On Memory for Everyday Symbols

by

Brady R. T. Roberts

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Psychology

Waterloo, Ontario, Canada, 2023

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	Dr. Ken McRae Professor, Department of Psychology University of Western Ontario
Supervisors	Dr. Myra A. Fernandes Professor, Department of Psychology
	Dr. Colin M. MacLeod Professor Emeritus, Department of Psychology
Internal Members	Dr. Evan F. Risko Associate Professor, Department of Psychology
	Dr. James Danckert Professor, Department of Psychology
Internal-external Member	Dr. Jay Dolmage Professor, Department of English Language and Literature

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

This research was conducted at the University of Waterloo by Brady Roberts under the supervision of Dr. Myra A. Fernandes and Dr. Colin M. MacLeod. Experiments 1-4 and 6 of this dissertation have been accepted for publication at the journal *Cognition* (Roberts et al., in press, see references). That article is co-authored by myself, Colin M. MacLeod, and Myra A. Fernandes. Experiment 5 of this dissertation was published by myself, Myra A. Fernandes, and Colin M. MacLeod in the *Journal of Applied Research in Memory and Cognition* (Roberts et al., 2023, see references). Portions of this dissertation have been derived from or taken verbatim from these manuscripts.

This research was supported by a Natural Sciences and Engineering Research Council (NSERC) of Canada Postgraduate Scholarship awarded to me, Brady R. T. Roberts, and by NSERC Discovery Grants A7459 to Colin M. MacLeod and 2020-03917 to Myra A. Fernandes.

Abstract

This thesis investigated the memorability of common graphic symbols (e.g., !@#$%) and logos. In an initial set of 4 experiments, participants were presented during study with symbols or words (e.g., $ or 'dollar'). In Experiment 1, memory performance assessed using free recall demonstrated markedly better memory for symbols relative to their word counterparts, manipulated within-subject. Experiment 2 systematically varied whether symbols and words were presented at encoding and during a subsequent recognition test, manipulated between-subjects. Results conceptually replicated the findings of the first experiment, showing superior memory for symbols even when the retrieval test and study design were changed. Furthermore, by analyzing group data based on which stimuli (words or symbols) were used in the encoding and retrieval phases of the experiment, symbol superiority in memory was determined to be driven by encoding-based mechanisms. An alternative explanation holds that symbols may benefit memory as a result of their smaller overall set size compared to words. Experiment 3 addressed this potential issue by restricting the to-be-remembered set of words to a single category (common kitchen produce) whose set size was like that of the symbols that I used. Once again, symbols were better remembered than the words. This experiment showed not only that set size was not likely to be driving the previously seen memory benefit for symbols, but also that representing abstract concepts with symbols successfully reversed the concreteness effect in memory: Symbols were better remembered even when compared to highly concrete nouns. A fourth experiment directly tested a dual coding account by comparing memory for symbols, pictures, and words. There, symbols and pictures were both better remembered than words, and memory for symbols and pictures did not differ. Symbols not only were remembered just as well as images—as I predicted based on dual coding theory—but they also entirely eliminated the concreteness advantage in memory for pictures as well: Memory for symbols representing abstract concepts was equivalent to that for pictures depicting concrete objects. In Experiment 5A and 5B, I compared memory for professional sports teams presented in three encoding conditions: team names only, team logos without team names, and team logos with integrated team names. Across two experiments, while memory was often best for logos relative to team names, familiarity moderated this relation. When assessing memory for team names, the magnitude of the benefit for the logos-only condition depended on whether participants knew what the logos represented. In the sixth and final experiment, 337 naïve participants rated the set of symbols used in Experiments 1-4 on their meaning-based familiarity with each symbol and on their frequency of encountering it. Machine learning estimations of inherent stimulus memorability were provided by the ResMem residual neural network. These computer-derived memorability estimates correlated with memory for symbols, but familiarity and frequency ratings did not. Hierarchical linear regression revealed that inherent memorability estimates explained significant portions of variance for symbol memory, over and above effects of familiarity and frequency. This dissertation is the first to present evidence that, like pictures, graphic symbols and logos are better remembered than words, in line with dual coding theory and with distinctiveness accounts. Symbols offer a visual referent for abstract concepts that are otherwise unlikely to be spontaneously imaged. Symbols also provide visual stimuli that are often both physically and conceptually unique. It is the visual nature of symbols that confers the impressive memory performance benefits.

Acknowledgements

First and foremost, absolutely no part of this dissertation would have been possible without the unwavering support and dedication of my graduate supervisors, Dr. Myra Fernandes and Dr. Colin MacLeod. Their feedback, inspiration, funding support, guidance, and general life advice was nothing short of instrumental for the past 6 years of my life. They are fair and honest, always pushing me to be a better scholar. Myra and Colin are excellent academics and even better mentors. They are what I want to be when I grow up.

On the home front, the love and support from my family has been crucial to my success. While not always exactly sure about what I do or when I am expected to finish, they never faltered in their support for my pursuit of this degree. Thanks, Mom and Dad.

I also owe considerable thanks to my collaborator, mentor, and friend, Dr. Jeffrey Wammes. Without his willingness to take me on as a research assistant almost ten years ago, I surely would not be where I am today. He taught me much of what I know, and what I still need to learn.

I am also grateful to my dedicated research assistants over the years for their help gathering materials, collecting data, and providing feedback on my studies: Batul Karimjee, Carissa Ng, Lauren Homann, Philip He, Saad Qadeer, Tori Dickey, and Torin Young. Thanks, too, to Dr. Christopher Lee, Dr. Craig Zhou, Dr. Daniel Todorovic, Dr. Melissa Meade, and Dr. Ryan Yeung for their valuable feedback on my work while we shared a lab together. Further thanks go to my current lab mates who support me every day with critical feedback, coffee runs, and ample venting time: Adam Cox, Pelin Tanberg, Sophia Tran, Xinyi Lu, and Yadurshana Sivashankar. Special shoutout to my day-one cohort friends from beyond the lab for some much-needed time off with drinks, chicken fries, slo-pitch, and murder mysteries: Ethan Meyers, Kaiden Stewart, Laura Bianchi, Rebecca Trossman, and Shirley Ong.

Finally, none of this would have been possible without the dedicated support of Victoria Antinucci. Through thick and thin, she has been there for me and supported my pursuit of academia with admirable patience; her genuine kindness and generosity were what kept me going all these years. I look forward to the next chapter of life together.

Quote

"Yeah, yeah, but your scientists were so preoccupied with whether or not they could that they didn't stop to think if they should."

– Dr. Ian Malcolm

Table of Contents

List of Figures

List of Tables

Chapter 1 – General Introduction

1.1 Historical Context

Since the time of ancient Greece, philosophers have contemplated the definition and use of symbols. In those times, it was commonly held that ‘signs’ referred to aspects of nature that conveyed meaning, whereas ‘symbols’ referred to cultural conventions that had meaning. To these thinkers, the most obvious examples of symbols were words, which Aristotle referred to as the symbols of the inner images (Modrak, 2001). The notion of words as symbols has persisted for thousands of years. Writing in 1690, John Locke considered the ‘doctrine of signs’ to be fundamental to the communication of science and knowledge to others (da Costa e Silva, 2019). In Shakespeare’s *Romeo and Juliet*, Juliet laments that she cannot marry Romeo due to his family name, claiming “that which we call a rose by any other name would smell as sweet” (Shakespeare, 1597/2011, 2.2.890-891). In essence, Juliet is arguing that the words used to name things are entirely arbitrary, and if only Romeo was born under another family name, the two would be allowed to wed.

While most philosophical writing about symbols has been focused on words, contemporary definitions of symbols have explored other ways in which humans ascribe significance to graphic symbols that they have created (e.g., !@#$%). The purpose of the present dissertation was to investigate how these graphic symbols are represented and remembered. In so doing, empirical research on human cognition was used to bridge the gap between philosophy and psychology.

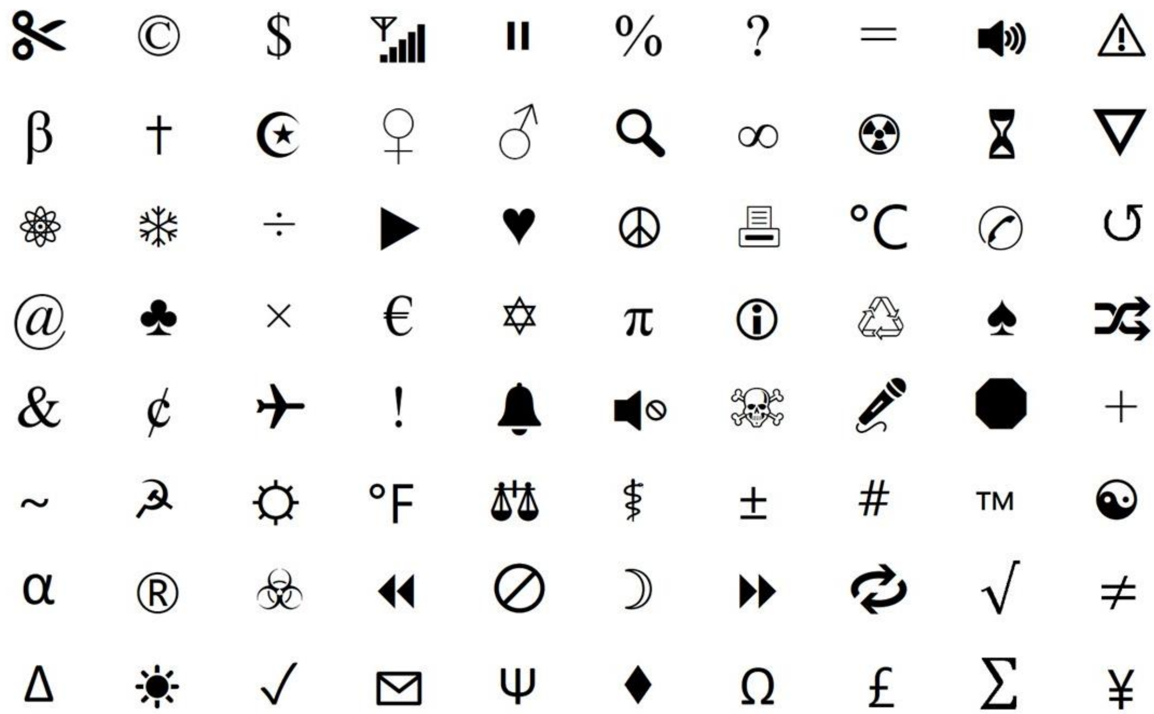
The philosophical field of semiotics is entirely devoted to the study of signs. American philosopher Charles Peirce is often considered the father of semiotics, so it is not surprising that Peirce's definitions of signs and symbols have dominated modern discussion of these concepts. By 1867, Peirce had developed a triadic structure of elements that were all encompassed under the category of 'signs' (Atkin, 2013). These three elements consisted of the index (associations based in logic: e.g., ☠ to represent poisonous), the icon (associations based in physical resemblance: e.g., ✆ to represent a payphone), and the symbol (associations derived from culture: e.g., $ to represent dollar).

