Aiding Semantic Memory Creation with Navigational Context

Thomas Benjamin Lyle Wasden
Brigham Young University

Follow this and additional works at: https://scholarsarchive.byu.edu/etd
Part of the Life Sciences Commons

BYU ScholarsArchive Citation
https://scholarsarchive.byu.edu/etd/9398

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.
Aiding Semantic Memory Creation with Navigational Context

Thomas Benjamin Lyle Wasden

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Steven Luke, Chair
Brock Kirwan
Jeffery G. Edwards

Neuroscience Center
Brigham Young University

Copyright © 2022 Thomas Benjamin Lyle Wasden
All Rights Reserved
ABSTRACT

Aiding Semantic Memory Creation with Navigational Context

Thomas Benjamin Lyle Wasden
Neuroscience Center, BYU
Master of Science

While we have traditionally understood the hippocampus to be involved in memory and navigation, it also appears that it has a role in language processing, creation and prediction. An obvious explanation for this is that language is impossible if linguistic signs cannot be remembered and retrieved. Because linguistic signs are definitionally biologically neutral or arbitrary, we must use the brain’s apparatus for learning and storing information from the external world to store and retrieve them. Although plausible, this explanation fails to take into account the hippocampus’ role in navigation as a contributing element in the processing, storage and retrieval of linguistic signs. Because the hippocampus also represents non-physical spaces through the same basic cognitive mechanisms with which it represents physical space, it is possible that the semantic content of linguistic signs is encoded in a fundamentally similar way to how navigational information is encoded. If true, this could have implications for education in general, and second language acquisition specifically. These experiments test whether there might be a learning benefit to presenting information in consistent spatial locations by having participants learn word associations in a 3-dimensional virtual environment. The experiments found that this was not the case. These findings have implications for education. Some educational paradigms stress learning in relevant contexts. These results suggest that physical location may not be an important component of a learning environment.

Keywords: navigation, memory, hippocampus, word association
# TABLE OF CONTENTS

- Aiding Semantic Memory Creation with Navigational Context ........................................... i
- ABSTRACT ................................................................................................................................ ii
- TABLE OF CONTENTS ........................................................................................................ iii
- TABLE OF FIGURES ........................................................................................................... iv
- Introduction .......................................................................................................................... 1
  - Context-Dependent Memory ............................................................................................... 1
  - Memory, Language, and Navigation ..................................................................................... 2
  - Semantic Vectors and Meaning ........................................................................................... 4
  - Implications .......................................................................................................................... 4
- Methods ................................................................................................................................. 6
  - Hypotheses ......................................................................................................................... 6
  - Participants .......................................................................................................................... 6
  - Materials ............................................................................................................................... 7
  - Experimental Design .......................................................................................................... 7
    - Experiment 1: .................................................................................................................... 7
    - Experiment 2: ................................................................................................................... 20
- Results .................................................................................................................................... 22
  - Experiment 1 ..................................................................................................................... 22
    - Experiment 1 Summary: .................................................................................................. 29
  - Experiment 2 ..................................................................................................................... 29
    - Experiment 2 Summary: .................................................................................................. 30
- Discussion ............................................................................................................................... 31
  - Experiment 1: ..................................................................................................................... 31
  - Experiment 2 ..................................................................................................................... 33
- Conclusion ............................................................................................................................... 34
- References .............................................................................................................................. 36
# TABLE OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>8</td>
</tr>
<tr>
<td>Figure 2</td>
<td>10</td>
</tr>
<tr>
<td>Figure 3</td>
<td>11</td>
</tr>
<tr>
<td>Figure 4</td>
<td>12</td>
</tr>
<tr>
<td>Figure 5</td>
<td>13</td>
</tr>
<tr>
<td>Figure 6</td>
<td>14</td>
</tr>
<tr>
<td>Figure 7</td>
<td>15</td>
</tr>
<tr>
<td>Figure 8</td>
<td>16</td>
</tr>
<tr>
<td>Figure 9</td>
<td>18</td>
</tr>
<tr>
<td>Figure 10</td>
<td>20</td>
</tr>
<tr>
<td>Figure 11:</td>
<td>23</td>
</tr>
<tr>
<td>Figure 12</td>
<td>24</td>
</tr>
<tr>
<td>Figure 13</td>
<td>25</td>
</tr>
<tr>
<td>Figure 14</td>
<td>26</td>
</tr>
<tr>
<td>Figure 15</td>
<td>27</td>
</tr>
<tr>
<td>Figure 16</td>
<td>28</td>
</tr>
</tbody>
</table>
Introduction

