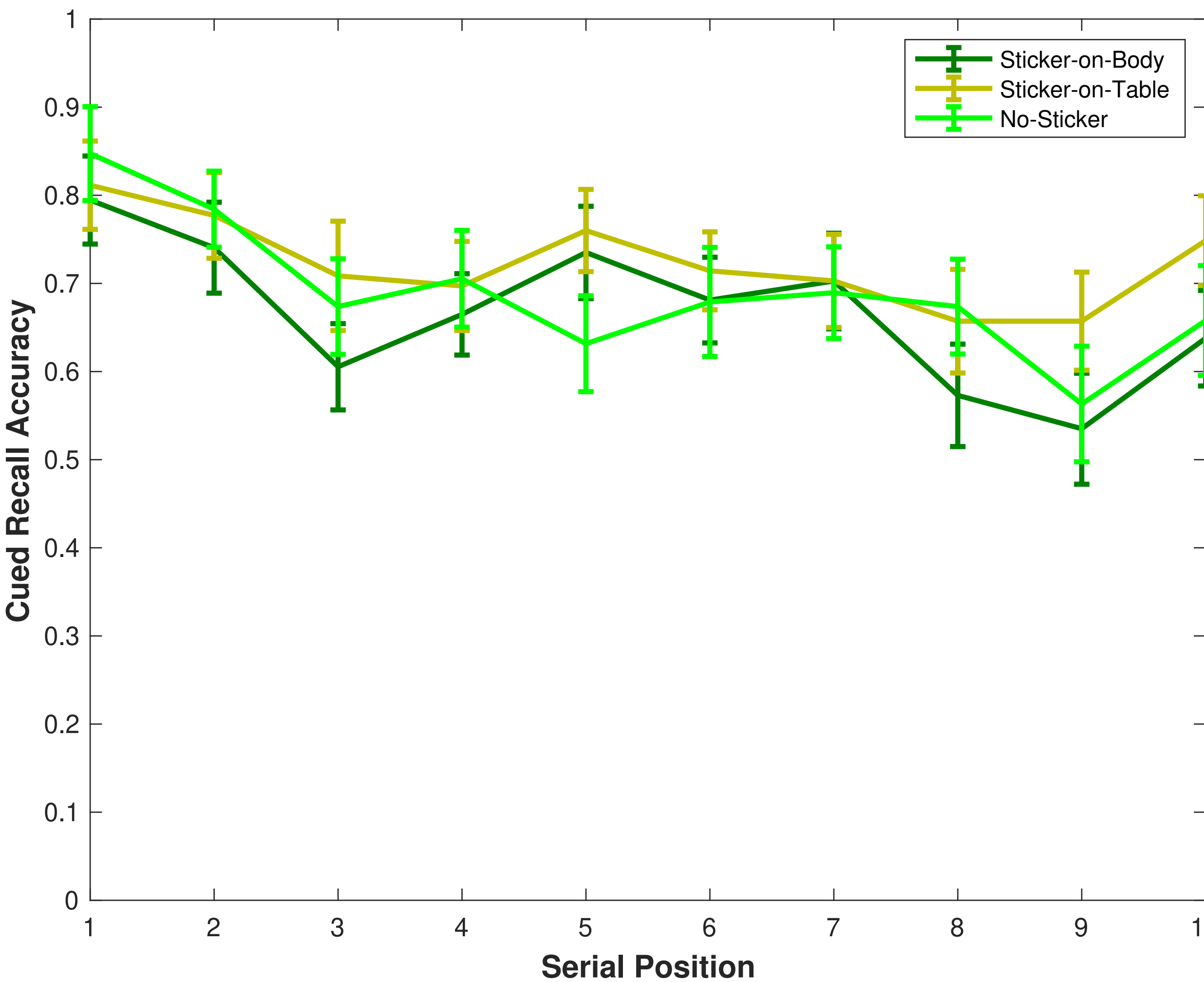


Figure 0.5. Experiment 2: Serial position curves of cued recall accuracy for each group. Error bars are standard errors of the mean corrected for subject variability (Loftus and Masson, 1994).

