Effectiveness of the method of loci is only minimally related to factors that should influence imagined navigation

Jeremy B. Caplan[1,2], Eric L. G. Legge[3,1],

Bevin Cheng[1], and Christopher R. Madan[4,1]

[1]Department of Psychology and [2]Neuroscience and Mental Health Institute,

University of Alberta, Edmonton, AB, T6G 2E9, Canada.

[3]Department of Psychology, MacEwan University, Edmonton, AB, T5J 4S2, Canada.

[4]School of Psychology, University of Nottingham, Nottingham, NG7 2RD, UK.

Author Note

Corresponding author: Eric Legge. eric.legge@macewan.ca. Department of Psychology, City Centre Campus, MacEwan University, Edmonton, Alberta T5J 4S2, Canada, Tel: +1.780.497.5044.

We thank James Paterson for feedback on the manuscript, Isabel Lek for help with participant testing and data-scoring, and Marcia Spetch and Vadim Bulitko for discussions that led to the design of the study. Supported in part by the Natural Sciences and Engineering Council of Canada (NSERC) and the Alberta Ingenuity Fund. We thank Valve Corporation for providing a license for Half-Life 2 for educational purposes. ELGL and CRM were supported by NSERC Alexander Graham Bell Canada Graduate Scholarships (Doctoral-level).

Abstract

The method of loci is arguably the most famous mnemonic strategy, and is highly effective for memorizing lists of non-spatial information in order. As described and instructed, this strategy apparently relies on a spatial/navigational metaphor. The user imagines moving through an environment, placing (study) and reporting (recall) list-items along the way. However, whether the method relies critically on this spatial/navigation metaphor is unknown. An alternative hypothesis is that the navigation component is superfluous to memory success, and the method of loci is better viewed as a special case of a larger class of imagery-based peg strategies. Training participants on three virtual environments varying in their characteristics (an apartment, an open field, and a radial-arm maze), we asked participants to use each trained environment as the basis of the method of loci to learn five 11-word lists. Performance varied significantly across environment. However, the effects were small in magnitude. Further tests suggested navigation-relevant knowledge and ability were not major determinants of success in verbal memory, even for participants who were confirmed to have been compliant with the strategy. These findings echo neuroimaging findings that navigation-based cognition does occur during application of the method of loci, but imagined navigation is unlikely to be directly responsible for its effectiveness. Instead, the method of loci may be best viewed as a variant of peg methods.

Effectiveness of the method of loci is only minimally related to factors that should influence imagined navigation

Introduction

Unlike memory strategies participants come up with spontaneously in the lab, several mnemonic strategies are known to produce superlative performance, even after modest training (e.g., Roediger, 1980), and are adopted by world memory champions (e.g., Maguire, Valentine, Wilding, & Kapur, 2003). To understand the full range of human memory performance, it is important to understand the cognitive mechanisms by which such superior memory strategies operate. The method of loci is arguably the most well known mnemonic strategy, which predates written history, that produces excellent memory for verbal serial lists (Bower, 1970; Roediger, 1980; Yates, 1966) and is used by many of the best memorizers in the world to remember extremely large amounts of information in order (including, remarkably, 2^{16} digits of π; Raz et al., 2009). In the method of loci, the memorizer imagines moving through a familiar environment, placing list-items (typically, words) in locations (loci) along the path. To recall, one imagines navigating along the same path, reporting objects along the way. The method of loci is thus thought to rely on an imagined spatial/navigational substrate to support memory for materials that need not, themselves, be spatial at all.

Here we consider a line of thought by Bower (1970), who cast doubt on whether navigation-like cognition is central to the success of the method of loci. He argued that the effectiveness of the method of loci should not be attributed to the cues being self-generated, spatial locations, nor even imaginable. He noted the method of loci is formally similar to peg techniques, whereby one links list items to a pre-defined set of peg words or images. Bower argued that the fact that numerical peg systems can produce equivalent performance to the method of loci (see also Roediger, 1980; Wang & Thomas, 2000) suggests the spatial and navigational characteristics of the strategy may not be essential (see also Bouffard, Stokes, Kramer, & Ekstrom, 2018). This raises the possibility that even when participants apply the method of loci, navigational aspects of the strategy may play no direct role in serial-list memory. Alternatively, as recently argued, for example, by Rolls (2017), the method of loci may be effective because it activates the neural navigation system, which, in turn, is specialized for memory (see also Bouffard et al., 2018, who presented similar reasoning comparing the method of loci to autobiographical and everyday procedural peg systems). Indeed, a wealth of evidence has suggested that the hippocampus and neighboring regions contain neurons that are selective to environmental features (e.g., place cells, border cells) that also code for memories for features and events encountered at those locations (e.g., Moser, Rowland, & Moser, 2015), suggesting a synergy between navigation processes and episodic memory, that might underlie the effectiveness of the method of loci. Several neuroimaging studies have reported navigation-like brain activity during application of the method of loci. For example, Müller et al. (2018) found that memory "athletes" had higher correlations between activity in hippocampal and caudate regions of interest, in turn, proportional to their world ranking. This could reflect integration of cognitive-map/allocentric and stimulus–response/egocentric navigation systems. However, the functional connectivity measures were not related to within-subjects performance, leaving open the possibility that this activity accompanies the method of loci, but may not necessarily be directly responsible for its effectiveness. Other neuroimaging studies can be viewed in a similar way; that is, evidence that participants imagine navigating while applying the method of loci, but without directly tying brain activity to successful application of the strategy (e.g., Dresler et al., 2017; Kondo et al., 2005; Maguire et al., 2003; Mallow, Bernarding, Luchtmann, Bethmann, & Brechmann, 2015; Müller et al., 2018; Nyberg et al., 2003).

Legge, Madan, Ng, and Caplan (2012) found that for naïve participants, the method of loci was as effective when used with a novel, just-learned, extremely familiar environment such as the participant's own house. extremely familiar environment such as the participant's own house. Here, we take advantage of this finding, by training participants on virtual environments with particular navigation-relevant characteristics, and then asking them to use those newly learned environments as the basis for the method of loci to learn multiple word-lists. To maximize the chance of observing an effect, we exaggerated the differences in topological and other spatial characteristics of three environments, with the goal of varying how amenable they would be to imagined navigation. The first environment resembled an apartment, with multiple rooms, and numerous lines of sight from one room to another. The function of each room was self-evident (bathroom, bedroom, etc.). The second environment was an open field with no delineated sub-regions, but by turning within the environment, nearly all objects were visible from all standing locations. The third environment was an eight-arm radial-arm maze, with no distinctive features apart from distinct objects placed at the end of each arm. To see those objects, one had to navigate to the very end of an arm and turn. Thus, multiple potential loci were never visible simultaneously in one view, and no locus was visible from any other locus. All environments had multiple distinct landmarks (objects), with no repetitions.

More specifically, the Apartment, Open Field, and Radial Arm environments differed from one another in three critical ways:

1. *Conceptual familiarity:* Familiarity and experience with an environment is a strong predictor of wayfinding accuracy (Li & Kippel, 2016; O'Neill, 1992), and has been shown to influence how quickly survey knowledge of a space is acquired through navigation (Thorndike & Hayes-Roth, 1982). The Apartment environment was designed to align with participants' preconceptions of a home, thus allowing participants to benefit from their conceptual familiarity with such spaces when navigating. It was composed of multiple rooms with distinct, familiar functions (bathroom, bedroom, etc.). The Open Field environment was one single room, with clusters of objects that were not related to one another in any obvious way. However, such open-space environments are commonly encountered in everyday life (i.e., warehouses), but likely are less familiar to participants than the Apartment environment. The Radial Arm environment would have been the least

similar to typical environments participants may have experienced, and thus was considered to have low conceptual familiarity.