Although other definitions of symbols certainly exist (e.g., Goodman, 1976; Huttenlocher & Higgins, 1978; Saussure, 1916/1998), Peirce's are among the most prevalent in semiotics. Peirce's triadic structure of signs provides a useful framework to conceptualize visual communication, but it is not without its faults. Notably, the boundaries between the three types of signs are not as clear as one might expect: Signs can certainly share attributes of multiple triadic elements. The sign for 'cut' (✂) could be considered an icon because it resembles scissors, but it also could be considered an index because it is used to indicate a cutting action in virtual space (in word processing software) rather than signifying the presence or use of real scissors. Although Peirce acknowledged this potential for sign overlap in his writings, trying to fit signs into neat categories can cause confusion as multiple interpretations often are correct. Further, because words can be considered Peircean symbols, there is understandable confusion between what a philosopher might call a symbol and what a lay person might think a symbol is. Even in the case of words, which many consider to be the original 'symbols,' research on sound symbolism has found that some words (e.g., 'buzz') are more iconic than others (Vinson et al., 2021), suggesting that language is not always arbitrarily or culturally derived (meaning language is not always purely symbolic). Therefore, in this thesis, I have chosen to adopt DeLoache's (2004) definition of a symbol being "something that someone intends to represent something other than itself" (p. 1). This definition is admittedly vague, yet it is useful insofar as it highlights the importance of intent to convey meaning by using a proxy for the real thing. Returning to a point Aristotle raised, a broad definition like that from DeLoache (2004) necessarily includes words as a type of symbol as well. To clarify then, this thesis involves the comparison of 'graphic symbols' (e.g., !@#$%; hereafter referred to simply as 'symbols') to words and pictures.

Symbols can be seen everywhere around us, even in unexpected places: Lines on a highway can be considered symbols insofar as they indicate whether it is appropriate to change lanes. In the present study, I chose to focus on a mix of common graphic symbols (Figure 1) to investigate how symbolic content is represented in memory, allowing parallels to the more established literature surrounding memory for images. Whereas Peirce likely would have claimed that the present study involves a mix of three types of 'signs,' his strict definitions are avoided here to prevent a misunderstanding that the current research concerns memory for physical signs (e.g., traffic signs, billboards) rather than graphic symbols like those found on a keyboard (e.g., !@#$%).

Figure 1

Graphic Symbols Used in Experiments 1-4



Note. In Experiments 1 and 4, only the first 50 of these 80 symbols were used.

1.2 What Do We Know About Symbols?

Despite philosophers having been intrigued by symbols for millennia, the discipline of psychology has been remarkably ambivalent toward them. What little cognitive research does exist on symbols often does not make use of symbols that participants would recognize. For instance, Lupyan and Spivey (2008) found that low-level perceptual processing of symbols could be moderated by top-down feedback from conceptual understandings, but they used entirely novel symbols that participants had to learn during the study. Similarly, Coppens et al. (2011) demonstrated that the testing effect—the improvement in memory from testing rather than restudying (Bjork, 1975; Roediger & Karpicke, 2006)—is generalizable to symbol-word pairs, but the symbols that they used were West African 'Adinkra' characters that were novel to participants. Aggravating this problem, because symbols are culturally specific, some very useful and relevant work has been limited in its generalizability. For instance, Prada et al. (2016) developed the Lisbon Symbols Database which contains normative ratings for over 600 symbols. Unfortunately, many of these symbols, while common in the Portuguese society in which they were normed, are less familiar to participants in other parts of the world.

Ironically, perhaps the most common use of meaningful symbols in psychological research to date has been for the purpose of filler stimuli, as in fMRI experiments (e.g., using '#####' or other symbols as a visual mask). Some of the more thorough work in this area, however, reports findings of neural activation during presentation of these 'filler' stimuli, providing insight into the underlying neural representations of symbols. For instance, Reinke et al. (2008) found that words and symbols both activated the left visual-word form area (VWFA) in the brain, whereas numbers and letters did not. In line with this finding, Kronbichler et al. (2004) demonstrated that symbols are processed in areas associated with perception of abstract visual stimuli.

Carreiras et al. (2014) concluded that processing of letters, numbers, and symbols, while sharing some overlap, largely involve the recruitment of distinct neural regions. For example, they reported that symbols activated unique brain areas—namely, the right superior parietal and the right middle and inferior temporal cortices—different from those activated by letters. These are regions thought to be involved in semantic processing (Binder et al., 2009; Binder & Desai, 2011). In addition, neural activation in the right middle temporal gyrus was higher when symbols, as opposed to numbers or letters, were presented (Carreiras et al., 2014). Importantly, damage to this specific brain region has been implicated in agraphia and alexia for Japanese kanji (logographic characters akin to symbols; Sakurai et al., 2008). The aforementioned fMRI and neurological patient studies suggest that symbols not only hold semantic information similar to words, as might be expected, but that they also evoke neural mechanisms distinct from semantically-void single-character stimuli like letters.

Despite the rarity of psychological research on common symbols, there is an entire sub-field devoted to processing of numbers (most often, Arabic numerals such as 1, 2, 3, etc.). From this work, one can draw parallels to symbols that permit speculation concerning their cognitive processing. Much of the work exploring numbers has argued that there is a universal abstract numeric processing system that handles both quantity, size, and numeric symbols (often localized to parietal and prefrontal cortices; Diester & Nieder, 2007, see Sokolowski et al., 2017 for a recent meta-analysis). The idea is that there is a single system that handles an array of 3 dots, an item that is 3rd largest, and the Arabic numeral ‘3’ itself. Recent work (e.g., Cohen Kadosh et al., 2007, 2008) is, however, starting to question whether symbolic numerals (e.g., ‘3’) are processed separately in the brain from non-symbolic numerical magnitudes (i.e., quantities) and/or non-numerical magnitudes (i.e., physical size). This would suggest that there is a network for the semantics of quantity, but separate areas for interpreting abstract visual stimuli like Arabic numerals. It is possible that symbols are handled by abstract visual processing systems similar to those involved in processing numbers. For instance, in Peircean terms, one might consider '3' to be symbolic, the third largest in size to be indexical, and three dots to be iconic. There are many parallels to be drawn between existing work on numeric cognition and emerging work on symbolic representations in memory.

Research by Lyons and Ansari (2009) indicated that activity in the left parietal cortex was associated with how well participants were able to associate numeric values with novel symbols. Further work suggests that cognitive processing for Arabic numerals (e.g., 1, 2, 3) shifts from frontal to parietal sites later in life, which could be indicative of greater automaticity in the mapping between number and magnitude (Ansari et al., 2005). Similar character-to-meaning mapping processes could be at play for symbols, too. Perhaps as one becomes increasingly familiar with a symbol, the learned association between the perceptual symbol and the concept it refers to becomes increasingly strengthened. I examine this possibility in the present dissertation.

1.3 Are Symbols More Like Images or Words?

Despite symbols having received relatively little attention in psychological research, pictures have benefitted from a rich history of work going back to the late 1800s (e.g., Bergstrom, 1893). A significant body of work emerged in the 1960s to suggest that pictures (i.e., graphics representing concrete things) are often better remembered than words. This phenomenon—the picture superiority effect (PSE)—has now been extensively studied for almost six decades. Early on, Allan Paivio (1969, 1971) proposed his influential dual coding theory—that pictures are remembered better than words because pictures often are represented by two distinct codes—verbal and image—whereas words tend to evoke only the verbal code.

Because symbols are often used alongside text, is it possible that they are processed like words? This could be. But symbols are not made up of meaningless graphemes (i.e., letters) as words are. Instead, symbols are more holistic and unique in their formation, and carry a meaning of their own, more like images (a term hereafter used interchangeably with 'pictures'). Therefore, when conceptualizing symbols within the dual coding framework designed for pictures, one could argue that symbols may also elicit an image code in memory as a result of their holistic, non-verbal nature.

Paivio's dual coding theory is also often used to explain why concreteness effects occur in memory: The referents of concrete words are more easily imaged than are those of abstract words, making concrete words better remembered as they are more likely to elicit spontaneous imagery (Ding et al., 2017; Fliessbach et al., 2006; Hamilton & Rajaram, 2001; Jessen et al., 2000; Khanna & Cortese, 2021; Paivio et al., 1994; Roberts & Wammes, 2021). Therefore, according to dual coding theory, any word that is imaged (either in the physical world or via imagination) necessarily becomes 'concrete' insofar as it affords the stimulus an image code in memory.