Effect of self-paced study time. We were interested in whether the different instructions with varying degrees of bodily engagement affected study time (Figure 0.7). 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

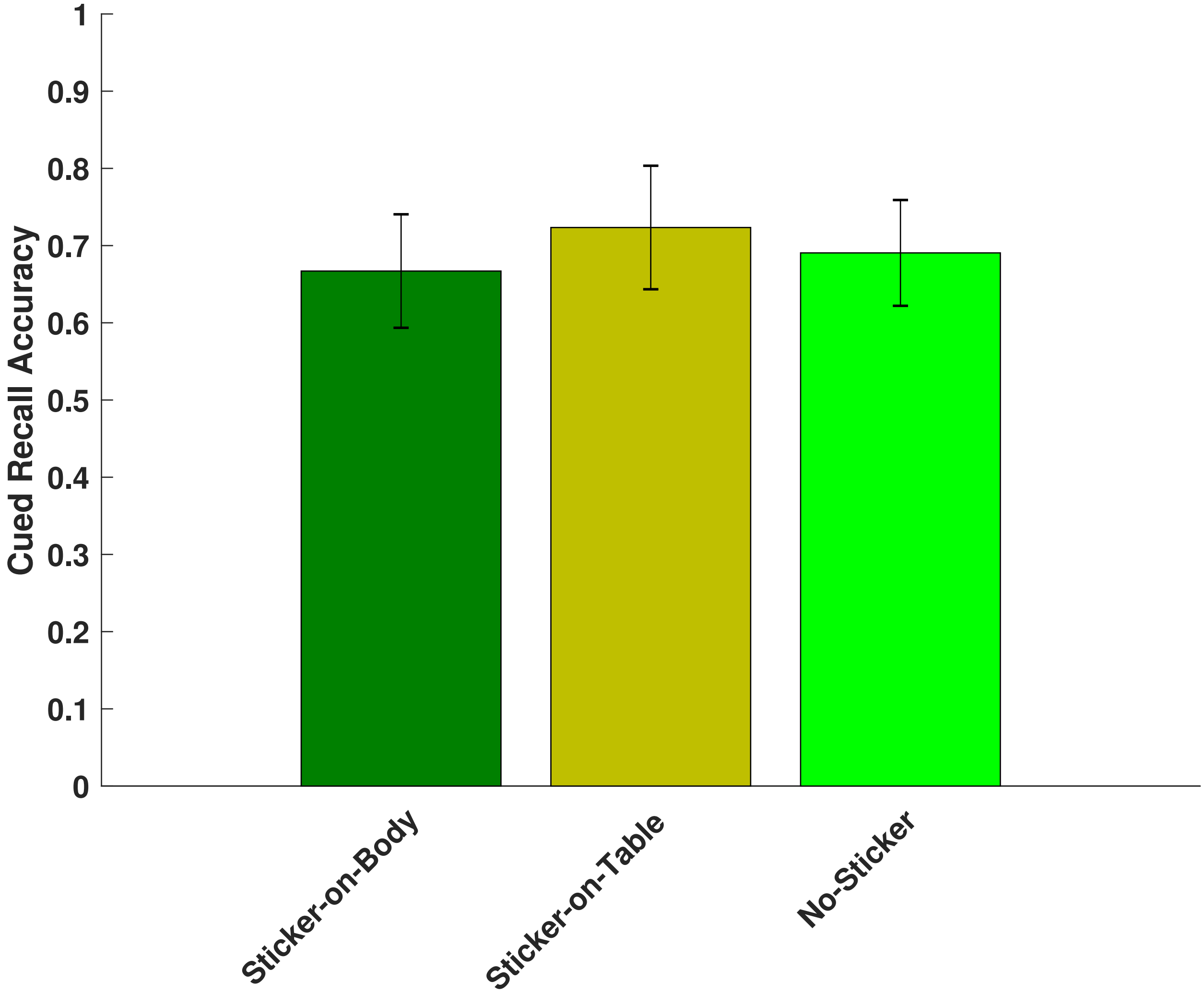


Figure 0.6. Experiment 2: Scaffold-cued recall accuracy for each group. Error bars are standard errors of the mean corrected for subject variability (Loftus and Masson, 1994).

