

DEVELOPING A MEMORY REPRESENTATION: DO WE VISUALIZE OR DO WE
'VERBALIZE' OBJECTS AS WE EXPLORE THEM?

BY

CARLY PENROSE

A thesis submitted to
Department of Psychology
Mount Allison University
In partial fulfillment of the requirements for the
Bachelor of Arts degree with Honours
April 18, 2019

Abstract

Participants typically perform better when testing conditions match learning conditions, a phenomenon labeled encoding specificity. Interestingly, recent findings in visuo-haptic object identification violate this principle: participants who learn to recognize objects haptically perform just as well when asked to identify objects by sight as by touch. One possible explanation is that participants who explore objects haptically visualize the objects they explore, creating a multisensory memory trace equally accessible to vision and touch. I evaluated this possibility by asking healthy undergraduate participants to learn to recognize novel objects either by sight or by touch. Participants completed sequences of learning trials where they explore each object, and test trials where they recall the name of each object. During learning trials, some participants were presented with a visual distractor (half of participants viewed letters and the other half viewed nonverbal characters) they had to recognize later, while other participants completed a distractor-less control condition. Consistent with past findings, my results violated encoding specificity for participants who learned to recognize objects haptically. Interestingly however, only the verbal distractors interfered with learning. These results suggest that the creation of memory representation for novel objects is mediated by a verbal code rather than through visualization.

Acknowledgments

So many people have played important roles in the completion of this project, from my family in Halifax, to my family here in Sackville. Nothing I write here could do justice to how wonderful these people are, but I will do my best!

First and foremost, my incredible supervisor, Dr. Genevieve Desmarais. The high expectations you have for your students motivated me to expect the same from myself and has helped me accomplish so much more than I thought possible when I started this year. I can honestly say I have never learned so much from an educator in my life, and I am so glad I had the opportunity to work with you. Your teaching will continue to influence my life long after I've left Mount A, and I want to sincerely thank you for that. To my second reader, the wonderful, warm and hard-working Dr. Louise Wasylkiw. Your dedication to helping students is inspiring and I appreciate the help and encouragement you have offered me throughout the writing process. And to my research assistant, Sarah Park, your help in the lab has been so important for the completion of this project. You have all helped me significantly with the content of this project, and I cannot thank you enough for the hours you all have put in for my benefit.

I would like to acknowledge the world's best roommates for their countless offers to help and for being important emotional supports all year. Finally, I could not have had such a rewarding year without my personal cheerleader, Morgan Kelter. You have been so invested in my project from the start, offering support and helping me get other people interested in the work I was doing. You've told me before that you "just want to see me succeed", well you have, but only because you were there with me every step of the way.

Table of Contents

Method	8
Participants	8
Materials	8
Procedure	12
Results	16
Distraction Task Performance	16
Blocks to Criterion	16
Number of Errors	20
Discussion	21
Verbalization Processes at Encoding	22
Visualization	25
Active Rehearsal Versus Passive Interference	27
Conclusion	29
References	30

List of Figures

<i>Figure 1.</i> Apparatus used	10
<i>Figure 2.</i> Yi symbols used as non-verbal distractors	11
<i>Figure 3a.</i> Sequence of events of learning trials in the control condition (no distractors)	15
<i>Figure 3b.</i> Sequence of events of learning trials with a visualization distractor.....	15
<i>Figure 4.</i> Mean number of blocks to criterion in each distraction condition.....	18
<i>Figure 5.</i> Mean number of blocks to criterion by learning and identification modality	19

Developing a memory representation: Do we visualize or do we verbalize objects as we explore them?

Think about a time when you have walked through a dark room. With the inability to see, you may have resorted to feeling the walls for familiar landmarks to help you identify where you are. From this example, it is evident that one can recognize the same objects using either touch or vision to help orient themselves and learn about the world. This ability to recognize objects across senses is known as cross-modal memory. The processes that enable cross-modal memory have been of interest to researchers, and research on cross-modal learning and identification of objects has demonstrated that while cross-modal recognition is possible, an object learned in a certain context (such as a sensory modality) is more likely to be accurately remembered or recognized when it is retrieved in the same modality or context. This is the cognitive principle known as encoding specificity (Tulving & Thompson, 1973). Recent research indicates that an encoding specificity advantage may apply to retrieval accuracy for visual information, but this advantage may not be as robust for objects that are learned by touch, (i.e., haptically) (Desmarais, Wells, Meade, & Nadeau, 2017; Newell, Woods, Mernagh & Bühlhoff, 2005). One possible explanation for this may be that participants create a visual representation of objects in their memory by visualizing objects during haptic encoding and exploration of objects. Therefore, I aimed to investigate how interfering with visualization processes impacted participants' ability to recognize novel objects.

A simplistic view of sensory perception may assume that each sense leads to differential processing and memory formation for an object, however, this is an oversimplification and does not consider the sensory integration that is necessary to

accurately recognize and perceive sensory information. A significant amount of sensory integration research in recent years has investigated the possibility that object representation is not confined only to the modality in which they were learned. Instead, objects may be learned, stored and represented through a multisensory code (Erdogan, Yildirim & Jacobs, 2014; Lacey & Campbell, 2006). The multi-sensory theory suggests that haptic and visual senses can encode and retrieve information from the multi-sensory representations of objects equally efficiently and accurately. This would account for cross-modal memory ability found in object recognition experiments. Supporting the multi-modal representation theory, Erdogan et al. (2004) found that regardless of whether objects were presented visually, haptically or bimodally, participants' categorization of novel objects was similar. In addition to this, participants used similar exploratory processes to make similarity judgements in all conditions. This led Erdogan et al. (2004) to conclude that the processes involved in object shape perception were similar and that the representations themselves are similar as well. Thus, they hypothesized that the congruency between visual and haptic perceptual abilities in shape learning, categorization and recognition is the result of an overlapping, multi-modal or amodal representation of these objects.