Consider an example. Dual coding theory would predict that the word 'apple' will be better remembered than the word 'peace' but that an image representing peace, such as a picture of white doves or of olive branches, should be remembered just as well as a picture of an apple. Recent work from our laboratory has found evidence consistent with this prediction: Drawing a picture related to an abstract word brings memory performance for that word up to the level of written concrete words (Roberts & Wammes, 2021). In that study, when asked to draw the referents of certain abstract words (e.g., ‘love’), many participants chose to draw common symbols (e.g., ♥), rather than more elaborate scenes, to depict the concepts.

This pattern of findings was a direct motivator for the current study. Is it possible that drawing a symbol could serve as an effective way to represent an abstract concept without picturing something else that is already concrete? It seems plausible: Drawing ‘☮’ should serve to concretize the word ‘peace’ as well as olive branches or white doves but should simultaneously avoid the confound of using other related concrete things to depict the abstract concept. That is, representing an abstract idea with a graphic symbol (e.g., ☮) provides the participant with a pictorial (and therefore concrete) representation of that abstract concept, likely eliciting a secondary image code in memory to accompany the verbal code of the word itself (e.g., ‘peace’). This logic formed the basis of my principal prediction: As a result of their dual codes in memory, symbols and images should both be better remembered than words but memory for symbols and images should not differ.

Related work from our lab investigating memory for emojis points to a similar dual coding mechanism. Using a divided-attention technique, Homann, Roberts, et al. (2022) showed that recall of words was attenuated when a concurrent 1-back task demanded processing of verbal (words) or verbal and image (emojis) stimuli, but not when the 1-back task involved pure image stimuli (stars with different fill patterns). Recall of emojis, on the other hand, was diminished when attention was divided using a 1-back task with pure verbal (words), pure image (stars with different fill patterns), and verbal and image (emojis) stimuli. Therefore, while recall of words can be influenced by concurrent tasks drawing on verbal processing mechanisms, recall of emojis is impeded by both concurrent verbal and visual processing demands. The major conclusion is that emojis are likely encoded and subsequently retrieved by use of dual codes in memory, much as I suggest is the case for symbols.

Of course, other theories have been put forth to explain picture superiority in memory, most of them rooted in some variant of a distinctiveness account (Hunt & McDaniel, 1993). The typical definition of distinctiveness used in these accounts follows that of Hunt (2013, p. 1): Distinctiveness is “the processing of difference in the context of similarity.” A conceptual distinctiveness account contends that pictures are conceptually more distinctive than words because they provide semantic characteristics that are diagnostic of a given item (e.g., large eyes are diagnostic of an owl because they are uncommon among other birds; Hamilton & Geraci, 2006; Nelson et al., 1977). In contrast, physical distinctiveness accounts argue that pictures vary more in visual appearance than words (since words, at least in the English language, are made up of the same 26 recycled letters; Ensor et al., 2019b; Mintzer & Snodgrass, 1999). It is thought that items that are distinct among other items during encoding will yield more precise specifications of the item or encoding event (Hunt, 2013). This distinctive information is then used as a heuristic during memory retrieval: Participants attempt to remember the unique, diagnostic features of an item to aid in recall, or they compare those features against a test image on a recognition test.

When applied in the present case, these physical and conceptual distinctiveness accounts should also predict better memory for symbols than for words. Relative to words, symbols are more physically distinct: They consist of varying lines, shapes, orientations, and degrees of symmetry. Furthermore, whereas words in the English language use the same 26 letters repeatedly, there are many more than 26 unique symbols, most of which share few visual attributes. In addition, symbols are also conceptually distinct insofar as they have hardly any semantic neighbours. For example, whereas the word 'play' (in the context of watching television) has many related words, such as 'begin,' 'start,' 'commence,' etc., it is further crowded in semantic space due to its nature as a homonym (e.g., 'play' can also refer to games or to theatrical performances). The symbol for 'play' (▶), on the other hand, stands in relative semantic isolation from other symbols; no other symbol that conveys a similar meaning comes to mind. Therefore, the symbol for play would be considered to bear high conceptual distinctiveness (Hamilton & Geraci, 2006), whereas its word counterpart would not. Importantly, however, more recent work (Ensor et al., 2019b) has suggested that distinctiveness and dual coding may each contribute to picture superiority depending on the type of retrieval test that is employed, an idea upon which I expand in the General Discussion.

1.4 What Other Factors Might Contribute to Symbol Processing?

A graphic symbol—like any other sign—must be known to the interpreter before it can convey meaning. Therefore, it is likely that memory for a symbol would be moderated by one's familiarity with it. That is, if one did not know that '%' means 'percentage,' dual coding theory would suggest that the symbol would fail to elicit a verbal code in memory. Rather, the symbol would simply appear as a meaningless shape and therefore would be less likely to be retrieved on a memory test. It is possible, however, that one could generate a verbal label or descriptor for the unknown symbol (e.g., % = 'two circles with a diagonal line between them'), and this could boost memory for it via elaboration, but any benefit of doing so would be lost in a context in which the meaning of the symbol must be known to permit interpretation of the information being conveyed (e.g., in sentences). At the same time, abstract words that are highly familiar (e.g., 'love') may more easily elicit image codes in memory, thus making them more like concrete words (e.g., 'apple'). Therefore, it stands to reason that, if symbols are indeed dual coded and consequently lead to better memory than words, this performance discrepancy ought to be attenuated when one's familiarity with a symbol is so low that it cannot be labelled.

The predictions made thus far for symbol superiority in memory have been rooted in Paivio's dual coding theory. But what if the physical form of the symbol is the true driver of differential retention, as proposed by the physical distinctiveness account of picture superiority? One way to investigate this question would be to turn to the emerging area of intrinsic stimulus memorability. Memorability—the likelihood that something will later be remembered—has recently been explored as a perceptual attribute of a stimulus, a quality that is independent of stimulus type, aesthetics, emotionality, priming, and attention (Bainbridge, 2019, 2020, 2021, 2022; Bainbridge et al., 2013, 2017; Bainbridge & Rissman, 2018; Brady & Bainbridge, 2022; Goetschalckx et al., 2018; Isola et al., 2011, 2014; Xie et al., 2020). It could therefore be the case that, relative to words, symbols simply are unique in their physical and/or conceptual forms, and it is this property that underlies improved memory performance. These possibilities and others are examined in the present study.

1.5 The Present Investigation

Six experiments were conducted, all aimed at addressing how graphic symbols are represented in and retrieved from memory. It could be that symbols, being unitary characters not comprised of graphemes (unlike words formed from letters), are processed holistically like pictures. Moreover, because symbols are most often used to represent abstract concepts, they could serve to ‘concretize’ those ideas by providing quick and easy-to-recognize visual referents. Such a visual referent could aid memory by providing an image record (along the lines of Paivio’s dual coding theory) or by offering a distinct visual form that stands out in memory. In either case, memory for symbols should be superior to memory for words but should not differ significantly from memory for pictures.

In Experiment 1, participants were presented with symbols or their word counterparts (i.e., $ or ‘dollar’; see Appendix E) and were later tested on their memory for the studied items. Experiment 2 conceptually replicated the first study with a different experimental design. Experiment 3 addressed the potential confound of differing set sizes between words and symbols by reducing words to a single category. Experiment 4 compared memory for symbols, pictures, and words directly in a single investigation. It is also noteworthy that in Experiments 1 and 4 testing was via recall, whereas in Experiments 2 and 3 testing was via recognition. This also permitted consideration of the relevance of type of retrieval test to the underlying mechanism.

A fifth experiment completed in two parts investigated whether the findings found thus far would generalize to logos. Participants were asked to study a series of sports logos devoid of integrated text, sports logos with integrated text (e.g., the team's home city), or sports team names presented in a normal font. Afterward, recognition memory was tested by presenting either the same stimuli as were encoded originally, or sports team names only. Importantly, a sports familiarity questionnaire allowed for examination of how memory for an abstract concept (a sports team) represented by a concrete visual referent (a logo) can be modified by one's personal familiarity.

In Experiment 6, familiarity and frequency ratings were collected for the set of symbols used in the first four experiments, the goal being to determine whether general familiarity with a symbol would predict later memory for it. The ResMem neural network (Needell & Bainbridge, 2022) was also employed to assess whether symbols are better remembered than words because of their inherent memorability. This set of studies began with a straightforward initial experiment examining whether memory differs for symbols and their word counterparts.

Chapter 2 – Probing Memory for Everyday Symbols

2.1 Experiment 1 – Symbols vs. Words[1]

2.1.1 Introduction

Recall that the major prediction was that symbols are effectively processed as pictures insofar as they serve to concretize abstract concepts by providing a visual referent. That is, according to Paivio's dual coding theory, reading of abstract words is unlikely to evoke spontaneous mental imagery. By providing the participant with a picture-like visual stimulus rather than a word, however, an image code will be provided to the participant in lieu of their own mental imagery. The principal prediction for this first experiment was that if symbols are in fact 'mini-pictures,' and therefore are likely remembered like other images, then they should be better remembered than words as a result of dual coding and/or enhanced distinctiveness. To test this basic prediction, a simple two-block study was designed using nonoverlapping stimulus subsets in the two blocks. Participants were presented with symbols (e.g., '$') in one block and with words (e.g., 'dollar') in the other block. After a short delay, they were asked to recall the studied items. For the test of the word study block, participants were told to write down the words that they remembered; for the test of the symbol study block, they were asked to draw the symbols that they remembered.

---

[1] The work presented in Experiments 1-4 as well as in Experiment 6 is in press:

Roberts, B. R. T., MacLeod, C. M., & Fernandes, M. A. (in press). Symbol superiority: Why $ is better remembered than 'dollar'. *Cognition*. https://doi.org/10.1016/j.cognition.2023.105435

This two-block design was chosen to preserve the inherent encoding features of words and symbols and to minimize any ‘conversion’ of studied symbols into verbal labels (or vice versa) during study or during the memory test. Memory was tested by recall to ensure that when participants remembered a symbol but did not know its verbal label, they were still able to provide evidence of remembering it. Finally, to assess whether familiarity might moderate memory for the current set of symbols, I administered a matching task in which participants had to decide whether simultaneously presented symbols and words had corresponding meanings. This is the first experiment to have tested memory for common everyday graphic symbols and their word counterparts.