2. *Boundary:* The Apartment environment had boundaries resembling rooms within a typical home, thus providing robust intra-maze cues that could serve as additional navigational landmarks to the objects within (Chan, Baumann, Bellgrove, & Mattingley, 2012). As way-finding accuracy has been found to be influenced by the number and placement of landmarks (Heft, 1979; Jansen-Osmann & Fuchs, 2006), it is likely that these additional intra-maze cues could enhance how quickly and well participants learned the space. As well, the discrete boundaries provided by the various rooms provided a method of subdividing object clusters into smaller, schema-appropriate subsets, which may have aided in object recall. The Open Field environment had a single outer boundary. However, compared to the Apartment environment, the single, large room design of the Open Field environment did not allow for the formation of boundaries that could serve as additional intra-maze cues for navigation, or as a method of subdividing object clusters into smaller subsets. Lastly, the Radial Arm environment had rich, rotationally symmetric boundaries, and unique sub-boundaries provided by the alcoves at the end of each arm (not visible from the centre of the environment). As such, similar to the Apartment enviornment, the alcoves that contained the loci objects provided additional intra-maze cues that could serve as landmarks. However, due to the nature of the environment, wherein these alcoves were not visible from the center of the environment, and that from the center of the environment each arm looked identical, it is unlikely these additional intra-maze cues would be useful for aiding navigation and maintaining one's orientation in the space. This design also limited how well participants could subdivide object clusters found in each boundary into distinct units that could be used for navigation.

3. *Lines of sight:* The ability to see one location while standing in another location contributes to the "spatial syntax" of an environment; by enabling participants to understand the relationships of locations to one another, the prevalence of numerous lines

of sight can facilitate wayfinding within an environment (T. R. Herzog & Leverich, 2003; M. H. Herzog, Fahle, & Koch, 2001; Kim & Penn, 2004). The Open Field environment had the highest availability of clear lines of sight from one location to another, followed by the Apartment environment, with the Radial Arm environment having no visibility of one locus (cluster of objects) to another.

If the navigational metaphor implied by the method of loci is superfluous to its efficacy, we predict similar levels of verbal memory for all three environments when used as the basis of the method of loci. If, in contrast, the spatial or navigational aspects of the strategy do contribute to its effectiveness, we expect performance to differ substantially (a similar argument has been made by Rolls, 2017). Specifically, we expected the Apartment environment would be superior to the Radial Arm environment because many lines of sight would be available to participants to relate locations and superior to both the Open Field and Radial Arm environments because of the robust intra-maze cues provided by room boundaries and the fact that it should make the most spatial "sense" to participants, given its resemblance to a home schema. We expected the Open Field environment would be less effective than the Apartment environment, as it provides fewer intra-maze cues due to the lack of room boundaries, had less conceptual familiarity to participants, and offers less well defined loci (i.e., objects could not easily be categorized by room type as in the Apartment environment). However, we suspected the Open Field environment would be more effective than the Radial Arm environment due to greater lines of sight, the comparatively higher degree of conceptual familiarity, and rotational symmetry of the boundaries which could lead to disorientation. Thus, we expected the Radial Arm environment to produce by far the worst verbal memory.

Finally, prior studies have found that participants are not all compliant with strategy instructions. Moreover, for the method of loci groups, only participants who self-reported as compliant with the strategy instructions showed an advantage of the method of loci compared to an uninstructed control group (Legge et al., 2012). Here we went one step further. Rather than relying on subjective report, we devised an objective measure of compliance (which we also expected would increase compliance rates). We asked participants to talk out their strategy while studying the word lists, while the experimenter was present in the testing room, and rewarded participants for each list on which they were heard to implement the method of loci.

We sought to test four hypotheses motivated by the idea that the method of loci operates fundamentally through an imagined navigation mechanism:

H1) *Effect of environment.* If imagining navigating an environment functions much like navigating a real environment, then factors that influence ease of wayfinding and orientation should influence the efficacy of the method of loci. The prediction is large differences in serial-recall accuracy should be seen across the three environments, due to differences in conceptual familiarity with the layout of each environment, lines of sight, and available boundaries. Considering that, with similar methods, Legge et al. (2012) found advantages of 0.10–0.14 (proportion of words recalled) of participants applying the method of loci (compliant) over control participants, we expected effects of environment to be roughly this magnitude.

H2) *Effect of compliance.* The more compliant participants are with the method of loci strategy, the more they are imagining navigating the respective environment. The prediction is that the effect of environment should interact with compliance; thus, serial-recall accuracy will be more sensitive to environment for more compliant than for less compliant participants.

H3) *Influence of knowledge of the environment.* We are asking participants to use a particular environment as the substrate of the method of loci. If verbal memory is stored within a particular environment, then the quality of knowledge of the environment should influence recall accuracy. The prediction is that greater performance on a blueprint reconstruction task for an environment should correlate with greater serial-recall accuracy using that environment. However, one expects that there will be individual variability in overall memory skill, which would produce some level of positive correlation between any two memory performance measures. Thus, a more informative test is that the correlation between blueprint accuracy and serial-recall accuracy should be greater for high compliant than low compliant participants.

H4) *Effect of video game experience.* Finally, participants who have extensive experience with first-person perspective video games may be superior to non-gamers in virtual navigation tasks (Richardson, Powers, & Bousquet, 2011; Smith & Du'Mont, 2009). If the method of loci depends critically on imagined navigation, one would predict that video game experience and first-person game experience, in particular, would correlate positively with serial-recall accuracy using the method of loci (but see West et al., 2018, discussed below).

Methods

As advised by Simmons, Nelson, and Simonsohn (2012), we affirm that "we report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study."

Participants

Students in introductory psychology courses at the University of Alberta ($N = 179$) participated in a single experimental session lasting less than two hours, in exchange for partial course credit. Participants, run individually, were also remunerated with up to $3.25 CAD depending on their compliance level (described below). The procedures were approved by a University of Alberta ethical review board. Six participants reported experiencing motion sickness when exploring the virtual environment and were excluded from data analyses, leaving 173 participants (115 female, 52 male, 6 unreported; ages 17–35 years, mean $\pm$ SD $= 19.5 \pm 2.5$ years, 6 unreported). Each participant repeated this procedure with all three environments, with environment-order counterbalanced across participants ($N =$ RFA: 28, RAF: 29, FRA: 29, FAR: 30, ARF: 29, AFR: 28, where A =

apartment, F = open field, R = radial arm maze). Our sample size was not precisely determined beforehand, but was a convenience sample, pending availability of experimenters and testing rooms, with the general aim to collect substantially more data than in our earlier Legge et al. (2012) study, which had included 142 participants. The prior study included an uninstructed control group, whereas here, all groups were asked to apply the method of loci. We thus anticipated that effect sizes might be comparable to, or smaller than that in the 2012 study, for between-subjects comparisons (i.e., the first environment for each participant).

Materials

Environments. The basic virtual environment methods were based on Legge et al. (2012). Three environments (Figure 1) were created using Hammer (Valve Software; Bellevue, WA) and Garry's Mod (Facepunch Studios; Walsall, England), game-editing software for use with the game Half-Life 2 (Valve Software; Bellevue, WA). Objects were selected to be distinct, both within and between the three environments.

The Apartment environment (Figure 1a,d) was one of the environments used by Legge et al. (2012). It included the following identifiable rooms: two bedrooms, a bathroom, a kitchen, a living room, and a small laundry room. Participants started exploring in one of the bedrooms.

The Open Field environment (Figure 1b,e) consisted of an octagonal room with a high ceiling and objects placed throughout. Participants started exploring in the center.

The Radial Arm Maze environment (Figure 1c,f) consisted of a central area with eight arms, each of which had an alcove at the end; after turning right, objects could be seen in each alcove. By design, there were no distinguishing features outside of the alcoves and the maze was enclosed, so no extramaze cues were visible. Thus, participants had to rely on path integration more than with the other environments, to keep track of their current location and relative locations of the other arms and distinctive objects.

Participants started exploring in the center.