Previous work from a host of domains (which will be summarized below), from evolutionary biology to educational psychology, lends support to the idea that associating location information with linguistic or semantic information may aid encoding. Despite this, the specific question of whether this association aids encoding independently from other sensory stimuli in learning tasks has not been tested. Additionally, some research paradigms that lend support to this idea, like those behind context-dependent memory work, test a related, but essentially different question. The following sections present evidence that suggests that location associations with linguistic or semantic information may aid encoding, and also distinguish previous, related work from the current question.

Context-Dependent Memory

A large body of research has been accumulating for several decades showing that memory retrieval is somewhat dependent on context. The earliest work looked at whether information learned underwater was retrieved equally well above water as it was in the same context (i.e. underwater) that it was learned in (Godden & Baddeley, 1975). It turned out that, in this specific circumstance, consistent context aided retrieval. Since then, many partial replications have added weight to the main results of this study, showing that the presence of similar contextual clues during learning and retrieval improves the amount of information retrieved, perhaps because of increased consolidation. These replications have shown that the same pattern holds for virtually presented context clues (Shin et al., 2020), for sounds (Ostendorf et al. 2020), for smells (Hackländer & Bermeitinger, 2018), and even in different species (Maza et al., 2016). Replications have also focused on different participant groups and tasks, including university students (Parker & Gemino, 2001; Coveney, 2013; Sung et al., 2020), individuals in
stressful environments (Thompson et al., 2001), individuals learning motor programs (Schmidt et al., 2021), and children (Tippenhauer & Saylor, 2019).

Although the general trend in the research shows that context influences retrieval, context-dependent effects are not always seen. For example, two of the studies above (Parker & Gemino, 2001; Coveney, 2013) failed to show a relationship between place-based stimuli and improved retrieval outcomes. Also, Smith (1984) showed that certain techniques designed to stimulate deep processing with place-based stimuli did not improve retrieval of learned information. In these studies though, it is not clear what is meant by using places as context: did individuals identify their location by visual contextual features, and if so, what were those features? Was there any reason to assume that they were salient? Answers to these questions tend to be left out. Additionally, several other studies have shown that there may be other factors that mediate or moderate the effect of context-dependent memory retrieval, including the level of cognitive load involved in the stimuli or context (Albus et al., 2021; Martínez, 2021; Sung et al., 2020).

Memory, Language, and Navigation

Among the many types of contexts that may be useful in improving retrieval, place-based, spatial or navigational contexts have historically enjoyed a long association with practical memory techniques and mnemonics (Caruthers, 1993). One popular application has been termed the “method of loci”, in which individuals imagine salient stimuli visually (and sometimes with other senses as well) in certain locations (i.e. loci) within a maze, path or other three dimensional environment. Testing the efficacy of the method of loci and related effects, however, does not tell us about the role of navigation in semantic memory because the effect of navigation is confounded with the effect of salient stimuli. Context-dependent memory experiments show that
salient stimuli of various types can improve memory retrieval (Godden et al., 1975; Hackländer et al., 2018; Ostendorf et al., 2020), and studies that purport to show the effects of navigation or place-association on memory make either make little attempt to control for these stimuli, or else actively add them in (Holden & Sykes, 2011; Shin et al., 2020).

Additionally, there has been a large amount of recent work showing that the hippocampus, the structure known to be involved in spatial navigation, has an important role in comparison of language related information. For example, Solomon et al. (2021) show that the hippocampus is involved in processing the relatedness between words. In a conceptually related study, Brown-Schmidt et al. (2021) contested the idea that the hippocampus is required for assessing the semantic relatedness of closely related ideas whose relationships are already learned, suggesting that the hippocampus has a role in forming semantic connections rather than retrieving them. Supporting this, Pu et al. (2020) show that theta oscillations in the hippocampus code the semantic relationship between words. Additionally, temporal lobe damage, and hippocampal damage in particular, have been shown to be related to deficits in language (Race et al., 2021; Whitten et al., 2021), including predictive abilities in language that are thought to be based on an individual’s ability to determine semantic connections between words (Covington & Duff, 2016). Similarly, Kepinska et al. (2018) showed that acquisition of novel grammatical knowledge is hippocampus dependent.