2.1.2 Method

Participants. An a priori power analysis was not conducted because there had been no previously published work documenting memory for symbols. Instead, data were collected from a minimum of 50 participants, continuing until the stopping rule date (the end of the academic term). A total of 155 University of Waterloo undergraduate students took part in exchange for course credit. Participants had self-selected to participate in the study and had self-reported normal or corrected-to-normal vision as well as having learned English before the age of 9.

From this initial sample, participants’ data were removed if they were found not to have followed the experiment instructions correctly (e.g., if they wrote down items during encoding or wrote out labels of symbols during recall) or had corrupt data files (n = 24). Then, *R* (v. 4.1.1; R Core Team, 2020) statistical software was used to exclude participants whose performance was ±3 *SD*s from the mean for recall of symbols ($n = 0$), recall of words ($n = 1$), or accuracy on the symbol matching task ($n = 2$)[2]. The final sample of 128 participants used in the statistical analyses was 71.88% female, with ages ranging from 17 to 31 ($M = 19.7$, $SD = 2.1$).

Materials. Two sources provided a set of common everyday symbols: (1) the Wikipedia entry for "Miscellaneous Symbols" (https://en.wikipedia.org/wiki/Miscellaneous_Symbols), and (2) aggregate response data from several different pages on Sporcle, an online quiz-sharing website (https://www.sporcle.com/). Sporcle hosts many user-generated quizzes; in the instances I gathered stimuli from, participants had to match symbols to their definitions. Conveniently, the aggregate response data for all of the attempts made on any Sporcle quiz are freely available to use; four different Sporcle quizzes were found that assessed knowledge of common symbols. The aggregate data from these quizzes were from up to 216,000 plays. Based on the results of these quizzes, symbols that were poorly known were discarded.

In the end, a list of 101 symbols was assembled. For Experiment 1, the best 50 symbols were selected based on the preferences that symbols were (1) physically and conceptually distinct from one another, and (2) common in everyday life for the target population of Western society undergraduate students. All symbol stimuli were created by copying each item from the web using the Segoe UI Symbol font, then editing it to be 110 X 116 pixels in greyscale on a white background (Figure 1).

[2] Retaining data from these participants in the statistical analyses made no difference to the pattern of results.

All stimuli in the words block were single-word versions of the symbols on the master stimuli list (e.g., $ = 'dollar', % = 'percentage'). All word stimuli were presented at 5% of the total screen height, in black lowercase Calibri font, centered on a white background.

Procedure. The experiment was administered to groups of six participants at a time, each using a separate Windows computer, separated by wall panel dividers. Following informed consent and demographic data collection, each participant was randomly assigned to one of four counterbalanced conditions (the product of the two stimulus subsets crossed with the two block orders)[3] before being seated at a testing desk. Each experiment computer was running PsychoPy (v. 3.2.4; Peirce et al., 2019) experiment builder software, outputting to a 24-inch monitor with 1920 x 1080p resolution (60 Hz). Participants sat roughly 24" to 30" from the computer monitor.

Prior to the study phase, participants were told that they would see either words or symbols presented one at a time on the screen and were instructed to try to remember as many as they could for a later memory test. Participants were then presented either with 25 words or with 25 symbols sequentially in the center of the screen. Each stimulus was shown for 2 s, followed by a blank screen for 250 ms, a fixation dot for 500 ms, and finally another blank screen for 250 ms. Encoding stimuli were presented to each participant in a unique random order.

[3] Due to a programming error, stimuli were not properly counterbalanced across participants. In Experiment 1 only, the items in the symbols list were always presented as symbols and the items in the words list were always presented as words.

After all of the 25 stimuli for the block had been presented, participants completed a filled delay task. They were instructed to listen to a tone and then to respond by pressing '1', '2', or '3' if the pitch of the tone was 'low,' 'medium,' or 'high,' respectively. Examples of each pitch were provided in the task instructions. Tones were played for 500 ms, with a new tone played every 2 s or when a response was made, whichever came first. This task persisted for two minutes and was included to guard against potential ceiling effects by eliminating recency and by minimizing post-list rehearsal. Following the filled delay, participants were given two minutes to complete the free recall memory test for the items seen during the study phase. Participants were instructed to recall items 'as presented': Word-based items were to be written whereas symbols were to be drawn.

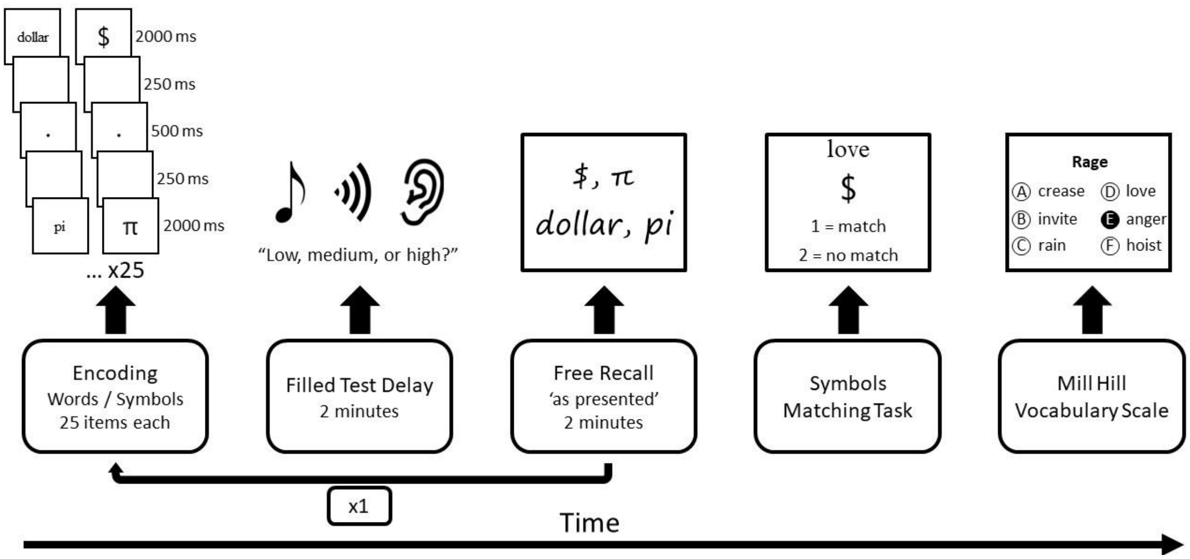
Finally, following the two study-test cycles, a 'symbols matching task' was administered to assess the participants' familiarity with the symbols. In this task, participants saw 25 symbol-word pairs that matched (e.g., '$ - dollar') and 25 that mismatched (e.g., '& - love') in a random order. The matching of symbols to words was counterbalanced across participants such that any particular symbol matched the presented word for half of participants. Participants were instructed to respond by pressing the '1' key if the symbol and word matched or the '2' key if they did not match. After each response, a new pair of stimuli immediately appeared on the screen (i.e., this task was self-paced).

Following the matching task, a paper version of the Mill Hill Vocabulary Scale (Set A; Raven, 1958) was administered to assess participants' English language competency. In this task, participants must select a synonym for a target word from a group of six possible choices. Afterwards, a feedback letter was provided that detailed the purpose of the study.

The procedures and materials for this study were approved by the Office of Research Ethics at the University of Waterloo (project #41594). All data, analysis code, experiment programs, and other materials are listed on the Open Science Framework (OSF; https://osf.io/e53z4/).

Figure 2

Experiment 1: Procedure



Note. 25 items (words or symbols) were presented at encoding in each of the two study-test cycles. Participants were never shown the same item in both symbol (e.g., $) and word (e.g., dollar) format.

2.1.3 Results

Memory Performance. I alone scored recall of symbols by determining whether the symbols drawn during recall had or had not appeared on the study list. Symbols written out with words were considered incorrect in this experiment because (1) it was desirable to avoid the possibility of confounds brought on by 'conversion' from symbol format to word format, and (2) writing symbols in word format was contrary to the experiment instructions and was therefore rare among participants (seven participants recalled all symbols by writing out their respective labels and were therefore not included in subsequent analyses)[4]. Word recall was scored using both lenient (synonyms acceptable) and strict (exact matches only) scoring methods before being converted into proportions by dividing by 25. To ensure the most conservative test of the predictions, recall of symbols was compared to recall of words using the lenient scoring criteria for the latter condition throughout the results reported here[5].