Serial lists. The serial-recall methods were based on Legge et al. (2012). Each list comprised 11 words[1] drawn from a 219-word pool[2]) based on properties from the MRC Psycholinguistic Database (Wilson, 1988) with Kucera-Francis written frequency 1–20 per million, concreteness and familiarity greater or equal to 550, length of 1–2 syllables, imageability at least 450, and nouns (both common part of speech and comprehensive synthetic category).[3] Words were assigned to lists with a new randomization for each participant.

Procedure

The main experiment (excluding the end-of-session questionnaire and blueprint recall) was run on one of two iMac (model: 5.1) computers with a 15” screen (1440×900 pixels). Prior to the main experiment, participants had a practice list (consisting of 11 words from the same word pool as the main task) for serial recall, and had practice moving around in a very simple virtual environment, as in Legge et al. (2012), a small, empty, square room, to train participants to navigate with the keyboard and mouse. Before exploring the first environment, participants were instructed on how to apply the method of loci using a written description of the method, and allowed to ask questions about how to apply the strategy, as in Legge et al. (2012). By design, we left the specific application of the

[1] List length eleven was chosen because it suggests no obvious chunking pattern.

[2] after exclusions; the words CEMENT, CLOCK, DRESSER, LAUNDRY, PIGEON, SHOWER, STATUE, STOOL, STOVE, STRAW, TOILET, and TRASH were removed because these named objects were found in one of the virtual environments. Numerous other uncontrolled word characteristics could influence probability of recall (e.g., Lau, Goh, & Yap, 2018), and random sequences of words might carry idiosyncratic meaning or vary in ngram frequency. However, because word lists were constructed in a new random order for each participant, this is expected to operate as random error, not likely to be systematically related to our Environment variable.

[3] See the appendix for the full word pool.

method, including choice of number and identification of the loci, up to the participants.

For each environment, participants completed three phases: 1) Environment training, exploring and learning the environment, 2) Serial recall using the just-learned environment, consisting of 5 novel lists of 11 words each, comprised of a study phase followed by one attempt at serial recall. 3) Blueprint recall, to test the participant's knowledge of the just-learned (and used) environment. This sequence was repeated for each of the three environments for all participants (within-subjects), with environment-order counterbalanced across participants.

At the end of the session, participants completed a questionnaire regarding their prior familiarity with the method of loci and experience with both computer games in general, and first-person games, specifically (7-point Likert scale).

Environment training. Participants were given up to five minutes to freely explore in order to learn the environment. If the participant did not terminate by the five-minute mark, it was terminated by the experimenter.

Serial Recall. As in Legge et al. (2012), words were presented visually, sequentially and centrally on the screen for 5000 ms each, followed by an inter-stimulus interval of 150 ms. Following the last word of a list, participants were asked to recall by typing the words, in order, to the best of their ability. Each word was followed by the "Enter" key, after which the screen cleared in preparation for the next response. Participants were instructed to type "PASS" to skip the current list position whenever they could not remember a particular word.

To check for strategy compliance, participants were additionally asked to continuously verbally describe how they were memorizing the lists during study (but not during recall). The experimenter, who was present in the testing room, scored compliance during testing to compute the bonus payment. If at least six of the eleven words (more than half) were mentioned as part of the method of loci, the list was scored as compliant (but participants were not told the criterion explicitly). Participants were rewarded with

$0.25 CAD per compliant list. Sessions were recorded with an audio recorder. Author BC later listened to the recording to verify compliance for the purpose of the data-analyses. Compliant vocalizations included both list words and words describing features of the environment. An example of a vocalization that was scored as compliant is: "**Flute** on the couch. **Nurse** nursing a patient on the couch. Patient dying on the couch. **Crypt**. **Spice** on the cabinet." (list words are in boldface, features of the environment are underlined). Non-compliant recordings were often silent; an example of a vocalization that was scored as non-compliant is, "Giant **birch** tree fell to make a **saloon**. Yeah, I don't really remember much from this one." (Silence for the rest of the recording.)

Our aim was to assess whether the participant attempted to apply the method of loci, without demanding excessive detail and elaboration. In scoring compliance, our criterion erred on the side of giving the participant the benefit of the doubt, in case participants sometimes failed to keep up with the overt verbal protocol, but were making a concerted attempt to apply the strategy. Thus: A word was judged compliant if the word was judged to have been mentioned verbally in some relation to the environment; e.g., how/where it was placed within the environment. A list was classified as compliant if at least half (≥ 6 words) were scored as compliant.

Blueprint recall. Following the five serial-recall lists with a given environment, participants were given a blank blueprint of that most recent environment (Figure 1a–c), and were asked to recall which objects were in various locations throughout the environment. The blueprint was labelled with the numbers 1–35 (Apartment) and 1–32 (Open Field and Radial Arm).[4] Each number corresponded to one distinct object or object cluster. The response sheet contained numbered lines, the numbers corresponding to

[4] We noticed too late that the number of distinct objects differed across environments. However, the difference of three objects is small compared to the differences in blueprint accuracy (Figure 4). Still, to maintain a more fair comparison, blueprint accuracy is always analyzed as a proportion of the total (35 or 32 objects, respectively).

number labels displayed on the blueprint, and participants wrote their remembered descriptions of the objects in each corresponding location. If at least one object was correctly recalled at the correct locus, that locus was scored as correct (strict scoring). For the Radial Arm environment, because of its rotational symmetry, the blueprint was re-scored at all eight rotations, and the highest score was used. For lenient scoring, a locus was correct if an object from the locus was written anywhere.

Data from the serial recall, blueprint recall, and questionnaires are available from `https://osf.io/yvtp2/`.

Data analysis. Analyses were conducted using repeated-measures ANOVAs in SPSS (IBM Corp.; Somers, NY). Effects were considered significant based on an alpha level of 0.05. Greenhouse-Geisser correction was applied wherever violations of sphericity were found. Post-hoc *t*-tests on significant Environment effects were Bonferroni-corrected. Data are analyzed both for the full sample, and for a highly compliant subset.

Results

We first report compliance rates, and integrate compliance level into the remaining analyses. Next, we report the effect of environment. Then we analyze performance on the blueprint tasks, looking for effects of environment, and then relate blueprint accuracy to serial-recall performance. Finally, we test for potential effects of prior experience, knowledge of the method of loci, and video-gaming experience, on application of the method of loci as well as the effect of environment.

Compliance

First we analyzed compliance rate as a function of the environment participants were to use as the basis of the method of loci. If any differences were to arise, our prediction was Apartment > Open Field > Radial Arm, part of which was supported: A three-way mixed, repeated-measures ANOVA, with design Environment Order[6] (between-subjects) $\times$ Environment[3] $\times$ List Number[5], where List Number (1–5) was for a particular environment, revealed a main effect of Environment (Figure 2), $F(2, 334) = 4.97,\ MSE = 0.19,\ p < 0.01,\ \eta_p^2 = 0.029$. The only significant post-hoc pairwise t-tests found compliance was lower for the Radial Arm environment ($M \pm SEM = 0.624 \pm 0.032$) than for the Apartment environment ($M = 0.688 \pm 0.031$; $p < 0.01$; Open Field environment: $M = 0.670 \pm 0.031$). The three-way interaction approached significance ($p = 0.07$) but given the large sample size and small effect size ($\eta_p^2 = 0.04$), we did not follow this up further. The remaining effects were non-significant ($p > 0.3$). These null effects suggest little effect of either fatigue or practice on compliance levels (no change over successive lists), and that the order in which participants experienced the three environments did not substantially influence their compliance (no interaction with Environment Order).

For the remaining analyses, two sets of analyses were conducted: first, using all participants, regardless of compliance; and second, only with the 66 participants who were fully compliant, having used the latest environment as the basis of the method of loci on all 15 lists.