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

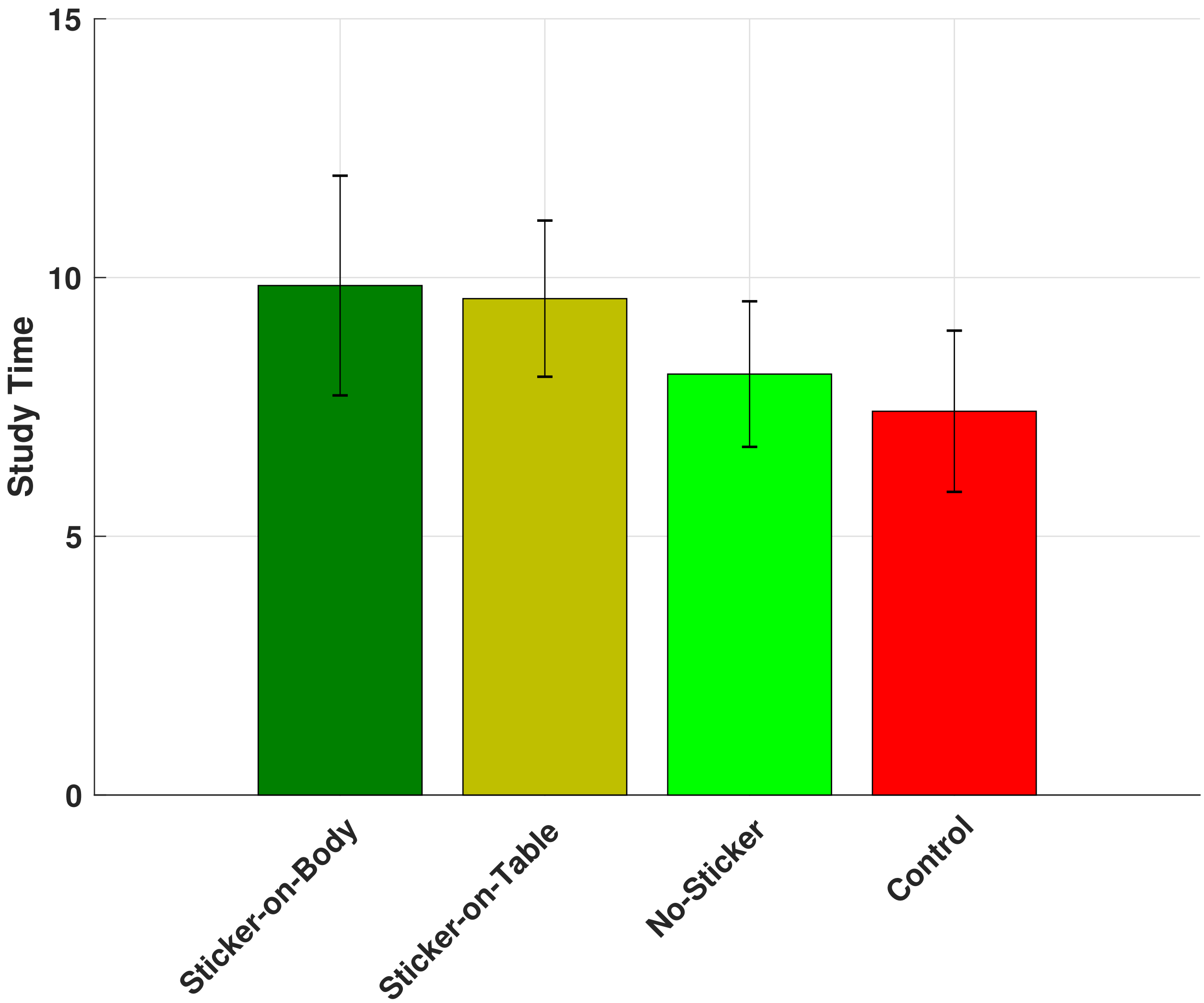


Figure 0.7. Experiment 2: Study time for each group. Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994).

advantage provided by the mnemonic scaffolds was not confounded by a pre-existing relationship in pre-instruction baseline memory, we conducted classical linear 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 A1
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 A2
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 A3
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 A1). b) connecting words with body parts made remembering the words easier, (Table A2), and whether c) they had used this technique before (Table A3). 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 0.8.