Critically, the results of a paired-samples *t*-test demonstrated that the proportion of symbols recalled was greater than the proportion of words recalled (Figure 3)[6], $t(127) = 8.75$, $p < .001$, $d = -0.77$, $CI_{95}[-0.97, -0.58]$, $BF_{10} = 6.10e+11$, with extreme Bayesian evidence in support of the alternative model.[7]

Exploratory Correlations. To investigate the possibility that one's knowledge of symbols, or of the English language, might enhance recall performance for symbols or words, a series of exploratory Pearson correlations was conducted. Mill Hill Vocabulary Scale scores were weakly correlated with the proportion of words recalled, $r(126) = .17$, p =

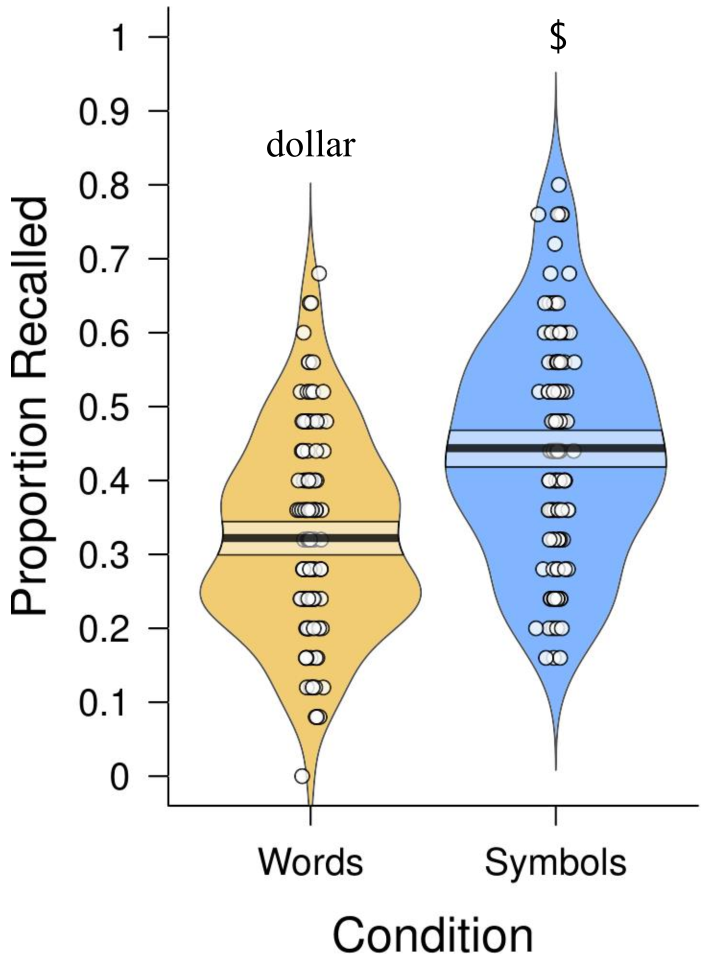
---

[4] This data trimming procedure allows for the most conservative measure of symbol recall performance, as including the trimmed responses should only strengthen the memory performance benefit seen for symbols.
[5] Using a strict scoring criterion for word recall instead yielded no change in the pattern of results: The difference between words and symbols simply became larger, $t(127) = 9.36$, $p < .001$, $d = -0.83$, $CI_{95}[-1.03, -0.63]$, $BF_{10} = 1.61e+13$.
[6] This analysis was also conducted with the pre-registered minimum sample size of 50 participants by randomly selecting data from the full data set. The pattern of results was identical.
[7] Throughout this dissertation, Bayes factors were calculated using the *BayesFactor* (Morey et al., 2011) package for *R*, enlisting a default Jeffreys-Zellner-Siow (JZS) prior with a Cauchy distribution (center = 0, r = 0.707). This package compares the fit of various linear models. In the present case, Bayes factors for the alternative (BF_{10}) are in comparison to intercept-only null models. Bayes factor interpretations follow the conventions of Lee and Wagenmakers (2013). Bayes factors in favor of the alternative (BF_{10}) or null (BF_{01}) models are presented in accordance with each preceding report of NHST analyses (i.e., based on a $p < .05$ criterion) such that $BF > 1$.

.062, but not with proportion of symbols recalled, $r(126) = .08$, $p = .363$. Performance on the symbols matching task was not significantly correlated with recall performance in either condition ($rs \leq |.14|$, $ps \geq .126$).

Figure 3

Experiment 1: Recall of Words and of Symbols



Note. Lenient scoring for word recall is presented. Errors bars = 95% confidence intervals. Examples of encoding stimuli for each trial type are provided above their data points.

2.1.4 Discussion

In Experiment 1, a basic hypothesis was assessed—that symbols would be better remembered than their word counterparts. This hypothesis was derived from the dual coding account of how pictures are represented in memory (Paivio, 1969, 1971). This prediction clearly was borne out: Symbols were indeed better remembered than their word counterparts. In addition, because accuracy on a symbol matching task did not correlate with memory performance for symbols or words, there was no support for the prediction that familiarity would moderate memory performance. However, there were two limitations to this experiment. First, as noted in Footnote 1, due to a programming error, items from the master word list were not properly counterbalanced into the symbols and words groups. Second, and more critically, because participants were asked to *draw* symbols and *write* words during the recall test, it is possible that the superior memory for symbols observed here was due to some drawing-related boost, along the lines of memory benefits known to operate following drawing at encoding (i.e., the drawing effect; see Wammes et al., 2016). The goals of Experiment 2 were to address these two limitations and to expand the generalizability of the findings by switching the setting, study design, and type of retrieval test.

2.2 Experiment 2 – Conceptual Replication and Extension

2.2.1 Introduction

Experiment 1 provided evidence that memory is indeed better for symbols than for words. There were, however, limitations to the experiment. To address these, Experiment 2 switched to a between-subjects design and used a recognition test rather than a free recall test to assess memory. The switch to recognition testing was intended to ensure that any difference in memory between symbols and words was not due to having to draw (the symbols) at retrieval. The switch to a between-subjects design was intended to ensure that participants did not have better memory for symbols simply because they saw these as more unusual or important compared to words when both types of stimuli were presented.

In Experiment 2, participants were randomly assigned to one of four groups in a 2 X 2 design, the factors being stimulus type at encoding and stimulus type at test. As a result, stimuli seen on the recognition test were either congruent (e.g., $ → $; dollar → dollar), or they were incongruent (e.g., $ → dollar; dollar → $) with those seen at encoding. The first prediction was that a pure symbols-symbols (encoding-retrieval) condition would still lead to better memory than a pure words-words condition because the hypothesized memory difference occurs due to better encoding of symbols than of words.

In addition, I predicted that, based on the assumption that dual codes are elicited for symbols at encoding (like images), participants who studied symbols would demonstrate superior memory even when the recognition test consisted entirely of words. Of course, some reasonable attenuation in performance was expected as a result of the incongruency of stimulus types between the encoding and retrieval phases. In essence, all four groups were expected to be significantly different from each other: Both congruent groups should exhibit memory performance superior to that of the two incongruent groups due to transfer appropriate processing (Morris et al., 1977), or encoding specificity in the former case (Tulving & Thomson, 1973). Nonetheless, groups that encoded symbols should display better memory performance regardless of the format in which the items were presented on the recognition test.

2.2.2 Method

Participants. An a priori power analysis was conducted using G*Power software (v. 3.1.9.7; Faul et al., 2007), targeting a medium-sized between-subjects contrast ($d = 0.50$, $\alpha = .05$, two-tailed). This indicated required sample sizes of 64 participants in each of the four groups to achieve 80% statistical power. Accordingly, a goal was set to collect a minimum of 64 participants per group ($N = 256$ total), with the ideal target sample size set higher at 70 participants per group ($N = 280$ total). These sample sizes should also have provided 98% and 99% power to detect medium sized ($f = 0.25$, $\alpha = .05$, two-tailed) main effects and interactions as well, respectively. In the end, data were collected from 363 University of Waterloo undergraduate students who each took part in a single session in exchange for course credit. Participants had self-selected to participate in the study and had self-reported normal or corrected-to-normal vision as well as having learned English before the age of 9.

From this initial sample, participants' data were removed in sequential steps if they (1) had corrupted or incomplete data files, or were duplicate attempts ($n = 34$), (2) took less than 5 minutes to complete the study ($n = 1$), (3) took more than 30 minutes to complete the study ($n = 13$), or (4) were ±3 *SDs* away from the mean of remaining participants for study duration ($n = 0$). Therefore, 315 valid data files entered statistical analyses. Then, *R* statistical software was used to exclude participants in successive steps who were missing more than ten recognition responses (due to rushing through the study causing technical errors; $n = 28$), as well as those whose performance was ±3 *SD*s away from the mean in their group on any memory performance metric (hits, false alarms, or accuracy; $n = 4$) or on the symbol matching task ($n = 10$)[8]. The final sample of 273 participants used in the statistical analyses was 78.39% female (Table 1), with ages ranging from 18 to 48 ($M = 20.1$, $SD = 2.8$; one participant declined to provide their age).

Materials. The same materials as used in Experiment 1 were used here, except that the master stimulus list was expanded to include 80 symbols and their word counterparts (Figure 1). Five of the newly added symbols required two words to be used in the words condition (e.g., ⏩ = 'fast-forward'). Also, word-based encoding stimuli were now presented in Times New Roman size 24 lowercase black font, centered on a white background.

Procedure. The procedure generally followed that of Experiment 1, except that now—due to the COVID-19 pandemic—the study took place online. The experiment was built using Qualtrics software (https://www.qualtrics.com/) and was administered through Prolific (an online data collection platform; https://www.prolific.co/). As a result, the study was administered on participants' personal computers. This change in setting led to four minor procedural updates: (1) demographic data were collected at the end of the study (rather than at the beginning), (2) the tone classification filler task now involved using a mouse to click on a response of 'low,' 'medium,' or 'high' for each tone instead of responding using keypresses, (3) participants now responded with the 'n' and 'm' keys on the symbol

[8] Retaining data from these participants in the statistical analyses made no difference to the pattern of results.