Serial recall

Serial recall was scored in two ways: (a) strict scoring, in which an item was correct if it was recalled in the correct position, sensitive to order-errors, and (b) lenient scoring, in which an item was scored as correct if it came from the current list, regardless of order. To assess Hypothesis H1, we were first interested in whether accuracy, particularly with strict scoring, would depend on environment, predicting Apartment > Open Field > Radial Arm. Regarding Hypothesis H2, we tested whether compliance would modulate (increase) the effect of environment.

Strict scoring. A mixed, repeated-measures ANOVA on strict serial-recall accuracy (Figure 3), with Environment as a within-subjects factor and Compliance and Environment Order as between-subjects factors, found only a significant main effect of

Environment, although with a small effect size, $F(2, 322) = 8.74,\ MSE = 0.010,\ p < 0.001,\ \eta_p^2 = 0.051$. Post-hoc pairwise t-tests supported part of our prediction; the Radial Arm environment produced lower serial-recall accuracy ($M \pm SEM = 0.203 \pm 0.012$) than the Apartment ($M = 0.249 \pm 0.014$) and Open Field environment ($M = 0.232 \pm 0.014$), $p < 0.05$, which did not differ from one another. Note that the largest difference (0.046) was far smaller than the anticipated magnitude (0.10–0.14 as in the contrast between method of loci and control reported by Legge et al., 2012).

The same outcome was obtained when analyzing the Compliant participants alone, Environment[3]×Environment Order[6]: Only the main effect of Environment was significant, again with a small effect size, $F(2, 120) = 3.49,\ MSE = 0.011,\ p < 0.05,\ \eta_p^2 = 0.055$, but with the only significant post-hoc test being Apartment > Radial Arm, $p < 0.01$ ($M = 0.236 \pm 0.023,\ 0.229 \pm 0.022,\ 0.193 \pm 0.021$ for Apartment, Open Field and Radial Arm environments, respectively).

Lenient scoring. In the same analyses applied to lenient-scoring accuracy, the first ANOVA again revealed a significant main effect of Environment, $F(2, 322) = 22.5,\ MSE = 0.008,\ p < 0.001,\ \eta_p^2 = 0.12$, but this time, the Bonferroni-corrected post-hoc t-tests were all significant, with the predicted pattern Apartment > Field > Radial ($p < 0.05$; $M = 0.613 \pm 0.012,\ 0.573 \pm 0.012,\ 0.547 \pm 0.013$, respectively). This main effect was qualified by a significant Environment×Environment Order interaction, $F(2, 322) = 2.17,\ MSE = 0.008,\ p = 0.021,\ \eta_p^2 = 0.063$, but no other effects were significant. Simple effects at each level of Environment Order found the main effect of Environment significant for the two groups that had the Radial Arm environment first and for one group that had the Open Field followed by the Radial Arm environment, but no post-hoc tests reached significance. For the other group that received the Open Field first, and for both groups that had the Apartment environment first, the main effect of Environment was non-significant.

Following up with just the Compliant participants, only the main effect of Environment was significant, $F(2, 52) = 14.76,\ MSE = 0.006, p < 0.001,\ \eta_p^2 = 0.13$. Post-hoc t-tests found the Apartment environment ($M = 0.623 \pm 0.035$) was more accurate than both the Open Field ($M = 0.603 \pm 0.027$) and Radial Arm ($M = 0.532 \pm 0.026$) environments, $p < 0.005$, but the latter did not significantly differ, thus supporting part of our predicted pattern. Unlike the all-inclusive analysis, Environment Order did not significantly influence lenient-scored serial-recall accuracy, either as a main effect or interaction with Environment ($p > 0.2$).

In sum, environments did differ in efficacy. However, the effects were rather small in comparison to the large-size effects we expected due to the vast differences across environments. More concerning, the lack of interaction between Environment and Compliance, with both measures of serial recall accuracy, challenges Hypothesis H2, and casts doubt on the idea that the effect of environment was related to application of the method of loci. This suggests that the method of loci is resilient to vast changes in topological and visuospatial characteristics of the substrate-environment.

Blueprint recall

We expected participants' memory of the environments themselves to differ across environments, as Apartment > Open Field > Radial Arm. To evaluate memory for the environments (Figure 4), we analyzed blueprint performance, with strict-scoring of accuracy as the measure, and Compliance and Environment Order as between-subjects factors. Compliance was not a significant main effect ($p > 0.8$), nor did it interact with Environment ($p > 0.7$), suggesting blueprint recall was unrelated to participants' ability or willingness to implement the method of loci with the overt protocol. Environment Order was a non-significant main effect ($p > 0.7$) and the interaction of Environment Order with

Environment was also non-significant ($p > 0.6$). This suggests that there was no overall learning-to-learn, nor fatigue effect, on learning the environments. The main effect of Environment was significant, $F(2, 322) = 87.3,\ MSE = 0.065,\ p < 0.001, \eta_p^2 = 0.35$. Post-hoc t-tests found blueprint accuracy was higher for the Apartment ($M = 0.683 \pm 0.026$) environment than both the Open Field ($M = 0.357 \pm 0.026$) and Radial Arm ($M = 0.354 \pm 0.021$) environments ($p < 0.001$), but the Open Field and Radial Arm environments did not significantly differ from one another ($p > 0.05$).

The same ANOVA design using lenient scoring also found all effects to be non-significant ($p > 0.15$) apart from the main effect of Environment, $F(2, 342) = 62.6,\ MSE = 0.018,\ p < 0.001,\ \eta_p^2 = 0.28$, but unlike strict-scoring, all post-hoc tests were significant ($p < 0.05$), with rank-ordering: Apartment > Open Field > Radial Arm as predicted ($M = 0.814 \pm 0.020,\ 0.776 \pm 0.020,\ 0.656 \pm 0.018$, respectively).

In sum, independent of environment order and compliance, participants could remember more objects along with their spatial locations in the Apartment environment than the other two environments. When placement location was ignored, participants remembered the most objects from the Apartment environment and the fewest from the Radial Arm environment.

These large differences in knowledge of the environments confirm that the manipulation was successful in influencing spatial knowledge, and that a large superiority of the Apartment environment over the other two environments should have been expected in the analyses of serial recall accuracy. Thus, the small magnitude of the effect of Environment, and lack of interaction with Compliance in the previous section would seem to indicate that high-quality spatial knowledge is not critical to the success of this strategy.

Blueprint–serial-recall correlations. To test H3, that knowledge of an environment influences recall accuracy when using that environment as the basis for the method of loci, we first asked if mastery of an environment influenced serial-recall performance. We computed Pearson correlations between blueprint-accuracy (strict and lenient) and serial-recall-accuracy (strict and lenient) for each of the three environments. All twelve correlations were significant except one, which approached significance ($p = 0.052$), all with the same sign and similar magnitude, with $r(171)$ ranging from 0.148 to 0.320.

Hypothesis H3 implies that by restricting the correlation analyses to compliant-only participants, the coupling between environmental knowledge and serial recall accuracy should grow strong. However, the correlations generally reduced in magnitude, and all became non-significant for strict scoring of serial recall (r ranged from –0.004 to 0.213). For lenient scoring of serial recall, the only correlations that reached significance were: for the Radial Arm environment, strict-scoring of blueprint accuracy, $r(171) = 0.330,\ p = 0.007$ and for the Apartment environment, lenient scoring of blueprint accuracy, $r(172) = 0.248,\ p = 0.044$. A Wilcoxon rank-sum test found that the correlation significantly decreased ($p = 0.0024$) from the all-inclusive to compliant-only correlations (Figure 5).