A problem with the suggestion that the mental representations of objects are entirely multisensory in nature is that many studies report a within-modality accuracy advantage when encoding and recognition modalities match (Newell, et al., 2005; Norman, Norman, Clayton, Laineckhamy & Zelke, 2004), that is, these studies support encoding specificity. In a study by Newell et al. (2005) participants who were presented with scenes in one modality showed more accurate performance when recognizing the

object or scene in the same modality compared to cross-modally. They also found that participants in the visual learning and visual recognition condition had the most accurate performance, while participants who learned scenes visually and recognized scenes haptically produced the most errors. This suggests that the transfer of visual information—which is important for haptic scene recognition in this experiment—is quite poor. The asymmetric accuracy performance between within-modality and cross-modal identification conditions in object recognition experiments supports encoding specificity, especially in the visual modality. These results also suggest that while visual and haptic representations may be similar or overlapping, they may not be explained by a single, identical multimodal representation (Newell et al, 2005; Norman et al., 2004). If this were the case, recognition performance would be equivalent in both modalities regardless of learning condition.

While encoding specificity seems to be robustly supported in perceptual research, most of these supportive findings are more pronounced for the visual modality compared to the haptic modality (Norman et al, 2004; Newell et al., 2005, Desmarais et al, 2017). Some studies have found that participants who learn to recognize objects haptically are no faster at recognizing objects haptically compared to visually, while visual learning leads to superior visual recognition performance and equivalent or poorer recognition in the haptic modality (Lacey & Campbell, 2006; Lacey et al, 2006, Desmarais et al, 2017). Lacey and Campbell (2006) found that dynamic visual noise (a static-like moving visual stimulus) disrupted haptic and visual object encoding to the same extent. This suggests the transfer from haptic processing into visual information during the encoding of objects was disrupted by visual interference. They also found that visual interference at retrieval

resulted in poorer haptic recognition performance (Lacey & Campbell, 2006). This suggests that the representations formed from both visual and haptic learning of objects are multisensory, but are either primarily visual or that haptic information at processing can be easily transferred into visual mental representations.

To support the idea that the transfer from haptic to visual information is superior to the reverse, Desmarais et al. (2017) demonstrated encoding specificity for visual, but not for haptic recognition of novel objects. When participants learned to recognize novel objects visually, visual recognition performance was superior to haptic recognition. However, contrary to the assumptions of encoding specificity, haptic and visual recognition accuracy were equivalent when participants learned to recognize objects haptically. This finding demonstrates superior transfer from haptic learning to visual information, while transfer from visual learning into haptic information was not as efficient or accurate. These results refute encoding specificity, as haptic learning did not lead to superior performance during haptic recognition. The authors of the paper suggested this may be due to participants' use of visual imagery strategies during haptic encoding of stimuli, allowing for the creation of a dual (visual and haptic) representation of an object even though it was not visible to participants at encoding (Desmarais et al, 2017). Further supporting the visualization theory, participants who learned objects visually showed poorer performance during haptic recognition of objects, which may be explained by the fact that there is no reported haptic equivalent to visual imagery. Participants likely do not imagine what an object will feel like when they see it, but they may imagine what the object looks like when they are touching it.

Visual imagery can be defined as the generation of a mental representation of an object when the object is not physically present, or cannot be seen (Borst, Ganis, Thompson & Kosslyn, 2012). Researchers who have studied mental imagery have often used various techniques to encourage imagery such as peg word mnemonics (using a visual image to represent items on a list as a memory device for a sequence) (Quinn & McConnell, 1996). It is important to note, however, that in previous studies (e.g., Quinn & McConnell, 1996, Klatzky, Lederman & Matula, 1991) imagery processes were encouraged through verbal description of scenes or highly-imageable objects. Little research has been done on the effectiveness or use of visualization during haptic perception, although visual imagery abilities do appear to facilitate object learning. Research has found that superior performance on visual imagery tasks such as spatial location recall and mental rotation in children is related to better performance on cross-modal recognition of objects in visuo-haptic object learning (Stolz-Loike & Bornstein, 1987). As visual image generation is described as forming a mental representation for an object that is not visibly present, it stands to reason that when participants encounter and perceive objects or scenes haptically, they may generate mental images of the objects they are feeling but cannot see. This would promote the transfer of information from haptic to visual modalities during object learning, as imagery would give participants both haptic and visual information about objects at encoding.

If visual imagery is responsible for the lack of encoding specificity in haptic object recognition, researchers can use an interference paradigm to interrupt these imagery processes and isolate the effects it has on haptic encoding and cross-modal object identification. While few studies have used visualization interference for haptically

presented information, many studies have investigated the effects interference has on a variety of visual imagery tasks. These interference methods include Dynamic Visual Noise (Andrade, Kemps, Wernier, May, & Szmalec, 2002; Lacey & Campbell, 2006; McConnell & Quinn, 1996), colour discrimination tasks (Klauer & Zhao, 2004; Tresch, Sinnamon & Seamon, 1993), abstract paintings, and visual matrix patterns (Della Salla, Gray, Baddeley, Allamano & Wilson, 1999). Findings from these studies suggest that irrelevant visual information may disrupt visualization processes during encoding. Therefore, tasks that require visualization or occupy visual processing resources should adequately interfere with and disrupt any visualization that may occur during object learning.

While research has shown that concurrent, irrelevant visual noise sufficiently interferes with visual imagery for shape properties, this same interference may not occur when distraction tasks or target objects are familiar. Familiar objects allow for semantic or verbal labelling, which enables participants to avoid generating mental images. Thus, the use of familiar distractors may not adequately interfere with visualization. This has been shown by Klatzky et al. (1991) who asked participants questions based on highly imageable scenarios that were either difficult (i.e., “which is rougher, a plastic plate or a tomato?”) or easy (i.e., “what is rougher, sandpaper or binder paper?”) to answer. There was little or no reported use of visualization for easy questions, which could be answered by semantic information, while harder questions led to more active imagery processes, arguably due to lack of semantic knowledge and experience with these kinds of situations. The use of unfamiliar distractors tasks in a memory task may then require re-visualization for recognition, which could interfere with visual imagery information

coming from other modalities during object learning (Lacey, Flueckiger, Lava, & Sathian, 2010). Therefore, a concurrent visual memory task using unfamiliar distractor images or symbols may effectively occupy imagery resources and interfere with visualization during haptic encoding, allowing a researcher to examine the effects of visualization on haptic encoding of objects.