This relationship between the hippocampus, language and its well known function in navigation has led some to generalize about the hippocampus’ function, turning it into a cognitive map of meanings from any domain rather than interpreting it as solely a literal cognitive map (Solomon et al., 2019; Tavares et al., 2015). Although the human cognitive map appears to retain its simpler map-like function (Gruenewald, 2015), it may be that this was
merely a base for more complex meaning associates to evolve (Corballis, 2019). In fact, homologous structures in some of the most distantly related bilateral animals support the idea that this association between memory and spatial processing and navigation is very well conserved (Maza et al., 2016). Hupbach et al. (2008) were even able to show that an individual’s perception of being in a certain place aided the updating of episodic memories created at that place, implying that this relationship between navigational memory and other forms of memory may be fundamental.

**Semantic Vectors and Meaning**

The wide range of functions performed by the hippocampus, including navigation functions, memory functions, and language functions, raises an obvious question: could all of these functions actually be underlyingly the same, or at least very similar? Solomon et al. (2019) expressed that the underlying nature of semantic information could be vector based, implying that there is a cognitive map-like structure for word meanings that is analogous to the cognitive map for locations. In this paradigm, each word’s meaning can be represented conceptually by an ordered list of numbers (i.e. a vector) representing its relatedness to other words, concepts or contexts. Vector based representations of word meaning have been used in natural language processing for quite a long time to improve information retrieval (Tang et al., 2020) and natural language understanding (Jung, 2019), for example, using this basic architecture. The underlying biological system that would be analogous to this model would represent a word by a network’s functional connectivity with other networks which represent related concepts. This model and its implications form the primary subject for this thesis’ research questions.

**Implications**
The human cognitive map seems to encode the three dimensions of physical space and time (Gruenewald, 2003). If the underlying representation of semantic information is the same as the representation of physical location, we would expect these “vectors”, or overlapping networks to be equally able to incorporate physical dimensions into semantic meaning as they are able to incorporate other context-based stimuli or semantic information. Additionally, the association of navigational information to wholly new words might provide enough connectivity between newly forming networks and context to “jumpstart” the creation of a new entity (i.e. a new word vector). This would provide the nascent semantic vector with a convenient amount of information that would act as a handle to allow further manipulations that may increase the words connectivity to semantically relevant (i.e. non-navigational) information. This last implication is the subject of this thesis.

If validated, these implications could have a meaningful impact on teaching and learning. Many computer applications use virtual environments, and others could be designed that take advantage of these phenomena to improve learning. As noted by Lu (2013), inclusion of participatory environments, like games or mazes, has a real potential to improve teaching practices. Although some work has been done to try to demonstrate that place-based learning can be helpful in the field of language-learning, for example, by Holden & Sykes (2011), these studies do not limit context to navigational stimuli. Often their concern is about creating complete curricula or courses that leverage multiple known or putatively useful phenomena to improve learning. In these cases, it is not possible to establish that any one of the useful aspects of the course architecture is responsible for any observed improvement in learning. For example, Holden & Sykes (2011) spend a considerable amount of time trying to ensure that the dialogue used in their mobile application is realistic and appealing, and that the stories involved are
interesting. Because of this, there is a need for a study that appropriately controls for factors that are not strictly navigation related. In testing if navigational information is itself a salient enough feature to promote memory creation, stimuli presented to participants should be limited, wherever possible, to stimuli that are involved in determining position in space.

Methods

Hypotheses

1. Individuals perform better at tasks requiring them to learn new words when given consistent physical contexts that the participants must navigate for those words than when the words are presented unassociated with a navigation-requiring context. This hypothesis is based on the assumption that semantic information can be associated with navigational information in a way that aids retrieval and manipulation.

2. Performance on a quiz consisting of 3 dimensional mental rotations will be correlated to increases in the effect of navigational stimuli in aiding word learning. This will be the case because spatial reasoning skills will be required for participants to construct the spatial context that will aid encoding. This additional hypothesis is useful to provide additional evidence that the observed differences in learning across different environments are the effect of participants performing navigation rather than variations in the amount of time it takes to perform the tasks, or other variables that are difficult to control for.