These correlations explained, at the very most, only 11% of the variance. Combined with the observation that the correlations did not grow stronger when restricted to fully compliant participants, these small positive correlations may merely reflect generic individual differences in memory skill. Supporting this, when each correlation (including all participants) was rerun as a partial correlation, controlling for the corresponding blueprint accuracy for the other two environments, all correlations with lenient-scored blueprint accuracy became non-significant ($p > 0.05$). For correlations involving strict-scored blueprint accuracy, both the correlations remained robustly significant for the Radial Arm environment (strict-scored serial-recall correlated with strict-scored blueprint accuracy: $r_p(169) = 0.188,\ p = 0.014$; lenient-scored serial-recall correlated with strict-scored blueprint accuracy: $r_p(169) = 0.270,\ p < 0.001$) as well as for the Open Field environment, strict-scored serial-recall correlated with strict-scored blueprint accuracy ($r_p(169) = 0.191,\ p = 0.012$). This leaves room for the possibility that mastery of an environment influences serial-recall, but with a rather small effect size, explaining at most, $R^2 = 7.3\%$ of the variance— not what one would predict if the spatial-navigational properties of the environment influence the efficacy of the method of loci, and calling into question Hypothesis H3.

Gaming experience

Finally, we tested Hypothesis H4, that experience with video games, especially first-person perspective games, may make it easier for participants to learn the environments and to execute the method of loci with those environments, given increased practice imagining navigated virtual worlds. However, Pearson correlations between experience with video games in general, and first-person games in particular, with serial-recall accuracy and blueprint accuracy (both strict and lenient, separately for each environment, 24 correlations in total) produced only small correlations that were not significant, $|r| < 0.145,\ p > 0.05$, with two exceptions: first-person gaming experience correlated inversely with lenient- scored serial-recall accuracy with both the Apartment ($r = -0.151,\ p < 0.05$) and Radial Arm ($r = -0.227,\ p < 0.01$) environments, opposite our prediction. These largely null effects (illustrated for first-person games and serial recall accuracy in Figure 6) are inconsistent with Hypothesis H4, and further reinforce the idea that virtual navigation is not essential for the method of loci to be effective.

However, it has been shown that first-person video-game experience does not necessarily lead to only one form of learning. West et al. (2018) found that participants who applied a spatial strategy had increased hippocampal grey-matter volume, whereas a non-spatial, response-based strategy resulted in decreased hippocampal grey matter. This raises the possibility that our participants with first-person gaming experience may not have been compliant, or if they were, may still not have been able to virtually navigate effectively. To disentangle these factors,[5] we conducted a mixed, repeated-measures

[5] We thank an anonymous reviewer for this suggestion.

ANOVA on Environment, with between-subjects factors Compliance[2] and First-Person Experience[2] (simple yes or no response item, to facilitate interpretation of the results) on strict serial-position accuracy. First, the main effect of Environment was again significant, $F(2, 326) = 9.47,\ MSE = 0.010,\ p < 0.0001\ \eta_p^2 = 0.54$. The main effects of Compliance and First-Person Experience were both non-significant ($F < .3$, $p > 0.5$), but their interaction was significant, $F(1, 168) = 4.43,\ MSE = 0.28,\ p = 0.037,\ \eta_p^2 = 0.026$; simple effects suggested an advantage due to first-person gaming experience for compliant participants and a disadvantage due to first-person gaming experience for non-compliant participants, but neither was, itself, significant ($p = 0.067$ and 0.23, respectively). Most pertinent, the three-way interaction, Environment $\times$ Compliance $\times$ First-Person Experience was not significant, $F(2, 327) = 1.03,\ MSE = 0.010,\ p = 0.36,\ \eta_p^2 = 0.006$. The lack of interaction (also again, lack of two-way interaction, Environment $\times$ First-Person Experience, $F = 0.18,\ p = 0.84$) suggests that the magnitude of the effect of Environment was not dependent on prior first-person gaming experience, whether participants were compliant or not.

Discussion

With a large ($N = 173$) sample, we tested whether the navigational characteristics of the substrate-environment influences the efficacy of the method of loci for serial recall (e.g., Rolls, 2017). We succeeded in measuring a difference in serial-recall accuracy with a major manipulation of the layout characteristics of the environment. As detailed in the Introduction, our environments were expected to rank in order of spatial "sense" and navigability as Apartment > Open Field > Radial Arm; in some analyses, this predicted rank-ordering was found, but in other cases, the Apartment environment was far superior to the other two environments, which in turn, did not differ from one another. This could be due to numerous factors, the most obvious of which is that the Apartment environment may have been the most familiar, conceptually, and its characteristics related to familiar functions (bathroom, bedroom, etc.). However, the differences were small in magnitude (Figure 3), despite the large differences across the three environments. The small size of the effect of environment is at odds with what one would expect if the method of loci relied critically on imagined spatial navigation to succeed.

Blueprint accuracy correlated with serial-recall accuracy, but these correlations explained a small proportion of the variance, were not always significant, and were not larger when computed for perfectly compliant participants, suggesting at most, weak coupling of knowledge of the environment to effectiveness of the method of loci. As a caveat, survey knowledge is not necessarily correlated with first-person spatial knowledge (e.g., Montello, 1998; Rossano & Moak, 1998; Shelton & Gabrieli, 2002; Shelton & McNamara, 2004), so future studies testing knowledge of the environment differently may produce different results.

If mastery of the environment does little to determine how well participants apply the method of loci, and the effect of environment is quite small in magnitude, the implication is that the method of loci is *not* critically dependent on the spatial navigational properties of the substrate environment. The spatial-navigation premise of the method of loci may be unrelated to its success. Rather, the method of loci might be rather undistinguished among a broad set of peg methods, wherein the memorizer links new list items, often via mental imagery, to features of a pre-existing knowledge structure, such as peg lists. Still, the spatial navigation “cover-story” of the method of loci may, in part, explain its cultural resilience. The popularity of the strategy, and its survival over thousands of years, predating written history (Yates, 1966) may be due to the compelling nature of the idea of imagined navigation. Navigational imagery might even accompany application of the method of loci, as has been suggested by neuroimaging studies showing navigation-like brain activity associated with the method of loci (Dresler et al., 2017; Kondo et al., 2005; Maguire et al., 2003; Mallow et al., 2015; Müller et al., 2018; Nyberg et al., 2003), but such activity has yet to be directly linked to memory success. Navigational imagery may thus have little to do with the cognitive mechanisms by which people who apply this strategy actually perform verbal serial recall.

It should be noted that our participants were novices, and received only light instruction on the method of loci. With practice, a substantial dependence of serial-recall on the the spatial navigational properties of the substrate-environment might emerge. However, Legge et al. (2012) found that even novel environments produced enhanced serial recall in participants with little training or experience with the method of loci. If the method of loci is an evolutionarily conserved mechanism that supports memory by activating the hippocampus through navigation-like cognition, one would expect its effect to be immediately apparent, and not require significant expertise to emerge. Moreover, when Roediger (1980) asked participants to practice strategies at home, strict-scored serial-recall accuracy, for 20-word lists, was nearly as high for the peg-list method as for the method of loci. In the peg-list method, participants form an image combining each list-word with a word from a standardized, pre-memorized "peg list," comprised of highly imageable words. In Roediger's application, the pegs rhymed with corresponding numbers (1–gun, 2–shoe, . . .). In both the peg-list method and method of loci, participants imagine list-items alongside pre-memorized visual features (pegs or loci, respectively), a resemblance that Bower and Reitman (1972) noted. However, the peg-list method lacks any navigation metaphor. Bouffard et al. (2018), testing memory with free recall, found that peg strategies with little or no navigation component approached performance levels of the method of loci. The similar level of success between these strategies, thus, resonates with our findings, that environment-properties that should influence imagined navigation do not appear to exert a large influence on serial-recall. Interestingly, in a neuroimaging study, Fellner et al. (2016) reported brain-activity reminiscent of imagined navigation that was greater during method of loci than the peg-list method. However, they found no activity predictive of memory-outcome (subsequent-memory effect) that was unique to the method of loci. This converges with our findings, that the navigational demands of substrate environment have only small effects on serial-recall, suggesting the method of loci may be best understood as a special case of peg methods. This is also in line with Gross et al. (2013); in their meta-analysis of memory enhancement studies, with a focus on aging, they concluded that there is no special place for the method of loci among strategies.