The aim of the present study was to examine whether the previously found lack of encoding specificity in the haptic modality (Desmarais et al, 2017) was a result of visualization processes during haptic encoding. To evaluate this, I used an interference task to disrupt visual image generation during the encoding of novel objects. Participants were presented with sequences of learning trials (where they were presented with objects to explore and remember) and test trials (where they were asked to recall the name of each object) until they could identify all objects flawlessly (criterion). Half of the participants were presented with visual learning trials, while the other half of participants were presented with haptic learning trials. All participants were presented with both visual test trials and haptic test trials to evaluate encoding specificity. Crucially, in each group two-thirds of the participants were asked to complete a secondary task during learning trials (one-third did not complete a secondary task). Specifically, participants completing a secondary task were asked to hold in memory either unfamiliar visually-presented symbols, or visually-presented English letters. Based on previous research, I hypothesized that:

- (1) the participants who are asked to complete a secondary task will require more trials to reach criterion than participants who are not asked to complete a secondary task.

- a. This effect will be larger for participants in the haptic learning condition than for participants in the visual learning condition
 - b. This effect will be larger for participants who are asked to complete a secondary task using unfamiliar symbols than for participants who are asked to complete a secondary task using letters
- (2) For participants who learn to recognize objects haptically, completing a secondary task using unfamiliar visual stimuli will restore encoding specificity: participants will produce more errors when recognizing objects visually than when recognizing objects haptically, while there should be no difference between visual and haptic identification when participants complete a secondary task using letters, or when there is no secondary task.

Method

Participants

Seventy-one undergraduate students from Mount Allison University (50 female, three other, mean age = 20.63 years, $SD=2.48$) were recruited for the study. Participants were recruited from introductory psychology classes, as well as through in-class recruitment. As compensation, participants received 0.5 course credits per half hour of participation for introductory psychology students, or \$5 per half hour for students not enrolled in these classes.

Materials

I used eight novel psychophysically-scaled objects (Desmarais & Dixon, 2005). Each object was 81 mm long, made of light gray, hard and smooth PVC. The objects were designed to vary along three dimensions: curvature, thickness and tapering.

Each of the eight objects was paired with a non-word label not associated with any semantic meaning: “yoot”, “baiv”, “fint”, “malg”, “verp”, “grov”, “hong” or “jorl”. The specific label associated with each object was varied between participants, and the specific pairing was randomly selected as one of eight possible object-label pairings.

Objects were mounted on a black PVC apparatus measuring 45 cm x 40 cm by 44.5 cm (See figure 1). The apparatus allowed participants to perceive objects visually, haptically or bi-modally. Within the apparatus, there were two spinning central cylinders, each containing the full set of eight objects attached by spokes to the cylinders. The cylinders were stacked one above the other, with the top cylinder used for visual presentation, and the bottom cylinder used for haptic presentation. Each cylinder could be spun by the researcher to position objects in a window measuring 6.5 cm by 10 cm near the top of the apparatus for visual presentation, or positioning objects in a window at the bottom of the apparatus which participants were asked to reach into, with their view of the object they were grasping concealed by two broom brushes. The experiment was run using Superlab 5.0 and Cedrus SV1 Voice Key. Researcher commands were given through headphones worn by the participants, and the participants’ vocal responses were recorded by a headset microphone.

In the non-verbal condition, visual distractors were unfamiliar, difficult to verbally code ideographs from the Yi language (see figure 2). Since these objects could not be easily verbally coded, re-visualization should be necessary for the distractors to be remembered (Klauer & Zhao, 2004). The second distraction



Figure 1. Apparatus used. Left: Participant viewpoint. During visual encoding and identification trials, objects appeared in the top window, while in haptic encoding and identification trials, participants reached into the bottom opening which concealed any view of the object being grasped. In the top left, a card displaying the eight non-word labels was placed in participants' field of vision. Right: experimenter view, showing identical object sets attached to the central cylinders of the apparatus.



Figure 2. Yi symbols used as non-verbal distractors as they appear at the end of learning trials, prior to test trials.

condition used all 26 letters in the English alphabet, and were easy to semantically label. Target distractors in both conditions were first displayed alone, centered on a sheet of plain letter paper during encoding displayed on letter paper written in 210 pt. in either Lucida Handwriting (for verbal) or Microsoft Yi Baiti font (for non-verbal) font. For recognition, the four target distractors were displayed alongside four lures in a set of eight, with font size reduced to 110 pt. for each symbol.

Procedure

Participants completed a series of eight learning trials followed by eight test trials. During learning trials, participants explored objects in one of two learning modalities: visual or haptic. During testing trials, participants completed blocks of visual identification trials and blocks of haptic identification trials. Participants were also assigned to one of three distraction conditions: the letter condition, the Yi symbol condition, or a control condition. This yielded a 2x3 mixed subjects design where learning condition (haptic or visual) as well as distraction task condition (verbal, nonverbal or control) were between subjects variables, leading to six conditions; haptic learning with nonverbal distraction, haptic learning with verbal distraction, haptic learning with no concurrent distraction task (control), visual learning with nonverbal distraction, visual learning with verbal distraction, and visual learning with no concurrent task (control). Using a random number generator, participants were randomly assigned to each of the six conditions. All participants were tested in both modalities. Performance on identification trials in both the haptic and visual modalities was the within-subjects variable.

Visual learning. In the visual learning condition, participants sat in front of the apparatus at arm's length and were asked to close their eyes while the experimenter positioned the objects in the viewing window. When the participants heard the cue "open" they opened their eyes, while the experimenter read the label corresponding to the object presented. Participants viewed the object for two seconds, at which point the cue "end of trial" signaled participants to close their eyes while the experimenter set up the next trial. After each of the eight objects was shown once in a random order, participants moved on to test trials. A visual representation of the sequences of events in the learning phase is shown in figure 3.