Participants

Participants in this experiment were college students at least 18 years old. They were individuals with no diagnosed language deficits. Based on the effect size (Cohen’s d of 0.917) seen in Shin et al. (2020), a minimum of 32 participants were needed for each experiment. For
the first experiment, 43 participated. In the second, 37 participated. The participants were close
to half male and female, although there were more females who participated than males. In
experiment 1, 26 females and 17 males participated. In experiment 2, 22 females and 15 males
participated.

There were no exclusions based off of native language. Basic directions provided in
English were used in the administration of the experiments, and nonce words in the Latin
alphabet were used in the experiment. It was assumed that if participants could understand the
recruitment material in English they would be competent enough in English to handle the
directions and nonce words.

Materials

Both experiments were conducted entirely with the use of software built with the Unity
game engine and using the C# programming language. Fractal images were used in the tasks and
testing. Fractal images were obtained from a modified version of an online free fractal creator
found at https://sciencevsmagic.net/fractal/.

27 nonce words were used for the word association tasks. These words were created to be
simple, and were all one syllable in length, with a single onset consonant, one vowel, and one
coda consonant.

Tasks were presented with a wireless keyboard and mouse, and on a 1920 by 1080
resolution screen connected to a laptop by HDMI connection.

Experimental Design

Experiment 1:

Experiment 1 consisted of an object rotation task, three word association tasks, three reading
span tests, and three tests of the word associations learned in the word association tasks. Below
are descriptions of each of the tasks. Following that is a description of the order in which the tasks were presented to participants.

**Object Rotation Task Details.** The object rotation task presented participants with pictures of pairs of objects, and asked them to determine as quickly and accurately as possible if the pictures could be of the same object, or if they are of different objects. To complete the task, individuals have to mentally rotate the objects to determine whether the pictures differ only in the angle they are taken at. Participants were presented with two trial pairs before the test began in order to get acquainted with the task. They were given a maximum of 30 seconds to decide on each pair of objects. Figure 1 shows what this task looks like.

**Figure 1**

*Object Rotation Task*
Note. The image above shows the object rotation task as a participant would see it. The object pair is marked “trial item,” meaning that the response to this item would not be graded and that this pair is presented for participants to practice the task with.

**Word Association Task Details.** Three word association tasks were presented. They will be labeled for the sake of convenience A, B, and C. Each task was administered after a video tutorial on how to perform the task and a trial version of the task in which the fractal images used in the tasks were replaced with images of numbers.

Task A (shown in figure 2) was a word association task that required navigation in a virtual environment. In this task, participants would be presented with a prompt word. In a directory that was available to participants through the duration of the task, participants could find an object associated with the prompt word. Participants would then search the virtual environment for the object associated with the prompt. On finding the object, participants would be given a new prompt. There were 24 prompts to follow in total. Within the virtual
environment, there were nine objects. Participants were prompted to find three of the objects 5 times, three of the objects 2 times, and three of the objects only once. The purpose of this arrangement was to make some of the associations easier to remember than others in order to avoid floor and ceiling effects during testing. The objects to be found were cubes, each with a unique fractal image repeated on each of its faces.

**Figure 2**

*Task A*

![Image of task A](image)

*Note.* The image above is an image of task A as a participant might see it. The participant navigates the virtual environment. The prompt word is at the top of the screen. The number of matches completed and the number of matches total are recorded to the right.

Task B (shown in figure 3) was a word association task that consisted of clicking pictures of objects like those used in task A. In this task, participants were presented with a prompt word. In a directory that was available to participants through the duration of the task, participants
could find an object associated with the prompt word. Participants would then search an array of images for an image of the object. On finding the object and clicking it, participants would be given a new prompt word. There were 24 prompts to follow in total. The images were of cubes with fractal designs on them, but the fractal designs used in this task were not used in any other task. The images were of the objects in a context like the virtual environment in task A, except that the contexts did not form a coherent environment (i.e. they could not be pieced together to form a map of some possible virtual environment).

**Figure 3**

*Task B*

*Note.* The image above is an image of task B as it might be seen by a participant. The prompt word is at the top. Participants click the image with the object associated with the prompt word in the directory. The number of matches completed and the total to c are recorded to the right.
In task C (shown in figure 4), participants were presented with a prompt word. In a directory that was available to participants through the duration of the task, participants could find an object associated with the prompt word. Participants would then search an array of images for an image of the object. On finding the object and clicking it, participants would be given a new prompt word. There were 24 prompts to follow in total. The images just fractal designs without any other context given to them.