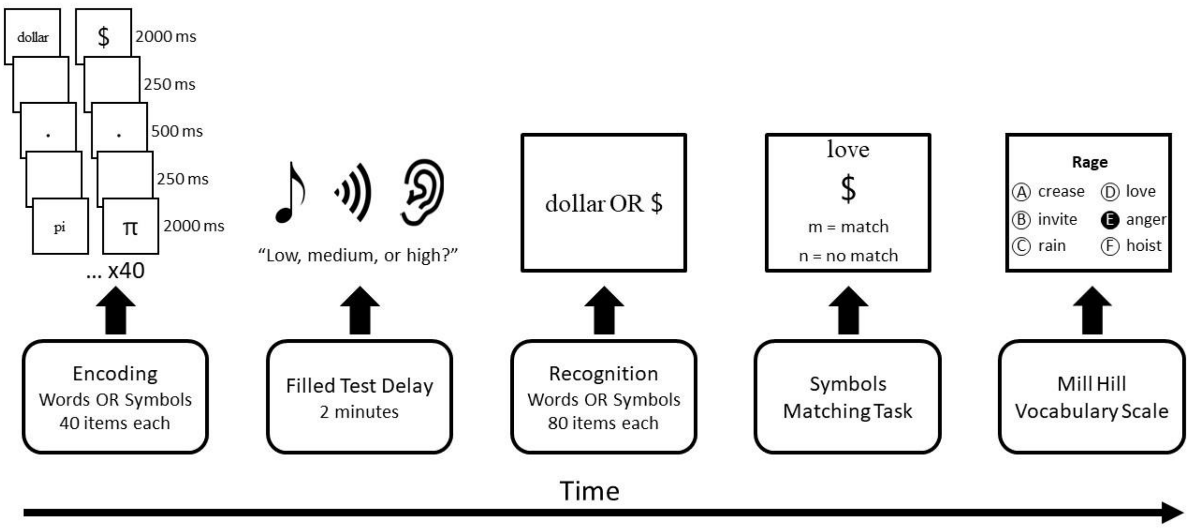
matching task (rather than the '1' and '2' keys), and (4) an electronic version of the Mill Hill Vocabulary Scale was used (rather than a paper copy).

In addition, to test the robustness of the findings from Experiment 1, Experiment 2 switched to a between-subjects design with an old/new recognition memory test. Because of the change to a 2 X 2 between-subjects design, participants were randomly sorted into one of four groups. In the Symbols-Symbols group participants studied and then were tested using symbol stimuli only at encoding and retrieval (e.g., $ → $), in the Symbols-Words group the test stimuli switched to words (e.g., $ → dollar), in the Words-Symbols group participants studied words and then were tested using symbols (e.g., dollar → $), and in the Words-Words group both the study and test stimuli were words (e.g., dollar → dollar). During the study phase, participants in the Symbols-Words and Words-Symbols conditions were unaware of the upcoming switch in stimulus format on the retrieval test. Before the test phase began, the participants in these two groups were told to discriminate between old and new items regardless of the stimulus format: "Items will be presented in either word or symbol format. For items you remember being presented in the study phase (regardless of format), press 'M'. For items you do NOT remember being presented in the study phase (regardless of format), press 'N'."

In each group, 40 items were randomly selected from the 80-item master stimulus list to be presented at encoding. All 80 items were included on the recognition test (40 targets plus the remaining 40 items that served as lures). Participants responded with the 'n' key if the item was 'new' (i.e., not studied previously), or the 'm' key if the item was 'old' (i.e., they remembered it from the study phase). The procedures and materials for this study were approved by the Office of Research Ethics at the University of Waterloo (project #41594). All data, pre-registrations, analysis code, experiment programs, and other materials are listed on the Open Science Framework (OSF; https://osf.io/e53z4/).

Figure 4

Experiment 2: Procedure



Note. 40 items (words or symbols) were presented at encoding to each participant, depending on condition assignment. Recognition test stimuli were either congruent (e.g., $ → $) or incongruent (e.g., $ → dollar) with encoding stimuli, depending on condition assignment.

2.2.3 Results

Memory Performance. To examine memory performance across groups, a one-way independent-groups analysis of variance (ANOVA) was conducted using the rstatix (v. 0.7.0; Kassambara, 2021) package for *R*, with Group as the independent variable with four levels (Symbols-Symbols, Symbols-Words, Words-Symbols, and Words-Words) and memory accuracy (hit rate minus false alarm rate) as the dependent measure. This analysis revealed a significant main effect of Group, $F(3, 269) = 56.31$, $p < .001$, $\eta_p^2 = .39$, $BF_{10} = 1.21e+25$. Welch-adjusted[9] pairwise comparisons with Holm corrections (Holm, 1979) revealed that all between-group contrasts were statistically significant ($ps \leq .011$, $ds \geq 0.44$; Table 1)[10].

Table 1

Descriptive Statistics for All Memory Performance Metrics as a Function of Experiment and Condition for Experiments 1-4

Experiment	**Condition**	***n***	**% Female**	**Age**		**Proportion Recalled**		**Accuracy**		**Hit Rate**		**False Alarm Rate**	
				M	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***
1	Words	128	71.88%	19.70	2.10	0.32	0.13						
	Symbols					0.44	0.15						
2	Words-Words	68	76.47%	19.87	1.65			0.56	0.22	0.72	0.13	0.17	0.15
	Words-Symbols	69	78.26%	20.16	1.95			0.24	0.16	0.47	0.18	0.23	0.15
	Symbols-Words	67	76.11%	19.54	2.15			0.36	0.20	0.64	0.15	0.28	0.13
	Symbols-Symbols	69	82.61%	20.65	4.36			0.66	0.24	0.84	0.13	0.18	0.16
3	Words	92	77.17%	20.77	4.53			0.59	0.21	0.83	0.12	0.24	0.15
	Symbols	97	76.29%	20.35	3.02			0.68	0.22	0.82	0.14	0.14	0.13
4	Words	218	50.46%	33.40	11.80	0.41	0.20						
	Symbols					0.47	0.17						
	Pictures					0.46	0.17						

Note. The demographic figures in Experiments 1 and 4 represent a single group of participants that saw each condition.

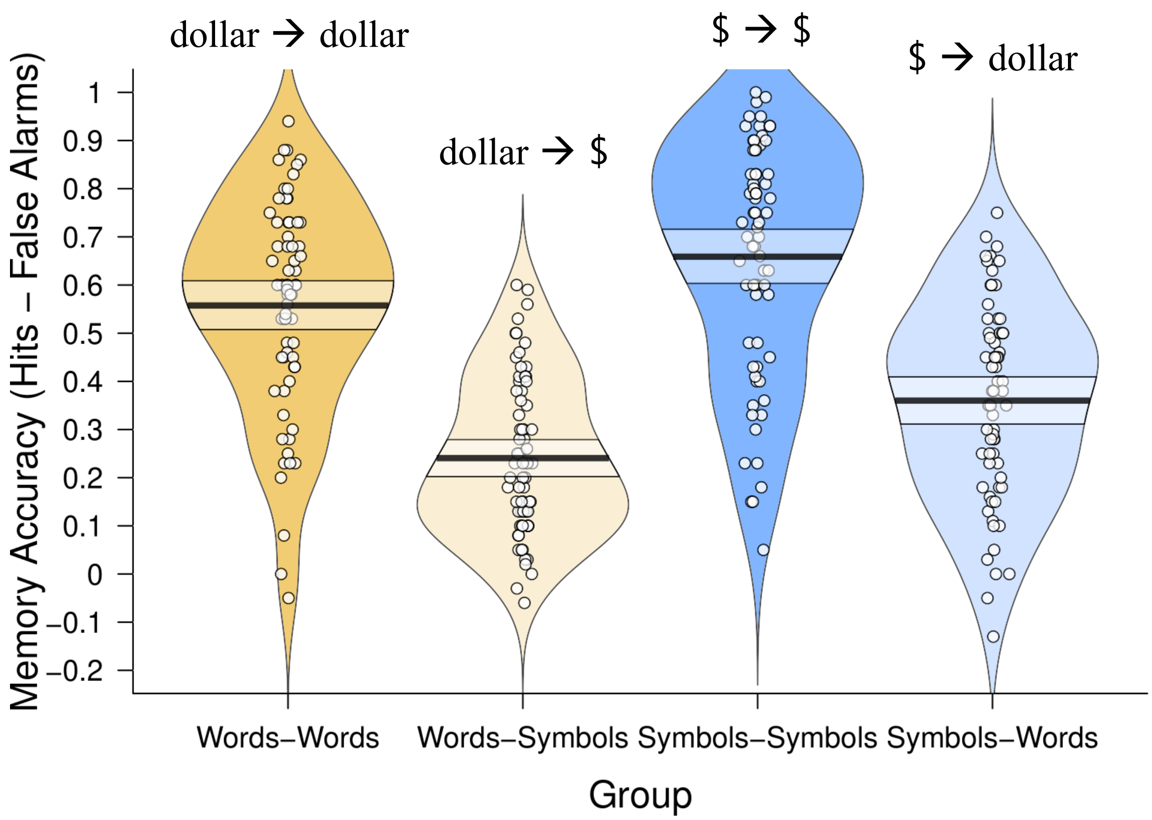
To examine whether there were differences in encoding or retrieval of symbols and words, two separate Welch-adjusted independent-samples *t*-tests were conducted, each with Holm corrections. First, the effects of the two stimulus sets at encoding were tested by contrasting participants who studied symbols (Symbols-Symbols combined with Symbols-Words) versus those who studied words (Words-Words combined with Words-Symbols). This analysis demonstrated that studying symbols on average led to superior memory than

[9] Welch-corrected *t*-tests are still valid and recommended as a default method for between-subjects comparisons (Delacre et al., 2017, 2020; Ruxton, 2006), even when homogeneity of variance is not violated.
[10] This analysis was also conducted with the pre-registered target sample size of 64 participants per group by randomly selecting data from the full data set. The pattern of results was identical.

did studying words (Figure 5), $t(269.01) = 3.63$, $p < .001$, $d = 0.44$, $CI_{95}[0.20, 0.68]$, $BF_{10} = 62.07$, with very strong Bayesian evidence for the alternative model. Next, the effects of the two stimulus sets at retrieval were evaluated by contrasting those tested with symbols (Symbols-Symbols combined with Words-Symbols) versus those tested with words (Words-Words combined with Symbols-Words). This analysis showed no significant difference in test stimuli format between symbols and words, $t(259.82) = -0.31$, $p = .761$, $d = -0.04$, $CI_{95}[-0.27, 0.20]$, $BF_{01} = 7.20$, with moderate Bayesian evidence in support of the null model.