Final factors worth considering are that our lists were relatively short; results might differ for much longer lists; and our participants were recruited blind to the experiment, and were thus not selected in any way related to their interest in memory ability. In contrast, memory enthusiasts who voluntarily adopt the method of loci may be a highly selected subgroup. For such people, it is still possible that imagined navigation leads to superior memory performance than strategies that omit navigation— a question that could be tested in a future experiment. Nonetheless, it is possible that extensive training (e.g., de Lange et al., 2017; Dresler et al., 2017; Maguire et al., 2003; Mallow et al., 2015) with the method of loci may use the strategy differently than our less experienced participants.

In sum, the method of loci, despite its popularity, may not rely critically on imagined navigation; rather, this strategy may be equivalent to a large set of peg-based strategies that do not incorporate navigation.

References

Bouffard, N., Stokes, J., Kramer, H. J., & Ekstrom, A. D. (2018). Temporal encoding strategies result in boosts to final free recall performance comparable to spatial ones. *Memory & Cognition*, *46*(1), 17-31.

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

Bower, G. H., & Reitman, J. S. (1972). Mnemonic elaboration in multilist learning. *Journal of Verbal Learning and Verbal Behavior*, *11*, 478-485.

Chan, E., Baumann, O., Bellgrove, M. A., & Mattingley, J. B. (2012). From objects to landmarks: the function of visual location information in spatial navigation. *Frontiers in psychology*, *3*, 304. doi: 10.3389/fpsyg.2012.00304

de Lange, A.-M. G., Bråthen, A. C. S., Rohani, D. A., Grydeland, H., Fjell, A. M., & Walhovd, K. B. (2017). The effects of memory training on behavioral and microstructural plasticity in young and older adults. *Human Brain Mapping*.

Dresler, M., Shirer, W. R., Konrad, B. N., Müller, N. C. J., Wagner, I. C., Fernández, G., … Greicius, M. D. (2017). Mnemonic training reshapes brain networks to support superior memory. *Neuron*, *93*(5), 1227-1235.

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), ENEURO.0184-16.2016.

Gross, A. L., Parisi, J. M., Spira, A. P., Kueider, A. M., Ko, J. Y., Saczynski, J. S., … Rebok, G. W. (2013). Memory training interventions for older adults: a meta-analysis. *Aging & Mental Health*, *16*(6), 722-734.

Heft, H. (1979). The role of environmental features in route-learning: two exploratory studies of way-finding. *Environmental Psychology and Nonverbal Behavior*, *3*(3), 172-185.

Herzog, M. H., Fahle, M., & Koch, C. (2001). Spatial aspects of object formation revealed by a new illusion, shine-through. *Vision Research*, *41*, 2325-2335.

Herzog, T. R., & Leverich, O. L. (2003). Searching for legibility. *Environment and Behavior*, *35*(4), 459-477.

Jansen-Osmann, P., & Fuchs, P. (2006). Wayfinding behavior and spatial knowledge of adults and children in a virtual environment: The role of landmarks. *Experimental Psychology*, *53*(3), 171-181. doi: 10.1027/1618-3169.53.3.171

Kim, Y. O., & Penn, A. (2004). Linking the spatial syntax of cognitive maps to the spatial syntax of the environment. *Environment and Behavior*, *36*(4), 483-504.

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

Lau, M. C., Goh, W. D., & Yap, M. J. (2018). An item-level analysis of lexical-semantic effects in free recall and recognition memory using the megastudy approach. *Quarterly Journal of Experimental Psychology*, *71*(10), 2207-2222.

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

Li, R., & Kippel, A. (2016). Wayfinding behaviors in complex buildings: The impact of environmental legibility and familiarity. *Environment and Behavior*, *48*(3), 482-510. doi: 10.1177/0013916514550243

Maguire, E. A., Valentine, E. R., Wilding, J. M., & Kapur, N. (2003). Routes to remembering: the brains behind superior memory. *NatureNeuroscience*, *6*(1), 90-95.

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

Montello, D. R. (1998). A new framework for understanding the acquisition of spatial knowledge in large-scale environments. In M. J. Egenhofer & R. G. Golledge (Eds.), *Spatial and temporal reasoning in geographic information systems* (p. 143-154). New York: Oxford University Press.

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.

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 & Function Journal*, *223*, 1379-1389.

Nyberg, L., Sandblom, J., Jones, S., Neely, A. S., Petersson, K. M., Ingvar, M., & Bäckman, L. (2003). Neural correlates of training-related memory improvement in adulthood and aging. *Proceedings of the National Academy of Sciences, USA*, *100*(23), 13728-13733.

O'Neill, M. J. (1992). Effects of familiarity and plan complexity on wayfinding in simulated buildings. *Journal of Environmental Psychology*, *12*(4), 319-327. doi: 10.1016/S0272-4944(05)80080-5

Raz, A., Packard, M. G., Alexander, G. M., Buhle, J. T., Zhu, H., Shan, Y., & Peterson, B. S. (2009). A slice of π: an exploratory neuroimaging study of digit encoding and retrieval in a superior memorist. *Neurocase*, *15*(5), 361-372.

Richardson, A. E., Powers, M. E., & Bousquet, L. G. (2011). Video game experience predicts virtual, but not real navigation performance. *Computers in Human Behavior*, *27*(1), 552-560. doi: 10.1016/j.chb.2010.10.003

Roediger, H. L. I. (1980). The effectiveness of four mnemonics on 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.

Rossano, M. J., & Moak, J. (1998). Spatial representations acquired from computer models: cognitive load, orientation specificity and the acquisition of survey knowledge. *British Journal of Psychology*, *89*, 481-497.

Shelton, A. L., & Gabrieli, J. D. E. (2002). Neural correlates of encoding space from route and survey perspectives. *Journal of Neuroscience*, *22*(7), 2711-2717.

Shelton, A. L., & McNamara, T. P. (2004). Orientation and perspective dependence in route and survey learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(1), 158-170.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). *A 21 word solution (October 14, 2012).* (Available at SSRN: `https://ssrn.com/abstract=2160588 or http://dx.doi.org/10.2139/ssrn.2160588`)

Smith, S. P., & Du'Mont, S. (2009). Measuring the effect of gaming experience on virtual environment navigation tasks. *IEEE Symposium on 3D User Interfaces.* doi: 10.1109/3DUI.2009.4811198

Thorndike, P. W., & Hayes-Roth, B. (1982). Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology*, *14*(4), 560-589. doi: 10.1016/0010-0285(82)90019-6

Wang, A. Y., & Thomas, M. H. (2000). Looking for long-term mnemonic effects on serial recall: the legacy of Simonides. *American Journal of Psychology*, *113*(3), 331-340.

West, G. L., Konishi, K., Diarra, M., Benady-Chorney, J., Drisdelle, B. L., Dahmani, L., ... Bohbot, V. D. (2018). Impact of video games on plasticity of the hippocampus. *Molecular Psychiatry*, *23*(7), 1566-1574.

Wilson, M. D. (1988). The MRC psycholinguistic database: Machine readable dictionary, version 2. *Behavior Research Methods, Instruments, & Computers*, *20*, 6-11.

Yates, F. A. (1966). *The art of memory.* London: Routledge & Kegan Paul.