Haptic learning. Haptic learning trials were identical to visual learning trials, except the haptic trials began with the auditory cue "reach", at which point participants reached into the bottom window to grasp the object being presented and after five seconds, the cue "end of trial" signaled participants to remove their hand and wait for the next trial to begin. Differences in exploration time for each modality were used in accordance with convention as haptic encoding, sequential in nature, generally takes longer than visual encoding, and was also used to replicate the methodology of Desmarais et al. (2017).

Test trials. After participants completed eight learning trials, participants were given one block each of visual and haptic identification trials. During testing trials, participants closed their eyes while the experimenter rotated the cylinder to set up one object for identification. When participants heard the auditory cue "open" (for visual trials) or "reach" (for haptic trials), participants opened their eyes or reached into the apparatus and identified the object presented. If they were unsure, participants were

told to make their best guess about the object's label. The experimenter recorded the participants' responses with a key press to analyze their accuracy. Each object was presented once in a random order, and the cue "end of trial", triggered by participants' verbal responses, signaled participants to close their eyes or remove their hand and wait for the next object to be presented. The order of the two testing conditions (either haptic or visual) alternated between testing blocks. At the end of 16 test trials (one block of eight visual identification trials and one block of eight haptic identification trials), participants were given eight more learning trials, followed by another block of 16 test trials. Learning and test blocks continued until participants could correctly identify all objects in both modalities for two consecutive testing blocks. The number of blocks required to reach criterion and the overall number of errors made in each condition were averaged to determine participants' accuracy.

Distraction tasks. Participants were randomly assigned to one of three distraction groups: unfamiliar Yi symbols, familiar English letters, and a control condition where there was no secondary task. The sequence of events in the two distraction conditions followed the same procedure. In these conditions, at the beginning of learning trials 1, 3, 5, and 7, participants were first presented with a visual stimulus for two seconds and were told to remember each of the four stimuli they saw. At the end of the eight learning trials, participants were shown an 8.5"x11" sheet of paper containing the four targets along with four novel distractor symbols and were asked to identify the four target items they had already seen. The experimenter recorded the participants' accuracy for analysis. The experiment took between forty-five and 90 minutes to complete.

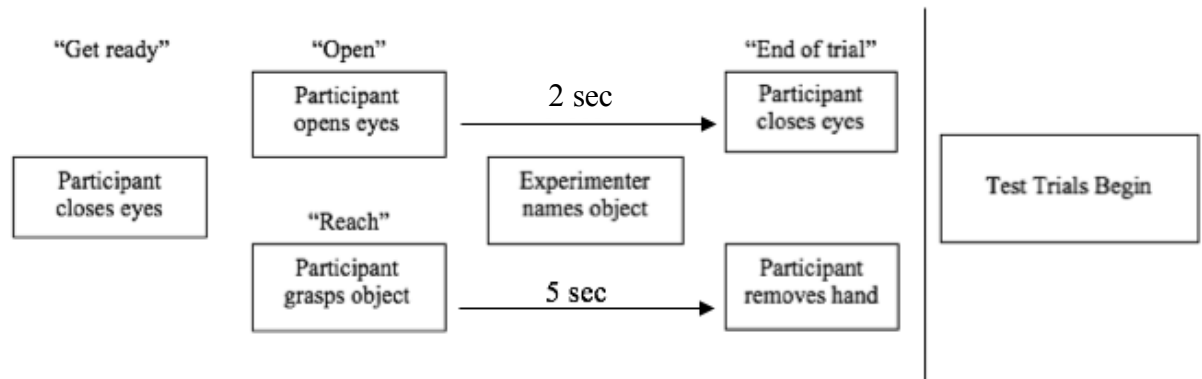


Figure 3a. Sequence of events of learning trials in the control condition (no distractors). All auditory cues given in the experiment are shown in quotation marks. The top set depicts the events for participants in the visual learning condition and the bottom set depicts the events for haptic learning conditions.

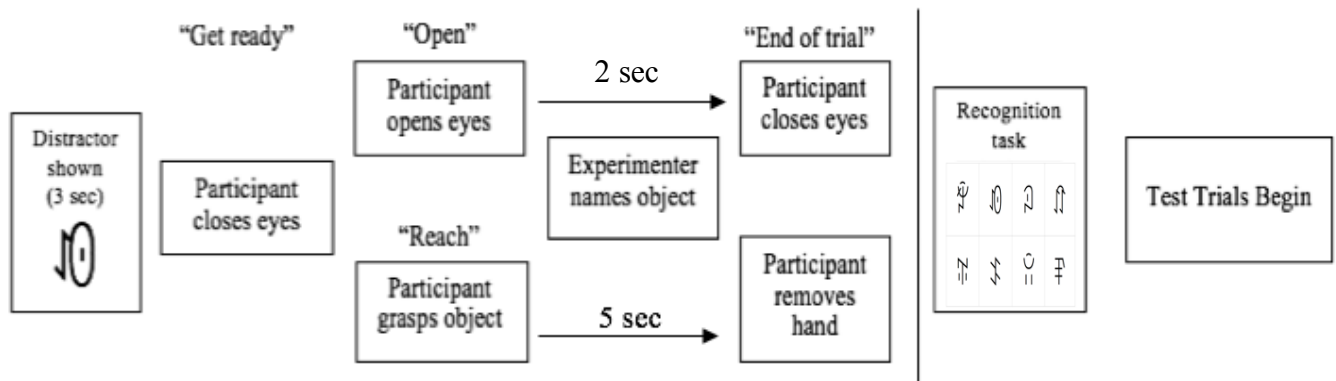


Figure 3b. Sequence of events of learning trials with a visualization distractor. Visual distractors were presented four times, at the beginning of trials 1, 3, 5, and 7. The recognition task (with four target and four lure distractors) was presented at the end of the eighth and final learning trial prior to test trials.