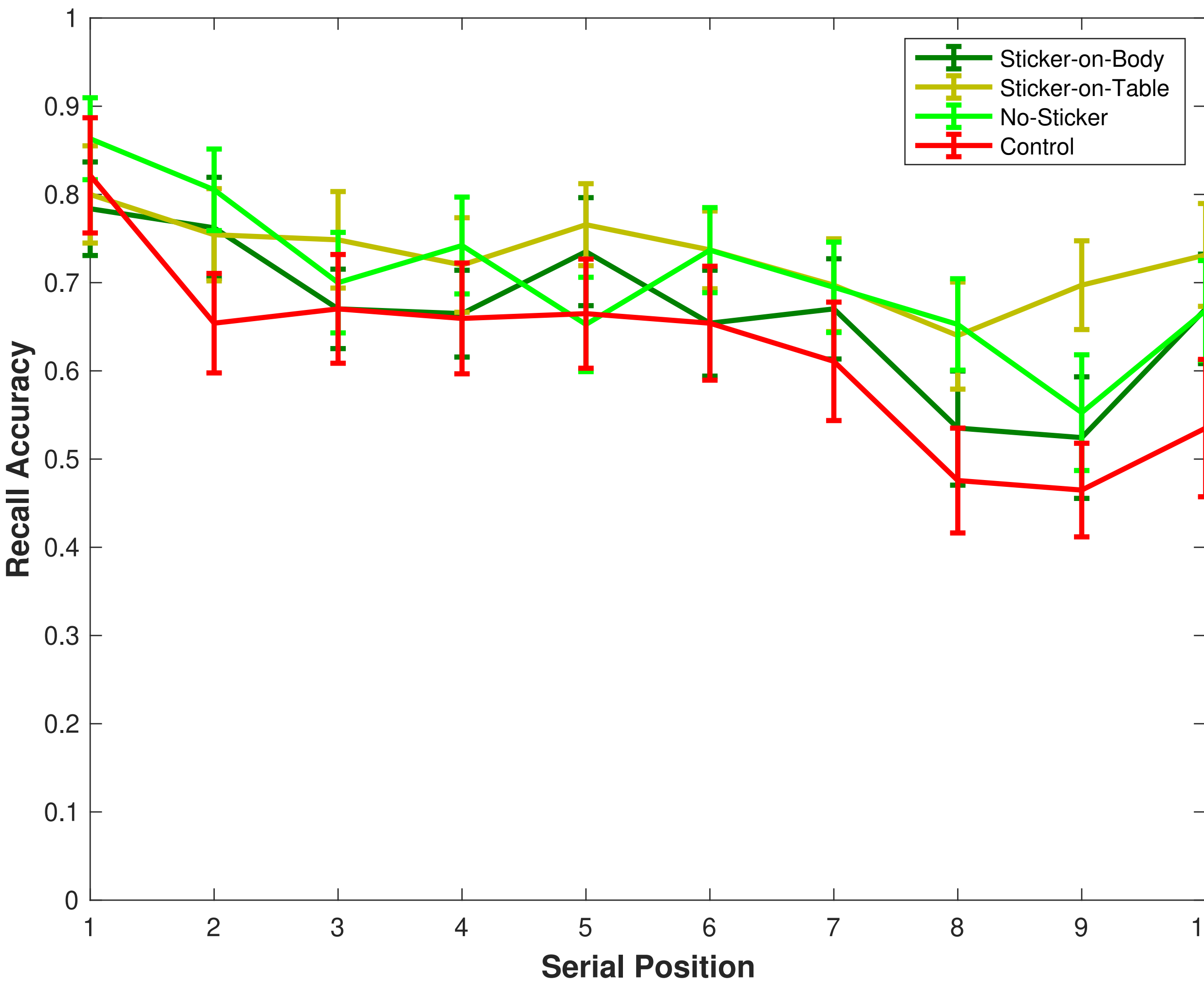


Figure 0.8. Experiment 2 - item memory (lenient scoring): Post-instruction recall accuracy for each group. Error bars are standard error of the mean corrected for subject variability (Loftus and Masson, 1994).