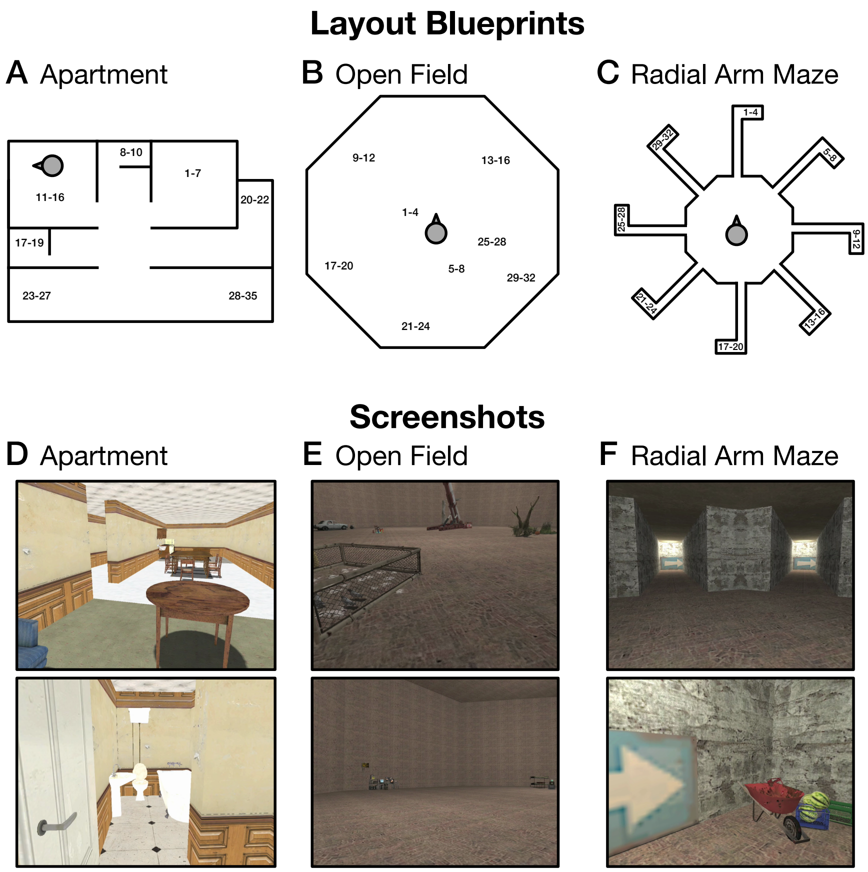

Figure 1. Layout blueprints (a–c) and sample screenshots (d–f) of the three environments. Grey circle with triangle marker denotes the starting position of the participant within the environment. Number ranges correspond to object clusters within the environment.

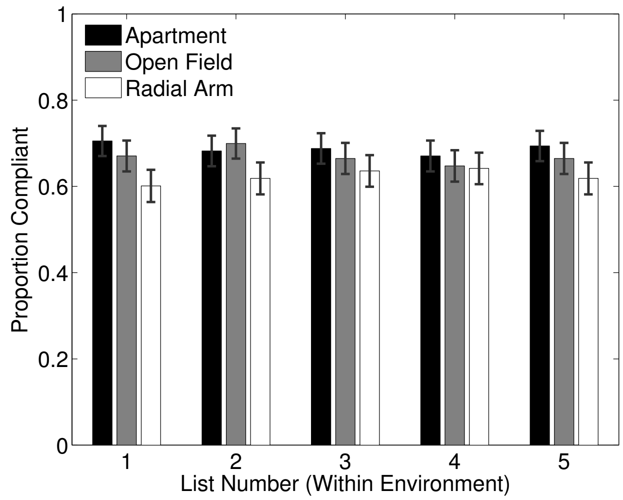

Figure 2. Compliance rates as a function of environment, averaged across all participants and all environments orders. Error bars plot standard error of the mean. The only significant post-hoc test was Apartment > Radial Arm.

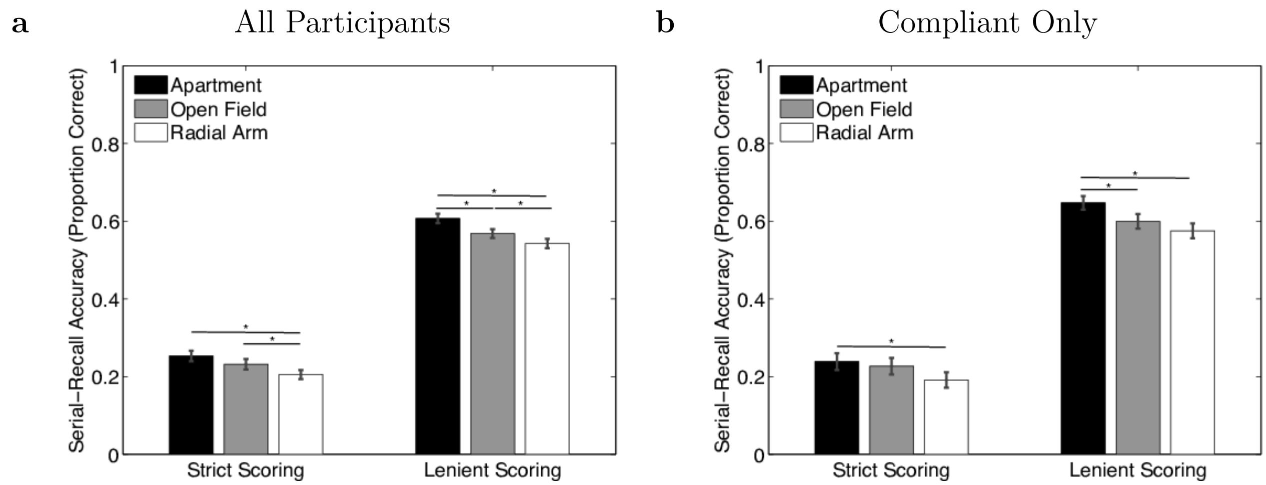

Figure 3. Serial-recall accuracy as a function of environment, averaged across all environments orders and all participants (a) or perfectly compliant participants only (b). Strict scoring— the word had recalled in the correct response-position. Lenient scoring— the word had to be from the current list, regardless of response order. Error bars plot standard error of the mean.

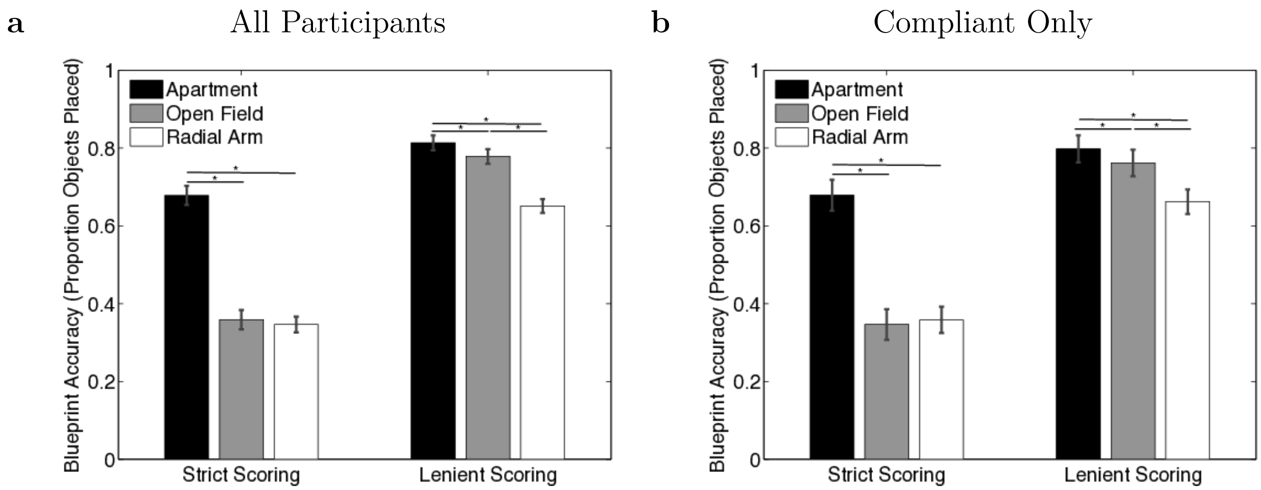

Figure 4. Blueprint accuracy, measured in proportion of objects correctly placed, as a function of environment, averaged across all environments orders and all participants (panel a) or perfectly compliant participants only (panel b). Strict scoring— the object had to be from the current environment and placed in the correct location. Lenient scoring— the object had to be from the current environment, regardless of where it was placed. Error bars plot standard error of the mean.