Results

Data from 10 participants were excluded from analysis due to inability to complete the memory task either within 20 blocks, or after 90 minutes of testing. Of the participants excluded, eight were in the haptic learning condition, and two were in the visual learning condition. Excluded data existed in all three distraction conditions; three participants were presented with non-verbal distractors, four were shown verbal distractors and three were in the control condition. The data analyzed are from the 60 participants that completed the task.

Distraction Task Performance

The participants who completed a non-verbal distraction task had an average accuracy rate of 91%, while the participants who completed a verbal distraction task had an accuracy rate of 95%. A t-test revealed no significant differences between performance on verbal and non-verbal distraction tasks, $t(19) = -1.72$, *ns*.

Blocks to Criterion

The number of blocks participants needed to reach criterion was entered in a 2 (learning condition) x 2 (identification modality) x 3 (distraction condition) mixed ANOVA where identification modality (visual or haptic) was the within-subjects factor, and both learning condition (visual or haptic) and distraction condition (non-verbal distractors, verbal distractors, or control) were between-subjects variables. Analyses revealed a main effect of identification modality, $F(1, 54) = 18.84$, $p < .001$, $partial \eta^2 = .259$. Generally, participants needed fewer blocks to reach criterion when identifying objects visually (mean number of blocks = 7.77, $SD = 3.31$) compared to haptically (mean = 8.68 blocks, $SD = 3.41$).

There was also a main effect of distraction condition, $F(2,54) = 3.55, p = .035$, $partial \eta^2 = .116$. To determine which condition was the source of this main effect, I conducted non-directional Bonferroni-corrected independent samples t-tests. These comparisons revealed that participants in the verbal distraction condition (mean = 9.70 blocks, $SD = 3.94$) required more blocks to reach criterion than participants in the control condition (mean = 7.13 blocks, $SD = 1.78$), $t(26.43) = 2.67, p = .013$, Cohen's $d = .84$. There were no significant differences between blocks to criterion for participants in the control and in the non-verbal (Mean = 7.85 blocks, $SD = 3.24$) conditions $t(29.45) = 0.88, ns.$, or participants in the verbal and the nonverbal conditions $t(38) = -1.62, ns.$ (see figure 4).

The main effect of identification modality was qualified by an interaction between identification modality and learning modality, $F(1,54) = 6.78, p = .012$, $partial \eta^2 = .112$ (see figure 5). Planned directional pairwise t-tests revealed that when participants learned to recognize objects visually, more blocks were required to reach criterion for haptic identification (mean = 8.47 blocks, $SD = 3.80$) than for visual identification (mean = 7.00 blocks, $SD = 3.50$), $t(29) = -3.77, p < .001$, Cohen's $d = .40$. When participants learned to recognize objects haptically, they did not differ on number of blocks to criterion for haptic identification (mean = 8.90 blocks, $SD = 3.02$) compared to visual identification of objects (mean = 8.53 blocks, $SD = 2.97$), $t(29) = -2.48, p = .019, ns.$

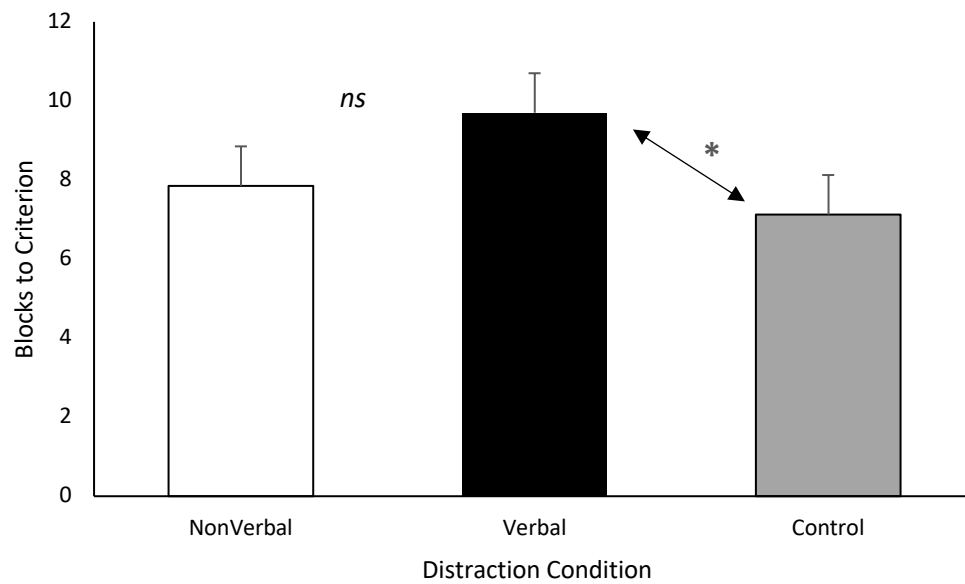


Figure 4. Mean number of blocks to criterion in each distraction condition. Asterisks indicate significant difference between the number of blocks to criterion in the verbal and control distraction conditions.