Scatter plots from individual difference questionnaires. Below are the scatter plots form the individual difference questionnaires (Figures 0.9, 0.10, 0.11). 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.

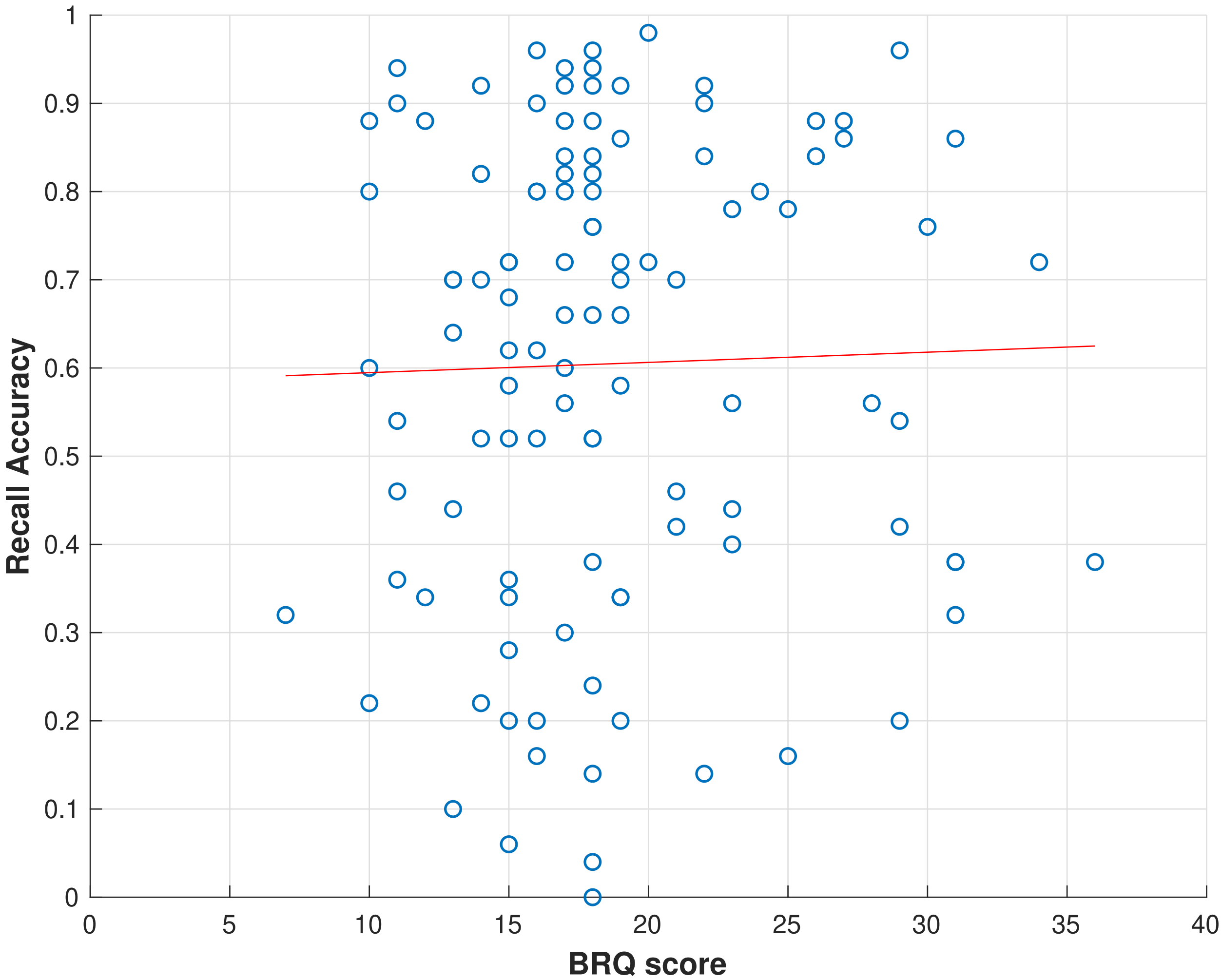


Figure 0.9. Experiment 2: Post-instruction recall accuracy by BRQ scores for all variants of the Body Scaffold combined