**Figure 4**

*Task C*

[Image of task C with 24 prompts and a prompt word at the top]

*Note.* The image above shows task C as it might be viewed by a participant. The prompt word is at the top. Participants click the image associated with the word in the directory. The number of tasks completed and the number of tasks to complete are noted to the right.
**Reading Span Test Details.** Participants were administered reading span tests (shown in figures 5, 6, and 7) with different content but identical presentation three separate times. During the reading span test, participants were presented with sets of sentences. After each sentence, participants decided whether the sentence made sense or not. Half of the sentences presented had semantic incongruities. After making this judgment, participants were presented with a word they were told to remember. Following the presentation of several sentences and words like this, participants were asked to reproduce a list of each of the words they were asked to remember. This was repeated several times, with 2 sentences the first and second time, the 4 sentences, then 6 sentences, and finally 8 sentences. Participants were told that the first set of sentences in each test would be practice. The function of the reading span test was primarily to provide a cognitively demanding task involving semantic judgments to provide an opportunity for some forgetting of the word associations learned before it.

**Figure 5**

*Reading Span Sentence Screen*
Does the sentence make sense?

She let out a long and painful cough as soon as she walked outside.

Press the right arrow for yes, and the left arrow for no.

*Note.* The image above is an image of a sentence in the reading span task. Participants respond based on whether or not the sentence makes sense to them by pressing the right or left arrow keys.

**Figure 6**

*Reading Span Word Screen*
Note. The image above is of the screen shown to participants during the reading span task after responding whether a sentence makes sense or not. Participants are asked to remember the word shown in the center of the screen.

**Figure 7**

*Reading Span Input Screen*
Note. The image above is of the screen shown to participants after they view multiple pairs of sentences and words to remember. Participants enter the words they were asked to remember in the input field, before being presented with a new set of sentence and word pairs.

**Word Association Test Details.** During the word association tests (shown in figure 8), participants were presented with a fractal design and asked which word it corresponded to in the task in which they then learned it.

**Figure 8**

*Word Association Test*
Note. The image above is an image of one of the items on the word association test. Participants are asked to reproduce the word associated with the image shown during the word association task preceding the test in the input field. Participants press “ENTER.”
**Break Details.** Participants were offered 5 minute breaks between word association tests and the following tasks. They were allowed to skip these breaks if they wanted to. The directions for the next task started immediately after 5 minutes ended if participants chose to take the full 5 minutes.

**Presentation Order.** Participants were given a consent form to sign on walking in to participate in the study. After reading and signing the form, video directions explained the general order of events during the experiment. The first task presented was the object rotation task, which was preceded by a video explaining how to perform it. Following the object rotation task, tasks A, B and C were presented in a random order, each preceded by a video tutorial and a trial version of the task. After each of these tasks a reading span test was presented. Following the reading span test participants were tested on the word association learned in the previous task. Between the word association test and the following task there were optional 5 minute breaks. Presentation order for experiment 1 is outlined in the flow chart in figure 9.

**Figure 9**

*Experiment 1 Task Presentation Order*
Note. Flowchart for the method in experiment 1. Tasks A, B, and C are administered in a random order with identical tests following each of the tasks. The only difference in tests is that the associations in the word association test are those from the most recent learning task (i.e. A, B, or C).
**Experiment 2:**

Experiment 2 used the same tasks as experiment 1 with one exception. That task, described below, will be called Task D. It was very similar to task A.

Task D was different from task A in only one way. After an objective had been reached in task D, all object locations were randomly switched. This meant that objects were not paired with consistent locations, even though they were found within a virtual environment.

**Presentation Order.**Tasks in experiment 2 were presented in the same order as in experiment 1, except that only tasks A and D were presented. Tasks B and C were not present. Additionally, there were only 2 reading span assessments and 2 word association tests because there were only two word association tasks, and there was only one break offered (between the two tasks). Presentation order for experiment 2 is outlined in figure 10.

**Figure 10**

*Experiment 2 Task Presentation Order*
Note. Flowchart for the method in experiment 2. Tasks A, and D are administered in a random order with identical tests following each of the tasks. The only difference in tests is that associations in the word association test are those from the most recent learning task (i.e. A or D).
Results

Python was used to collect user data from output files and the R statistical language was used to perform statistical comparisons.