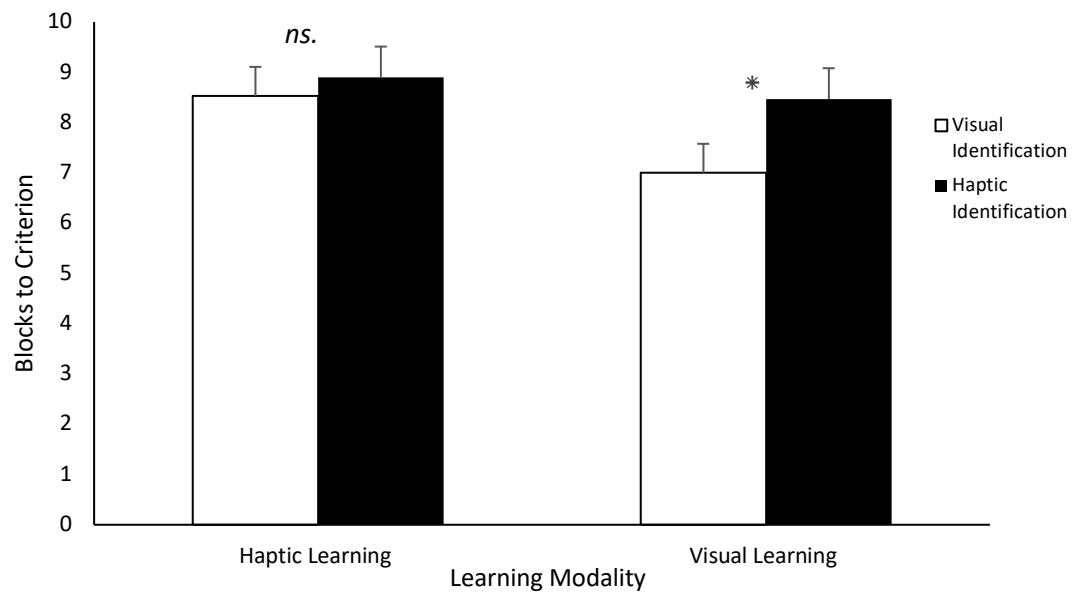


Figure 5. Mean number of blocks required for participants to reach criterion for haptic identification and visual identification. In the visual or haptic learning modality. Asterisks represent significant differences.

Number of Errors

Next, I entered each participant's number of errors during visual and haptic identification in a 2(learning condition) x 2(Identification modality) x 3(distraction condition) mixed ANOVA for which the repeated factor was identification modality (visual or haptic), and the between-subjects factors were learning condition (visual or haptic) and distraction condition (non-verbal distractors, verbal distractors, or control). The analysis revealed two trends. First there was a trend towards a main effect of learning condition $F(1,54) = 3.32, p = .074, \text{partial } \eta^2 = .058$. Generally, participants in the visual learning condition produced fewer errors (mean number of errors = 22.35, $SD = 14.81$) than participants in the haptic learning condition (mean = 28.83 errors, $SD = 12.98$). There was also a trend towards a main effect of distractor condition $F(2,54) = 2.42, p = .098, \text{partial } \eta^2 = .082$. To verify whether this effect was consistent with the one observed for blocks to criterion, I carried out post-hoc independent samples t-tests that revealed that participants in the verbal distraction condition produced marginally more errors (mean number of errors = 30.84 $SD = 16.28$) compared to participants in the control distraction condition (mean number of errors = 21.45, $SD = 7.94$), $t(27.56) = 2.32, p = .028$ Cohen's $d = .73$. Other pairwise t-tests were not significant; participants in the nonverbal distraction condition (mean number errors= 24.5, $SD=15.76$), did not produce more errors than participants in the control condition $t(28.06)=.77, ns$. Participants in the verbal distraction condition did not produce significantly more errors than participants in the verbal condition $t(38)=.219, ns$.

Discussion

I aimed to determine the effects of visualization on participants' ability to learn to recognize novel objects. Consistent with previous research, participants in the visual learning condition were better at visual identification than haptic identification, demonstrating encoding specificity for the visual modality reported by previous experiments (Desmarais et al., 2017, Newell et al, 2005). Additionally, as was observed by Desmarais et al. (2017), participants who learned to recognize objects haptically were not better at identifying objects haptically compared to identifying objects visually. Contrary to my hypotheses, the distraction condition which produced the greatest impact on object identification performance was the verbal condition, rather than the non-verbal condition, and this effect did not interact with other factors.

Interestingly, of the participants that were unable to identify the objects to criterion within the time limits of the study, and were thus excluded from analyses, no distraction conditions appeared to disproportionately affect the ability to complete the task. However, of the ten excluded participants, eight were assigned to the haptic learning condition, while only two were in the visual learning condition. This suggests that learning to recognize objects haptically may be more difficult for participants than learning to recognize objects visually.

An unexpected and interesting finding was that participants in the verbal distraction condition needed more blocks to reach criterion and produced more errors than the control distraction condition. The pattern of errors for participants in the verbal condition also showed a trend toward more errors than participants in the

nonverbal distraction condition. This effect of verbal interference occurred independent of learning modality or identification modality. These results suggest that there is a verbalization process through which we create a description of novel objects. This verbal description may facilitate the formation of memory representations for objects. While Desmarais et al. (2017) suggested that participants may visualize haptically encoded objects, my results do not support this explanation, as the condition that was designed to interfere with visualization processes—the nonverbal condition—did not impair identification performance. This suggests that the formation of mental representations for novel objects is mediated by verbal strategies such as the sub-vocal description of features of objects (e.g. fat/thin, curvy/straight, pointed/blunt). Indeed, some participants volunteered the information that they used verbal strategies to remember the objects, including pairing the names of similar objects together to aid in categorization and memory for the objects, demonstrating the clear involvement of verbal strategies for object encoding.

Verbalization Processes at Encoding

My results contribute to a growing body of research suggesting that there is a verbal component to the development of memory representations for objects (Lacey & Campbell, 2006; Postle, D'Esposito & Corkin, 2005; Uittenhove, Chaabi, Camos, & Barrouillet, 2019). Lacey and Campbell (2006) used an interference paradigm similar to that used in the present study to test participants' ability to learn to identify familiar and unfamiliar objects. They found that participants' identification performance for unfamiliar objects was significantly disrupted by both visual (Dynamic Visual Noise) and verbal interference (irrelevant speech) at encoding. The

authors concluded that the formation of mental representations for unfamiliar objects is at least partially facilitated by verbal description of the objects during learning (Lacey & Campbell, 2006). My study corroborates Lacey and Campbell's (2006) verbal account of memory representation formation for unfamiliar objects, as participants' performance was significantly impaired in the verbal interference task at encoding, regardless of either learning or identification modality. However, contrary to Lacey and Campbell's (2006) findings, the visual interference condition in my study did not lead to any reduction in performance

The difference between my results and those of Lacey and Campbell (2006) may be explained by the differences in complexity of the stimuli participants were presented with. Lacey and Campbell (2006) used complex hiking equipment which their participants were unfamiliar with. Conversely, the objects in my study were simpler, and therefore may have been easier to verbalize than the unfamiliar stimuli used by Lacey and Campbell (2006). The relative importance of verbal coding of objects may then be a function of the object's complexity. If an object is simple and easy to describe, using input from multiple modalities to remember the object likely would not be necessary for recognition. However, if an object is complex, a mental representation may be aided by inputs from different modalities that contribute to the understanding of the complex figure. If this is the case, participants may have avoided visualizing the stimuli in my experiment using verbal coding, while in Lacey and Campbell's (2006) study, their complex unfamiliar stimuli may have been harder to verbalize, requiring participants to use multiple sensory inputs—including the visualization of objects—to form a mental representation. This would explain why

completing the verbal interference condition led to poorer identification performance for participants than the nonverbal (visual) task in my experiment, while visual and verbal interference both had a detrimental effect on identification in Lacey and Campbell's (2006) study.