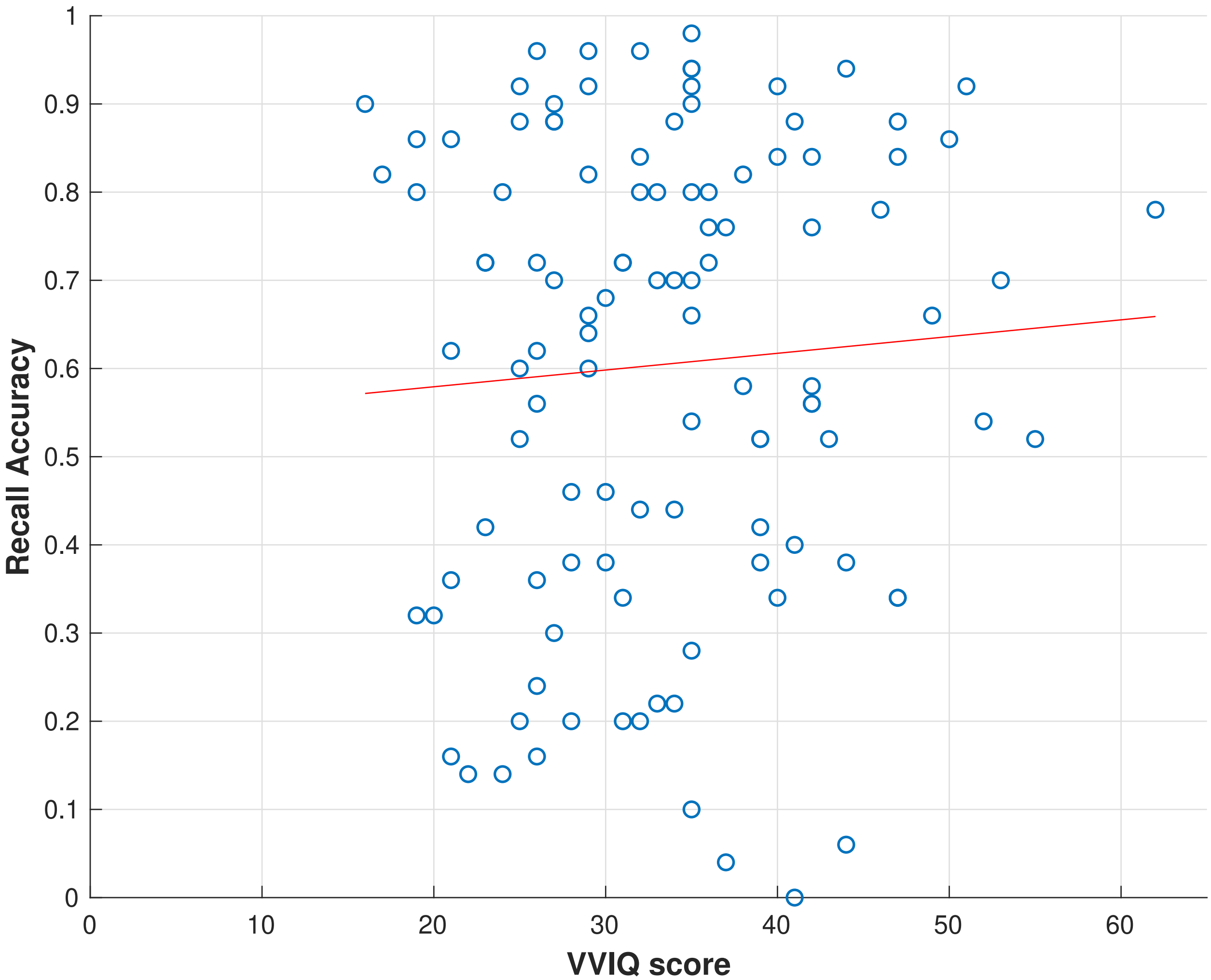


Figure 0.10. Experiment 2: Post-instruction recall accuracy by VVIQ scores for all variants of the Body Scaffold combined