Experiment 1

The first tests performed looked for a significant difference in learning in the three tasks, A, B and C. This will provide evidence for or against hypothesis 1. Information recalled from task A in a word association test was recalled on average significantly more poorly than information learned in the other tasks. The p-value from an ANOVA test looking for differences in the mean between the tests for tasks A, B, and C was 0.02. The lowest score, as reported above, was the score on the test after task A, showing a difference between learning during the tasks, but not the difference expected.

Next, it was important to look into hypothesis 2. Do individuals who perform object rotation better learn in physical contexts better? To do this, a correlation between object rotation and learning during task A was looked for. The correlation between task A information recall and object rotation performance was -0.02 with a p-value of 0.6. This provides no evidence for a relationship between object rotation ability and learning in a physical context.

The data for the various tasks were distributed in fairly different ways. The average score for the object rotation task was 6.0 correct answers with a standard deviation of 1.6 correct answers. The highest score possible was 9 correct answers. The average score for associations learned during task A was 2.2 points with a standard deviation of 2.2 and out of 9 points possible. The average score for associations learned during task B was 3.6 points with a standard deviation of 2.9 and out of 9 points possible. The average score for associations learned during task C was 3.4 points with a standard deviation of 2.4 and out of 9 points possible. The p-value
for a t-test looking for differences in distributions between the scores for tasks B and C was 0.8. figure 11 illustrates the differences in learning across the three tasks.

**Figure 11:**

*Test Scores*

![Box plot showing test scores for tasks A, B, and C.](image)

*Note.* The figure above shows the means for participant performance with the words learned in different tasks. There was significant difference between the groups. The average for task A was lower than the averages for task B and task C. Tasks B and C did not differ from each other significantly, indicating that the tasks were basically equivalent.

There was a significant correlation between learning in task A and the order in which the tasks were presented. Tasks were given an index (0 for first, 1 for second, and 2 for third) that identified the order of the tasks. Index distributions are given in figure 12. Task A had an average index of 0.8, meaning that it was most often presented early on. Task B had an average
index of 1.2, meaning that it was often presented late. Task C had an average index of 1.0, meaning that it was fairly well distributed in all three possible positions. Task A was correlated with its index across participants with a correlation coefficient of 0.3 and a p-value of 0.04. Task B was correlated with its index across participants with a correlation coefficient of -0.03 and a p-value of 0.9. Task A was correlated with its index across participants with a correlation coefficient of 0.04 and a p-value of 0.8. The p-value for an ANOVA test against the null hypothesis that there is no difference in average index across tasks was 0.08, indicating that the distribution could reasonably have been obtained by chance and was not likely a result of a programming error during randomization.

**Figure 12**

*Task Index*

![Diagram showing the distribution of indices for each task.](image)

*Note.* The image above shows the distribution of indices (i.e. order) for each task. The index indicates the order a task was in relative to other tasks. For example, an index of 0 indicates that
a task was first, an index of 1 indicates that a task was second, and an index of 2 indicates that a task was completed last. An average index of 1 was expected for each task since the order of tasks was randomized for each experiment. Although there was some random variation, the distribution could have reasonably resulted from chance variation.

Performance on word associations associated with task A was found to be non-unimodal using Hartigan’s dip test with a p-value < .001. Scores are shown in figure 13. Performance on word association tests after tasks B and C were found not to be unimodal using Hartigan’s dip test with p-values of 0.03 and 0.02 respectively. Scores for task B are show in figure 14, and scores for task C are in figure 15.

Figure 13

Task A Word Association Score
Note. The image above shows the distribution of scores in the word association test after task A. The distribution was shown not to be unimodal using Hartigan’s dip test, indicating that there may be an unaccounted-for variable that accounts for the higher scores of some individuals.

Figure 14

Task B Word Association Score
Note. The image above shows the distribution of scores in the word association test after task B.

The distribution was shown not to be unimodal using Hartigan’s dip test, indicating that there may be an unaccounted-for variable that accounts for the higher scores of some individuals.

Figure 15

Task C Word Association Score
Note. The image above shows the distribution of scores in the word association test after task C. The distribution was shown not to be unimodal using Hartigan’s dip test, indicating that there may be an unaccounted-for variable that accounts for the higher scores of some individuals.