Research with other complex stimuli may offer some more insight into the mechanisms by which the complexity of stimuli might be related to the sensory strategies used to encode them. For example, in studies where participants were shown scenes with multiple components, there was a within-modality advantage for scene recognition (Newell et al., 2005). Participants who learned scenes in a specific modality demonstrated more accurate scene recognition when tested in the same modality. As suggested in the discussion of Desmarais et al. (2017), the scenes used by Newell et al (2005) may have been more complex than the stimuli in my study, as the scenes had multiple separate components. This complexity may have made it harder for participants to verbalize these stimuli compared to a single object. If participants were less able to verbalize the scenes at encoding, they may have been less able to easily form a verbally-mediated representation that would be equally accessible to both visual and haptic modalities. This would require participants to rely more on sensory information from the modalities in which the scenes were encoded, potentially explaining why participants in the study by Newell et al (2005) demonstrated encoding specificity. Conversely, participants' ability to verbalize the objects in my study may have allowed them to avoid relying on modality-specific information in favor a verbal representation of the objects being encoded that could be easily accessed by both the visual and the haptic modalities.

Verbal coding appears to be important for the development of memory representations for objects specifically. Postle et al. (2005) found that performance on object recognition during an *n*-back memory test was sensitive to verbal interference, while spatial memory was not influenced by verbal distraction. This may explain why participants in both the visual and haptic learning conditions of my study saw a deficit in identification performance when they were given verbal distractors. Memory for objects seems to be facilitated by the verbal rather than the visual coding of features, even if the objects are first presented in the visual modality. It is important to note that the objects used by Postle et al. (2005) were simple objects, like those used in my study. My research combined with the findings of Postle et al. (2005), suggest that the development of a memory representation for simple objects relies heavily on verbalization, while the results of Lacey and Campbell (2006) and Newell et al. (2005) suggest that verbalization may be important, but less crucial for the encoding of complex scenes or objects.

Visualization

As mentioned earlier, the disruptive effects of visual interference on object identification found by Lacey and Campbell (2006) were not replicated in my study. While the conclusion that verbalization is more important for the formation of mental representations of objects than visualization certainly is possible, it may also be possible that the visual interference task (the non-verbal task) was not equivalent to the verbal task. It is true that the non-verbal symbols were more complex, in general, than the verbal distractors, which comprised familiar English letters. Therefore, one may conclude that the observed effects may have been the result of a lack of effort in

the non-verbal task, as the difficulty of completing both the distraction task and the encoding of unfamiliar objects may have been too much for participants to cognitively manage. This would suggest that participants may have stopped trying to remember the symbols altogether. However, statistical analyses revealed that participants' performance on the nonverbal distraction task was not significantly different from performance on the verbal distraction task, which indicates that the tasks were comparable in difficulty. In addition to this, both groups showed far better than chance performance on the distraction tasks, meaning that participants were likely actively participating in all aspects of the experiment, so one can rule out the possibility that the distraction tasks required unequal levels of attentional engagement or cognitive resources.

While the difficulty of the tasks may not have been the source of the differences found, it is still possible that the non-verbal task may not have adequately interfered with visualization processes. This presents a theoretical problem; if the nonverbal task did not selectively interfere with visualization, it becomes impossible to definitively rule out the role of visualization in the formation of representations for objects that were encoded haptically. Research suggests that there may be a functional dissociation between visual imagery and visual short term memory (Andrade et al., 2000). If this is the case, the visual interference condition in this study (a visual short term memory task) may not have disrupted visualization which would explain why it did not interfere with performance in the identification task.

Although the cognitive underpinnings of visual short term memory and visual imagery may be dissociable, research indicates that visual memory tasks are more

generally attentionally taxing, disrupting overall memory resources, and creating deficits in both visual short term memory and working memory performance (imagery), rather than selectively interfering with visual imagery (Andrade et al., 2000, Quinn & McConnell, 1996). If this was the case in my study, one would expect to see a general reduction in identification performance in both the visual and haptic identification conditions for the non-verbal distraction tasks. Contrary to this, I found no effect of the non-verbal distractors on performance, which again points to the possibility that verbalization is the most important sensory process for the formation of a mental representations of simple objects on my study, while visualization appeared to be less important.

Active Rehearsal Versus Passive Interference

The verbal distraction task may have been more conducive to the rehearsal of information than the non-verbal distraction task. The subvocal rehearsal of information can occur continuously, and may have interrupted the encoding process necessary to create memory representations for the objects in the main task of the study. In contrast, there does not appear to be much research regarding a mechanism for the visual rehearsal of information. Therefore, the active and continuous rehearsal of the distractor items in the verbal condition may have been more attentionally demanding for participants to complete when combined with the primary object learning task. This may explain the disruptive effects of the verbal interference task in both the visual and haptic learning conditions of this experiment. Rather than sensory interference for a certain modality, this finding may have more to do with attentional interference. Indeed, research with verbal distraction tasks consisting of

memory for letters show that active tasks involving active rehearsal and free recall of letters reduces accuracy performance on a primary visuo-spatial memory task for the location of a stimulus light on a screen. In contrast, the passive recognition task in this study in which participants were asked “did you see this letter, yes/no?”, was less detrimental for visuo-spatial memory accuracy on the stimulus location memory task. (Uittenhove et al., 2019).