Figure 5

Experiment 2: Recognition in Each of the Four Study-Test Groups



Note. Errors bars = 95% confidence intervals. For each group, stimuli examples from the encoding phase are presented before the arrow, and stimuli examples from the test phase are presented after the arrow.

Correlations of Memory with English Proficiency and Symbol Familiarity. Once again, a series of exploratory Pearson correlations was conducted to determine whether English language proficiency or symbol knowledge correlated with memory performance. The results of these analyses revealed that performance on the Mill Hill Vocabulary Scale was significantly correlated with memory accuracy in each group ($rs \geq .29$, $ps \leq .014$), except for the Symbols-Words group ($r(65) = .19$, $p = .134$). This effect seemed to be driven by a reduction in false alarms rather than by an increase in hits (Table 1), as the former measure correlated negatively with Mill Hill score in each group ($rs \leq -.26$, $ps \leq .031$), except for the Symbols-Symbols group ($r(67) = .22$, $p = .071$). Hit rate, on the other hand, did not correlate with Mill Hill score ($rs \leq |.14|$, $ps \geq .255$), except for the Symbols-Symbols group ($r(67) = .29$, $p = .012$).

Further analyses demonstrated that performance on the symbols matching task correlated significantly with memory accuracy in the symbols encoding groups (Symbols-Symbols, $r(67) = .39$, $p < .001$, and Symbols-Words, $r(65) = .35$, $p = .004$), but not in the words encoding groups (Words-Words, $r(66) = .17$, $p = .159$, and Words-Symbols, $r(67) = .22$, $p = .066$). Like the correlations conducted with Mill Hill performance, this pattern seemed to be driven by a reduction in false alarms rather than by an increase in hits (Table 1), as the former measure correlated negatively with symbols matching task performance in each group ($rs \leq -.25$, $ps \leq .038$). On the other hand, it only correlated with hit rate in the Symbols-Symbols group, $r(67) = .34$, $p = .004$ (remaining $rs \leq |.19|$, $ps \geq .115$).

2.2.4 Discussion

Experiment 2 had three primary goals. The first was to determine whether the results of Experiment 1 (i.e., symbols > words) could be conceptually replicated. The second was to ascertain whether any observed effect would be driven by stimulus format at encoding rather than at retrieval. The third was to explore whether there would be a transfer-appropriate processing benefit when stimulus format between encoding and retrieval was congruent as opposed to incongruent. To address these goals, four groups were created that each saw a unique combination of symbol and word stimuli presented at encoding and retrieval.

The predictions were supported by the data. The findings from Experiment 1 were replicated, plus there was an overall benefit to memory when stimuli were congruent relative to incongruent between encoding and retrieval. That is, a pure symbols-symbols condition led to superior memory performance compared to a pure words-words condition. In addition, the memory benefit for symbols over words clearly appeared to be driven by processes at encoding rather than at retrieval. Grouping together types of materials based on what was presented during encoding (words or symbols) showed that, overall, studying symbols led to better memory than studying words, regardless of the stimuli type shown on the later recognition test. In contrast, memory was unaffected by stimulus format on the recognition test (being tested on symbols or words made no difference). The results of this experiment therefore indicate that the effect of 'symbol superiority' is not only generalizable to between-subjects designs and to recognition testing, but that it also occurs primarily due to differences at encoding rather than at retrieval.

From the outset, a cost to memory performance was predicted to result from switching stimulus format between encoding and retrieval, based on the well-known concept of transfer appropriate processing (Morris et al., 1977). This predicted cost manifested as expected: Engaging in imagery-based (or word-based) processing at encoding benefitted memory on tests that also used imagery (or words) again at retrieval. More pointedly, this cost to memory also aligns with earlier picture superiority work (Mintzer & Snodgrass, 1999; Stenberg et al., 1995) which demonstrated significant costs when switching between words and pictures.

In this experiment, positive correlations were observed between English language proficiency (as measured by the Mill Hill Vocabulary Scale) and memory performance in each group. It is possible that familiarity with the English language tracks with immersion in North American culture. Therefore, insofar as the symbols used here are prominent in North American societies, it makes sense that greater familiarity with the dominant North American language would also indicate greater familiarity with the current set of symbols and words, which could serve to enhance memory. It is worth noting, however, that the significant correlations between Mill Hill score and memory performance observed in Experiment 2 were non-significant in Experiment 1, possibly speaking to the poor reliability of these particular observations.

Critically, the idea that familiarity with symbols could in part be driving superior memory performance was also supported in the second set of correlations showing that, whereas performance on the symbol matching task correlated positively with performance for groups that studied symbols, it had no relation to performance for groups that studied words. Most pertinent, hit rate in the ‘pure’ Symbols-Symbols group was positively related to performance on the symbol matching task whereas false alarm rate was negatively correlated with the same task. Thus, it is possible that familiarity with symbols could predict later memory performance for these items because of an increase in memory for old items as well as better protection from false memories (i.e., from false alarms). This pattern of results supports a dual coding account in so far as symbols need to be familiar to evoke secondary, verbal codes in memory.

At this point, the basic phenomenon—superior memory for symbols than for words—appeared to be on a solid foundation. A correlated factor could, however, be at play: Perhaps symbols are better remembered because they constitute a smaller overall set size than words. That is, there are fewer symbols in memory for people to search through relative to words, which could ease retrieval and improve memory for symbols. Because a similar argument had been made to explain picture superiority in memory (Nelson et al., 1985a, 1985b), this potential ‘set size’ confound was addressed in a third experiment.

2.3 Experiment 3 – Addressing a Set Size Explanation

2.3.1 Introduction

Following two successful experiments showing the basic effect of superior memory for symbols relative to words, a third experiment investigated a potential confound inherent in the previous two studies. The most prominent alternative explanation was that of set size, which had been used in the past to try to explain superior memory for ‘closed’ (items from a single category) relative to ‘open’ (unique items only) sets of words (Coltheart, 1993; Hulme et al., 1997; Poirier & Saint-Aubin, 1995; Roodenrys & Quinlan, 2000). A similar ‘set size’ explanation was initially put forth to explain picture superiority in memory: Pictures may belong to a smaller set size than words, making pictures easier to search through and select from during a memory test (Nelson et al., 1985a, 1985b; Nelson & McEvoy, 1979). By way of extension to the current study, symbols could be better remembered not because they evoke dual representations in memory but because of ease of retrieval due to reduced interference or search time resulting from being part of a smaller set than words. To illustrate, consider that any motivated individual likely could state thousands upon thousands of words that they know. In contrast, the same individual likely could come up with fewer than two hundred symbols. The actual values here are arbitrary, but the point is clear: Symbols may be easier to retrieve because there simply are fewer of them.

In Experiment 3, this possibility was assessed by reducing the set size of the words stimuli to roughly equate it to that for symbols. Because there were 80 symbols in the stimuli set of the previous study, a closed set of words was needed that contained roughly 80 items that most people would be able to recognize. The taxonomic category ‘common kitchen produce’ satisfied this requirement. In Experiment 3, memory for the set of symbols from the previous experiment was compared to memory for a list of words that are part of the fruit/vegetable category.

Because semantics were held constant between stimulus formats in the earlier experiments (i.e., $ vs. dollar), it seemed unlikely that differences in category set size were driving the previously reported memory benefit for symbols. Given this, the prediction for the current experiment was that, due to enhanced encoding for symbols relative to words, the symbol superiority benefit seen previously would persist even when set sizes were equated. Based on well-known concreteness effects in memory, however, the magnitude of the benefit for symbols was predicted to decrease as a result of higher concreteness for the words in the fruit/vegetable category, relative to the abstract words used in previous experiments. Experiment 3 also examined whether average familiarity with the current set of symbols would track with memory performance, given that this was the case in Experiment 2.

2.3.2 Method

Participants. An a priori power analysis was conducted using G*Power software (v. 3.1.9.7; Faul et al., 2007), targeting the smallest effect of interest in this study: a medium effect size of memory accuracy between the words and symbols groups ($d = 0.5$, $\alpha = .05$, two-tailed). This analysis indicated a minimum sample size of $N = 64$ per group or an ideal target sample size of $N = 86$ per group to achieve statistical power of 80% or 90%, respectively. Participants self-selected to participate in the study and had self-reported normal or corrected-to-normal vision, as well as having learned English before the age of 9. A total of 249 University of Waterloo undergraduate students each took part in a single session in exchange for course credit.

From this initial sample, participants' data were removed in sequential steps if they (1) had made duplicate responses (*n* = 20), (2) had corrupted or incomplete data files (*n* = 16), (2) took less than 5 minutes to complete the study (*n* = 0), (3) took more than 40 minutes to complete the study (*n* = 6), or (4) were ±3 *SDs* away from the mean of remaining participants for study duration (*n* = 10). After these exclusions, the sample consisted of 197 participants. Then, *R* statistical software was used to exclude participants whose performance was ±3 *SD*s away from the mean in their group on any memory performance metric (hits, false alarms, or accuracy; *n* = 6), or on the symbol familiarity task (*n* = 2)[11]. The final sample in the statistical analyses consisted of 189 participants, 76.72% female, with ages ranging from 17 to 49 (*M* = 20.6, *SD* = 3.8; Table 1).

Materials. Materials matched those in Experiment 2, except now the word stimuli were no longer word versions of the symbols. Instead, the word stimuli were 80 common kitchen fruits/vegetables (see Appendix A). Like the word-based stimuli in Experiments 1 and 2 (e.g., dollar), most of the items used here were one word (e.g., 'carrot'); 24 required two words (e.g., 'sweet potato').