To better understand the multimodality in the data, an “advantage” score was calculated. The distribution for this variable is shown in figure 16. This score was the score for task A words minus the scores for task B and C words. The implications of this score and its modality will be discussed below. The mean advantage was -4.9 points, with a standard deviation of 4.4 points. The advantage distribution was tested for unimodality, and the p-value against the null hypothesis of the data being unimodal was 0.7.

Figure 16

Advantage Score
Note. The image above shows the distribution of “advantage” scores. The advantage is the learning advantage of task A over tasks B and C. It is calculated as the task A score minus the task B and task C scores. The Hartigan’s dip test p-value for these data.

**Experiment 1 Summary:**

Learning in task A was significantly worse than learning in the other conditions. No variables other than the order in which the tasks were performed seemed to influence performance. Interestingly, task order only affected learning in task A.

**Experiment 2**

The distributions of the scores on tests of learning for the two tasks were very similar. The average score on the word association test associated with task D was 1.8 with a standard deviation of 2.0. The p-value for the two-tailed t-test against the null hypothesis that the two
distributions have the same mean was 0.9. This indicated that there were no significant
differences in the mean scores on word association tests after each task.

The one distribution characteristic that was different for the two association tests was the
significance of modality in the distributions. The p-value for Hartigan’s dip test against the null
hypothesis that task A word association recall was unimodal was < .001. The p-value for
Hartigan’s dip test against the null hypothesis that task D word association recall was unimodal
was 0.0077. The mean score on the object rotation task was 6 points, with a standard deviation of
1.4. The significance of these differences is not clear, but potential explanations will be
discussed below.

Finally, it was necessary to determine if object rotation ability predicted task performance
in any way by calculating correlations between object rotation performance and learning in tasks
A and D. The correlation coefficient between the object rotation task score and task A learning
was 0.3 with a p-value of 0.2. The correlation coefficient between object rotation task score and
task D learning was 0.6 with a p-value of 0.08. The correlation coefficient between the object
rotation task time and task A learning was 0.1 with a p-value of 0.5. The correlation coefficient
between object rotation task score and task D learning was -0.2 with a p-value of 0.2.

**Experiment 2 Summary:**

Learning in this experiment was not significantly different by task. Other variables
measured did not account for performance on either task. Learning in task A was much less
unimodal than in task D, but the significance of this is unclear.
Discussion

Experiment 1:

Experiment 1 does not lend support to either hypothesis. Rather than seeing an increase in performance in word learning in Task A, which placed words in virtual locations, there was a significant decrease in performance. Additionally, object rotation score did not predict learning in task A. There are, in addition to this, three interesting take-aways from the data to discuss in the following paragraphs. The first interesting point is that the data for learning in each task (but not in the advantage score) appear not to be unimodal. The second interesting point is that there seems to have been no significant difference in learning in tasks B and C. The third interesting point is that there was only a significant correlation between the index of the task and performance of that task for task A.

One question worth asking is why performance on the word association test after task A would be lower than performance on the other word association tests. It is possible that tasks B and C lead to superior learning because the visual search required to perform them is helpful in memorizing patterns. Task A may have required less visual search because object locations could be memorized rather than images. Data from experiment 1 do not rule these possibilities out.

Why were the data not unimodal? The modality of the data for tasks A, B, and C suggests that there were at least two groups of individuals who approached the tasks in significantly different ways. The p-value on modality was lowest for task A, which makes sense given that task A was significantly different, and possibly more complex, than tasks B and C. It could be that previous experience with video games or other experience navigating virtual environments contributed to the distribution of performance on these tasks. Because data on video game usage was not collected, this cannot be verified. This would also explain the decreased significance of
the multimodality effect in scores of learning on tasks B and C. These tasks more closely approximated the sorts of tasks that would be common outside of a video game or virtual environment, and so these data would be less likely to be multimodal. It is also possible that some groups of people tend to navigate differently, and that this experiment did not collect enough data (or the right type of data) to predict who these individuals are. Additionally, there could be differences in the amount of effort participants put into the tasks. Any of these explanations would also explain why the multimodality disappeared in the advantage score. Individuals who contribute to the higher or lower mode consistently in all tasks contribute to unimodality in the advantage score. Individuals who contribute to different modes across the tasks would be expected to contribute to multimodality in the advantage score distribution. Because the advantage score distribution is unimodal, it appears that participants tended to stay in the same mode (whether higher or lower) for all tasks.