Although the above argument suggests that attention is the mechanism by which verbal interference has its effects on the encoding of novel objects, Lacey and Campbell’s (2006) methodology and results indicate that this may not be the case. Their study used irrelevant speech as verbal interference. This information did not need to be remembered and was therefore not rehearsed by participants in their study. Yet, Lacey and Campbell (2006) still found that participants who were exposed to this passive verbal interference had lower recognition accuracy for the objects. Therefore, although attention is an important factor in learning and memory, it appears that the continuous attentional demand of the task was not the reason for the reduction in performance in the verbal interference condition.

Limitations and Future Research

The most likely explanation for the lack of effect of visual interference found in my study and the most significant limitation of my study is the difference in complexity of the objects I used compared to those used by Lacey and Campbell (2006). As the objects were rather simple in my study, verbal coding appeared to be sufficient to complete the experiment to criterion. Had objects been more complex and more difficult to verbalize, participants in the haptic learning condition may have

relied more on sensory information provided them during encoding, and performed in accordance with the expectations of encoding specificity. Unfortunately, because intramodal performance was not tested by Lacey and Campbell (2006), this hypothesis cannot be directly addressed. This presents an important gap in the literature that may contribute to the understanding of multisensory coding of novel objects, especially those learned in the haptic modality. A replication of the current study with more complex objects may help researchers address the relationship between the processes of verbalization, visualization and the complexity of objects in the formation of a mental representation for objects. If encoding specificity is found for complex objects, or if visual interference influences identification accuracy for more complex stimuli, this would allow researchers to understand the violation of encoding specificity that was shown both in my study and in Desmarais et al. (2017) which used the same simple stimuli.

Conclusion

I studied the effects of visual and verbal distraction on both haptic and visual identification of novel objects. I replicated that encoding specificity is observed when participants learn to recognize objects visually, but not when they learn to recognize objects haptically. Furthermore, contrary to my hypotheses, the verbal distraction condition interfered with identification performance more than either the nonverbal or control distraction conditions. This suggests that mental representations for unfamiliar objects are verbally mediated and that verbalization processes at encoding may be more important than visualization processes for the formation of mental representations of simple objects.

References

- Andrade, J., Kemps, E., Werniers, Y., May, J., Szmalec, A., (2002). Insensitivity of visual short-term memory to irrelevant visual information. *The Quarterly Journal of Experimental Psychology*, 55A, 753-774.
doi:10.1080/02724980143000541.
- Borst, G., Niven, E., Logie, R. (2012). Visual mental image generation does not overlap with visual short-term memory: A dual-task interference study. *Memory & Cognition*, 40, 360-372. doi:10.3758/s13421-011-0151-7.
- Della Salla, S., Gray, C., Baddeley, A., Allamano, N. & Wilson, L. (1999). Pattern span: A tool for unwelding visuo-spatial memory, *Neuropsychologia*, 37, 1189-1199.
- Desmarais, G., Meade, M., Wells, T., & Nadeau, M. (2017). Visuo-haptic integration in object identification using novel objects. *Attention, Perception & Psychophysics*, 79, 2478-2498. doi:10.3758/s13414-017-1382-x.
- Erdogan, G., Yildirim, I., & Jacobs, R. (2014). Transfer of object shape knowledge across visual and haptic modalities. *Cognitive Science Society*, 36, 2163-2168. <http://escholarship.org/uc/item/29b6w0hk>
- Klatzky, R. Lederman, S. & Matula, D. (1991). Imagined haptic exploration in judgements of object properties. *Journal of Experimental Psychology: Learning Memory and Cognition*, 17, 314-322.
- Klauer, K. & Zhao, Z. (2004). Double dissociations in visual and spatial short-term memory. *Journal of Experimental Psychology: General*, 133, 355-381.

doi:10.1037/0096-3445.133.3.355.

Lacey, S & Campbell, C. (2006). Mental Representation in visual/haptic crossmodal memory: evidence from interference effects. *The Quarterly Journal of Experimental Psychology*, *59*, 361-376. doi:10.1080/17470210500173232.

Lacey, S., Campbell, C., & Sathian, K. (2007). Vision and touch: Multiple or multisensory representations of objects? *Perception*, *36*, 1513-1521. doi:10.1068/p5850.

Lacey, S., Flueckiger, P., Stilla, R., Lava, M., & Sathian, K. (2010). Object familiarity modulates the relationship between visual object imagery and haptic shape perception. *Neuroimage*, *49*, 1977-1990. doi:10.1016/j.neuroimage.2009.10.081.

Newell, F., Woods, A., Mernagh, M., & Bühlhoff, H. (2005). Visual, haptic and crossmodal recognition of scenes. *Experimental Brain Research*, *161*, 233-242. doi:10.1007/s00221-004-2067-y.

Norman J., Norman, H., Clayton, A., Laineckhamy, J. & Zelke, G. (2004). The visual and haptic perception of natural object shape. *Perception and Psychophysics*, *66*, 342-351. doi:10.3758/BF03194883.

Postle, B., D'Esposito, M. & Corkin, S. (2005). Effects of verbal and nonverbal interference on spatial and object visual working memory. *Memory & Cognition*, *33*, 203-212. doi:10.3758/BF03195309

Quinn J., & McConnell, J. (1996). Irrelevant pictures in visual working memory. *The Quarterly Journal of Experimental Psychology*, *49A*, 200-215.

- Stoltz-Loike, M. & Bornstein, M. (1987). The role of imagery, language, and meta-memory in cross-modal transfer in children. *Psychological Research, 49*, 63-68.
- Tresch, M., Sinnamon, H., Seamon, J. (1993). Double dissociations of spatial and object visual memory: evidence from selective interference in intact human subjects. *Neuropsychologia, 31*, 211-219.
- Tulving, E. & Thompson, D. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*, 352-373.
- Uittenhove, K., Chaabi, L., Camos, V., & Barrouillet, P. (2019). Is Working Memory Storage Intrinsically Domain-Specific? *Journal of Experimental Psychology: General*. Advance online publication. doi:10.1037/xge0000566.