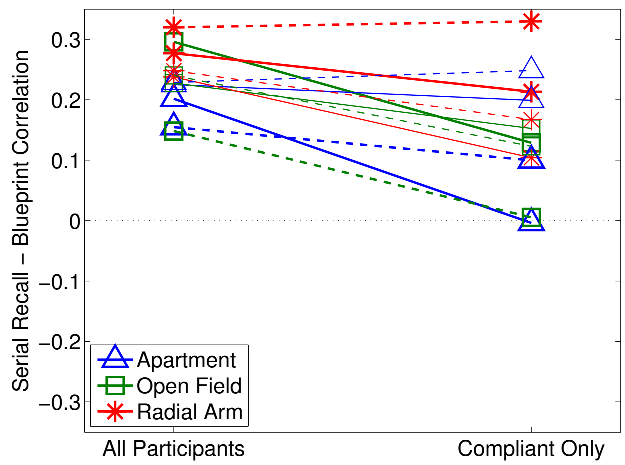

Figure 5. Correlation between serial-recall accuracy and blueprint-accuracy for a given environment. Correlations are computed over all participants (left-hand points) or over compliant-only (right-hand points). Point markers denote environment (triangle - Apartment, square - Open Field, star - Radial Arm). Thick lines denote calculations using strict scoring of serial recall and thin lines denote lenient scoring of serial recall. Solid lines denote strict scoring of blueprint recall and dashed lines denote lenient scoring of blueprint recall.

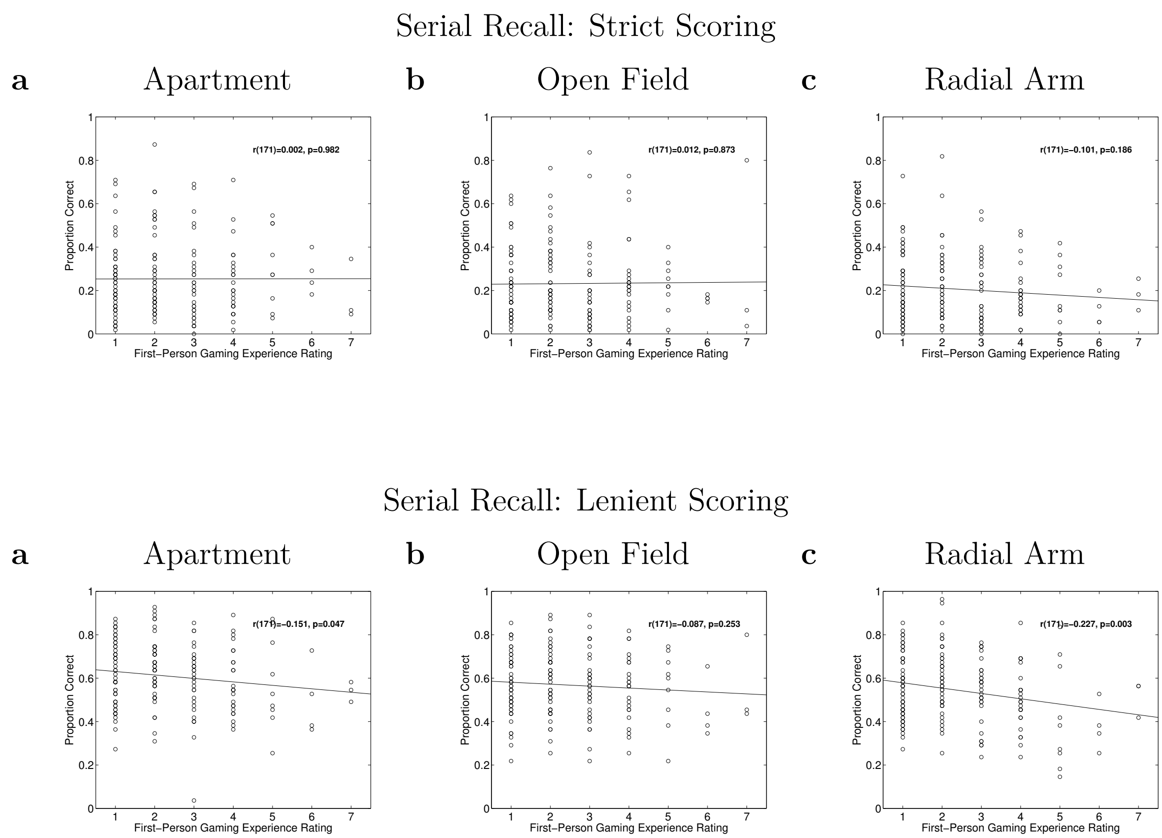

Figure 6. Scatter plots demonstrating the lack of positive correlation between prior experience with first-person video games and performance on serial recall. Each point represents a single participant.

Appendix

Word Pool

ALLEY	CANDLE	FIDDLE	LOCKER	PUPPY	STRIPE
ANCHOR	CANDY	FLAME	MANSION	PURSE	SULPHUR
ANKLE	CANNON	FLASK	MEADOW	RABBI	SUNBURN
APPLE	CANOE	FLOOD	MEASLES	RABBIT	SWAMP
ASPHALT	CARROT	FLUTE	MINER	RIBBON	SWORD
BAGPIPE	CEREAL	FROST	MONKEY	ROBIN	TENNIS
BALLOON	CHALK	FURNACE	MORGUE	RUBBER	THORN
BANDAGE	CHAPEL	GARLIC	MOUSE	SALAD	THREAD
BANNER	CHEEK	GEESE	MUCUS	SALOON	THUMB
BASKET	CHERRY	GIRDLE	MUSTARD	SATIN	TICKET
BEAST	CHISEL	GLACIER	NAPKIN	SAUCE	TIGER
BEAVER	CHOIR	GLOVE	NEEDLE	SCOTCH	TIMBER
BERRY	CIDER	GRAPE	NICKEL	SCOUT	TOAST
BIRCH	CIGAR	GRAVEL	NURSE	SEAWEED	TOOTH
BISCUIT	CLIFF	GRAVY	OATMEAL	SHAWL	TOWER
BLADE	CLOVER	GROCER	OLIVE	SHIELD	TRAILER
BLISTER	CLOWN	HAIRPIN	ONION	SHRIMP	TRIPOD
BLOSSOM	COFFIN	HAMMER	ORGAN	SHRUB	TROLLEY
BLOUSE	COLLAR	HEDGE	OTTER	SINGER	TROUT
BOSOM	CORAL	HELMET	PANTS	SKATE	TRUMPET
BOULDER	CORPSE	HOUND	PASTURE	SKULL	TULIP
BOUQUET	COTTAGE	HURDLE	PEACH	SLEEVE	TUNNEL
BRANDY	CRADLE	INFANT	PEARL	SLIPPER	TURTLE
BREAST	CRANE	INSECT	PEDAL	SNAIL	VAULT
BRICK	CREAM	JELLY	PEPPER	SPADE	VELVET
BRISTLE	CROWN	JEWEL	PIANIST	SPICE	WALLET
BRONZE	CRUISER	JOCKEY	PICKLE	SPIDER	WALNUT
BROOK	CRYPT	JUICE	PILLOW	SPIKE	WALRUS
BROOM	DENTIST	KENNEL	PLANK	SPINACH	WHEAT
BUBBLE	DITCH	KETTLE	PLATTER	SPONGE	WHISKEY
BUCKET	DOUGH	KITTEN	PLIERS	SPOON	WHISTLE
BUCKLE	DRIZZLE	KNIGHT	POSTER	STAIR	WILLOW
BUTCHER	DUMMY	LANTERN	PRIEST	STEAK	YACHT
BUTTON	DUNGEON	LEMON	PRUNE	STEAM	ZIPPER
CABBAGE	EAGLE	LINEN	PUDDLE	STEAMER	
CAMEL	ELBOW	LIVER	PUPIL	STEEPLE	
CANAL	FERRY	LOBSTER	PULPIT	STRING	