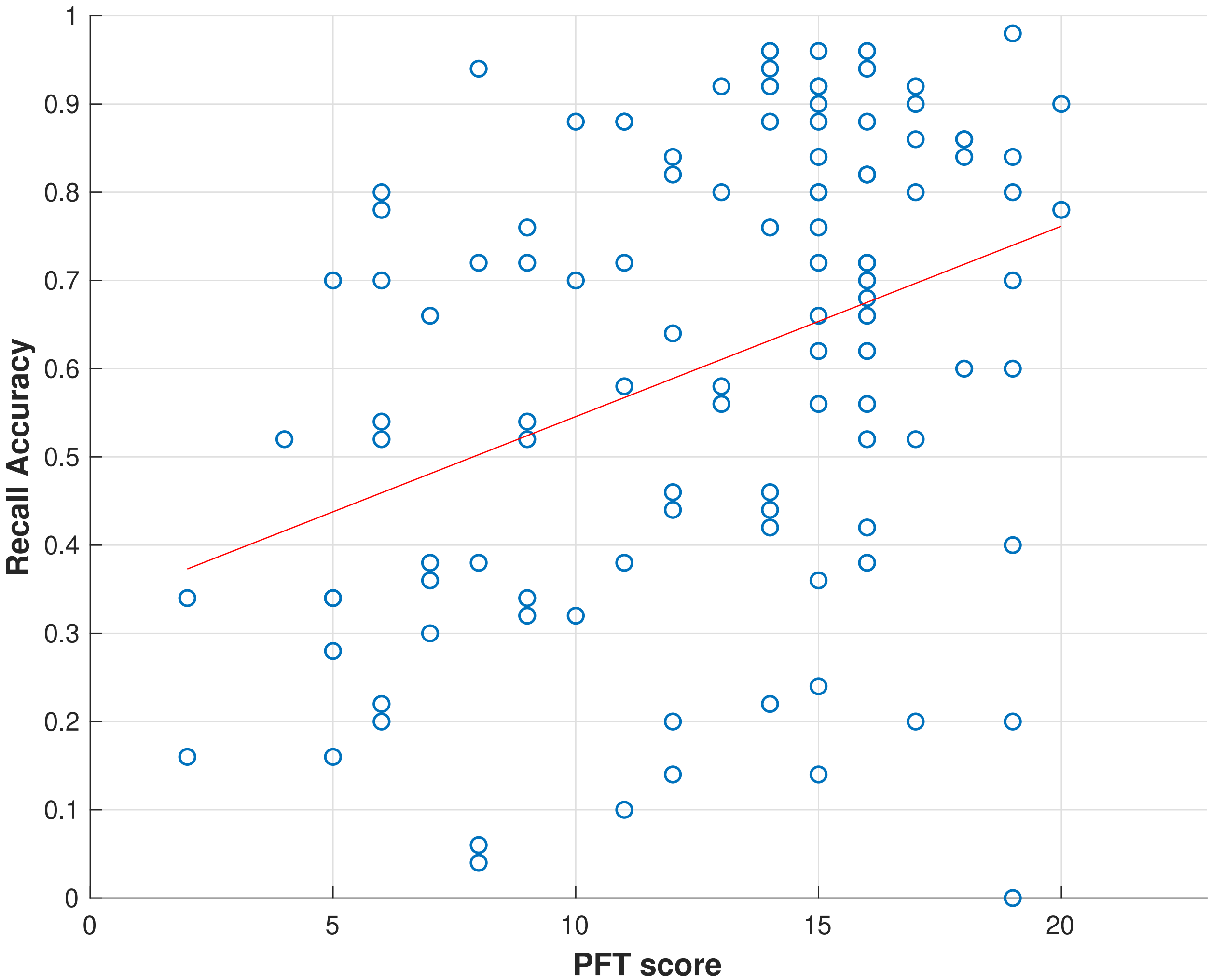


Figure 0.11. Experiment 2: Post-instruction recall accuracy by PFT scores for all variants of the Body Scaffold combined

Supplementary Materials B

Example Scaffolds

Body Scaffold example 1

- foot
- shin
- knee
- thigh
- 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 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 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