Next, it is interesting to note that there was no significant difference in learning between tasks B and C. These tasks were designed to give more external visual context in task B, and possibly aid performance in that way. The external context was designed to be similar to that in task A, which itself was intentionally minimal. Because of this, it appears that the presence of two controls was unnecessary.

The final point to discuss is that the order of the tasks was only correlated with performance for task A. In other words, participants who performed task A first probably performed significantly worse than they would have had they performed task A last, but that relationship did not exist for tasks B or C. This tendency should be taken into account in any future experiments using virtual navigation. It also complicates data analysis because it is harder to account for the effect of the order of the tasks when the effect is only significant for a minority
of the tasks. Additionally, it is worth wondering why this is the case. The most obvious explanation is that the cognitive load involved in learning how to navigate the virtual environment combined with the cognitive load involved in learning how the word association tasks work more generally impeded learning. Apparently, the tutorial and trial navigation task did not sufficiently address this problem. Albus, P., et al. (2021) found that cognitive load in virtual environments could significantly impair learning outcomes. Because of these concerns, and the fact that one of the control conditions proved superfluous, experiment 2 was designed to balance the cognitive load over two tasks: one experiment task and one control task.

**Experiment 2**

Experiment 2 failed to support either hypothesis as well. P-values for the difference in mean between the two learning conditions were not close to the level of significance. This is interesting because it makes it unlikely that the observed decreased learning during the navigation condition in experiment 1 would be due to participants having to perform more visual search in the control conditions, since increased visual search would be expected in task D as well. Neither navigation nor conditions requiring increased visual search seem to have any advantage over each other.

There was a large difference in the degree of significance in the Hartigan’s dip test of the distribution of scores for task A and D learning. This implies that there were two groups of participants, and their performance is more clearly different in the condition where object locations are consistent. What separates the groups is not clear, and may be effort, navigation ability, navigation or visual search strategy, video game familiarity, or some other variable. Additionally, this difference cannot be explained away in this experiment as an effect of unbalanced learning conditions as it was in the last experiment, since conditions were very
closely matched in experiment 2. If this difference is replicable, future work could investigate what is responsible for differences in distribution shape in these two conditions.

**Conclusion**

There is a long history, as mentioned before, of investigating how people learn in navigable environments. The idea that place learning may not be so different from word learning has also been around for a while (Pezdek, K., & Evans, G. W., 1979). It appears, however, that people do not simply incorporate physical dimensions into semantic schemas. This may be the case because removal of physical dimensions may be necessary in order to create an abstract representation of an object. Do individuals retain physical dimensions in their semantic representations for some things, like buildings, or objects that tend to be located in one place permanently? This could be determined with future work. The current study only looked at people’s representations of abstract shapes, which would normally be expected to occur in one place as easily as in any other place. Perhaps participants made a semantic distinction about the type of objects they were dealing with, and assumed based on that that location was not a relevant feature of the object.

This experiment could be improved in a couple of key ways. First, there were clear floor effects in the data. This could be ameliorated by making the learning tasks more effective. One way to do this would be to offer more learning trials for some of the fractals. For example, some of the prompts could have been given 10 times, rather than 5. Ceiling effects could be avoided by providing the other prompts the same number of times as in the current experiments. Another experiment that might effectively test the same thing as this one and avoid the floor effects by associating known words with pictures, or more nameable images with nonce words. This might be easier for participants. A pilot version of this experiment used nameable images, but ceiling
effects were seen so that method was abandoned. There may be an intermediate level of imageability that would work well to avoid both floor and ceiling effects. Another possible paradigm would dispense with the word directory by placing words in the virtual environment in addition to images. This may make the tasks simpler, and lead to better effort or increased learning.

Additionally, it is worth noting that most of the participants who commented on the tasks commented that the navigation tasks were very enjoyable. This is strange, primarily because all potentially interesting stimuli were intentionally kept out of the navigation tasks. Additional work could look at an innate drive to navigate and explore. This drive obviously exists in animals, since exploratory behavior is assumed to be the norm in many animal behavior tests. It would be interesting to determine what components of an environment are necessary to provide motivation for humans to explore it. It might be that the inherent human interest in exploration explains the efficacy of spatial mnemonics better than a vector semantic theory of semantic memory.
References


doi:10.1016/j.neuropsychologia.2020.107730


doi:10.1016/j.tics.2016.10.006

doi:10.1007/springerreference_302042