Procedure. The procedure was similar to that for the pure words and symbols groups of Experiment 2. Participants were randomly sorted into one of two groups—encode and test symbols or encode and test words. The only difference, other than the changed set of words, was the replacement of the symbols matching task with a 'symbols familiarity rating task', which was completed immediately following the recognition memory test. In this final task,

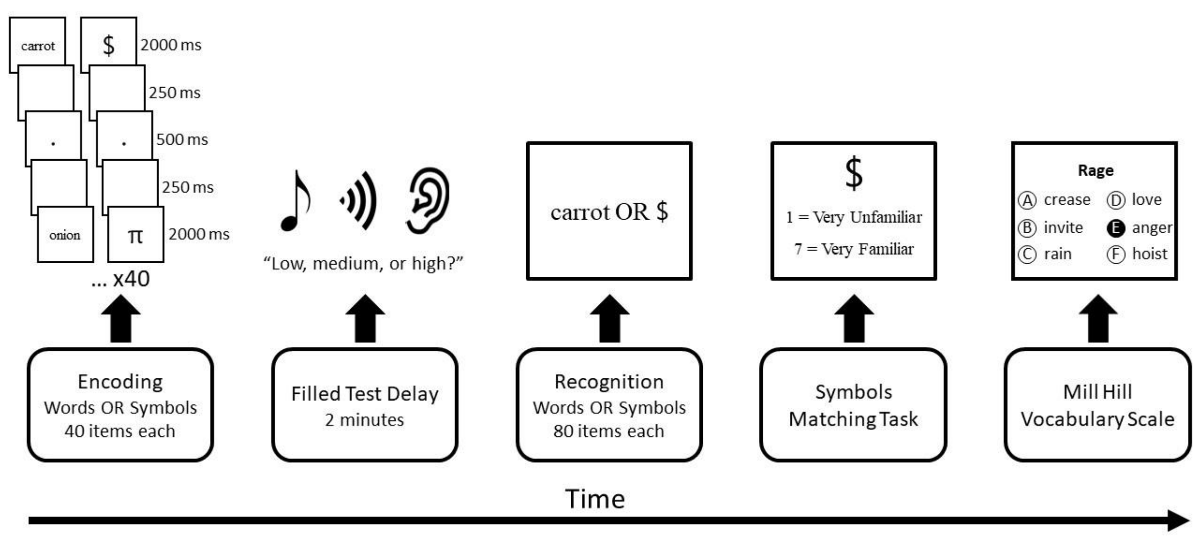
---

[11] Retaining data from these participants in the statistical analyses made no difference to the pattern of results.

all 80 symbols from the master stimulus list were presented one at a time in the center of the screen. Participants were instructed to indicate their familiarity with each symbol by clicking on one of seven response options along a Likert-style scale from '1: Very Unfamiliar' to '7: Very Familiar', with '4: Not Sure' as a middle response option. In the instructions for this task, a 'very familiar' symbol was defined as "you know what it means, personally use it, and/or see it used frequently." A definition of a 'very unfamiliar' symbol was also provided with the opposite description. The procedures and materials for this study were approved by the Office of Research Ethics at the University of Waterloo (project #41594). All data, pre-registrations, analysis code, experiment programs, and other materials are listed on the Open Science Framework (OSF; https://osf.io/e53z4/).

Figure 6

Experiment 3: Procedure



Note. 40 items (words or symbols) were presented at encoding to each participant, depending on condition assignment.

2.3.3 Results

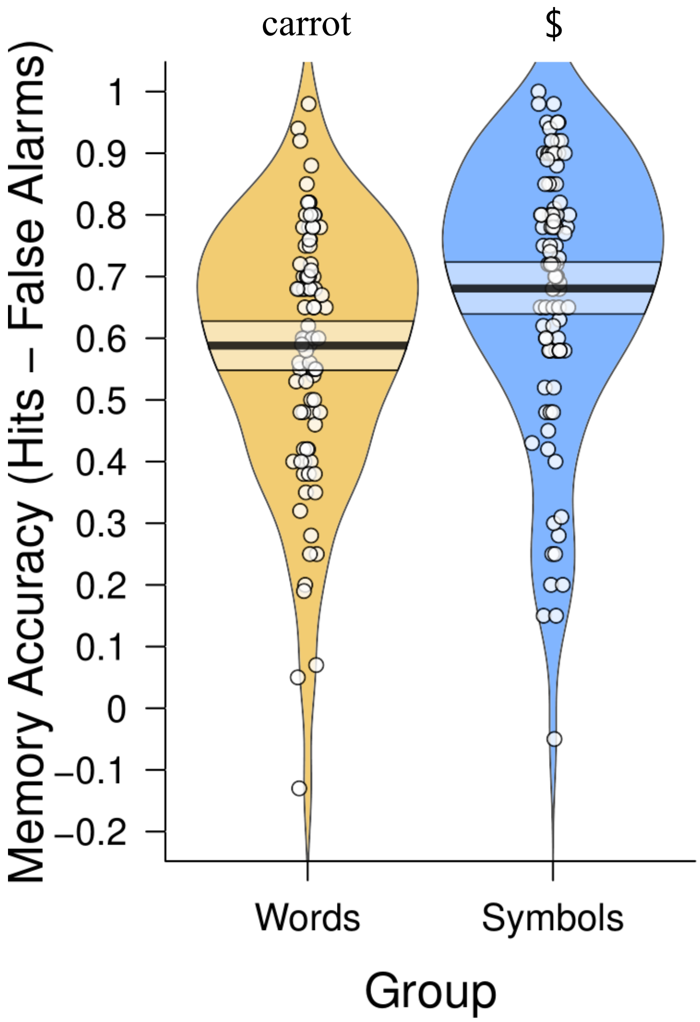
Memory Performance. To compare memory performance between words and symbols, a Welch-adjusted independent-samples *t*-test was conducted with memory accuracy (hit rate minus false alarm rate) as the dependent measure. Once again, the results showed that memory accuracy was higher for symbols than for words (Figure 7)[12], $t(187) = 3.03$, $p = .003$, $d = 0.44$, $CI_{95}[0.15, 0.73]$, $BF_{10} = 10.43$, with strong Bayesian evidence in support of the alternative model.

Correlations of Memory with English Proficiency and Symbol Familiarity. Correlations were used again to examine whether English language proficiency correlated with memory performance for symbols or words. Unlike in Experiments 1 and 2, these analyses failed to yield any significant correlations between Mill Hill scores and measures of memory performance in either condition (hits, false alarms, or accuracy; $rs \leq |.15|$, $ps \geq .154$). Next, I examined whether greater familiarity with the current symbols set would predict greater memory accuracy for symbols used in the memory study. Average symbols familiarity rating was, however, not correlated significantly with measures of memory performance in either condition (hits, false alarms, or accuracy; $rs \leq |.09|$, $ps \geq .391$).

[12] This analysis was also conducted with the pre-registered target sample size of 86 participants per group by randomly selecting data from the full data set. The pattern of results was identical.

Figure 7

Experiment 3: Recognition of Symbols and of Fruit/Vegetable Words



Note. Errors bars = 95% confidence intervals. Examples of encoding stimuli for each group are provided above their data points.

2.3.4 Discussion

This experiment addressed a potential confound of set size in accounting for my prior results (from Experiment 1 and 2) showing that symbols are better remembered than words. To do so, the set sizes of word and symbol stimuli were roughly equated by constraining the studied words to those in the fruit/vegetable category. The results here echoed those of the previous two experiments: Symbols were still better remembered than words. Contrary to my predictions based in dual coding theory and the results of Experiment 2, higher familiarity with symbols did not correlate with better memory for the same symbols in this experiment.

The findings of this experiment align well conceptually with those of Experiment 2 which showed that the effect is likely encoding-based, and therefore any easing of retrieval processes due to a smaller set size for symbols would likely be inconsequential. Thus far, the experiments have consistently shown that symbols have led to superior memory performance both relative to their word counterparts and now relative to words from a closed set of equivalent size.

Here, the concreteness effect in memory was eliminated by representing abstract concepts using symbols rather than words. That concrete words from the fruit/vegetable category were less well remembered than graphic symbols representing abstract concepts suggests that encoding of symbols more reliably leads to image codes in memory relative to the spontaneous imaging thought to occur with concrete words (Paivio & Csapo, 1973). This finding would be predicted by Paivio's dual coding theory: If the current set of symbols is indeed a set of 'mini-pictures,' then they should serve to 'concretize' their associated abstract concepts. That is, if concrete words lead to better memory because they are more easily imageable, as Paivio argued, then pictures must be inherently concrete. Therefore, rendering abstract concepts in symbol format may serve to make the to-be-remembered content imageable, thereby bringing about dual representations in memory despite the underlying association still being abstract. If symbols are indeed afforded dual codes in memory, akin to pictures, then memory for symbols should be similar to memory for images of even concrete objects. This prediction was directly tested in the following experiment.

2.4 Experiment 4 – Comparing Memory for Symbols and Pictures

2.4.1 Introduction

Recall that Paivio's classic dual coding theory postulates that pictures are better remembered because they routinely evoke two codes in memory: verbal and image. If one were to conclude based on the results of Experiments 1 through 3 that symbols are also reliably dual coded, then symbols should lead to memory superior to that for words (as has now been demonstrated) but equivalent to that for pictures. Critically, all pictures are inherently concrete insofar as they depict something visually and therefore immediately provide an image code in memory (Paivio, 1969). For images of easily recognizable content (e.g., a picture of an apple), spontaneous verbal labelling should also occur, providing a second code in memory (Paivio, 1969). Importantly, however, there should be no extra benefit to memory for images that depict physical things relative to those that represent abstract concepts (such as symbols). That is, the concreteness effect should be eliminated for familiar visual stimuli: When pictures are compared to other recognizable graphics (e.g., symbols), both should be highly likely to elicit dual codes in memory, regardless of the concreteness of their underlying concepts.

It is important, however, to disentangle the related ideas of 'abstract pictures' and 'pictures representing abstract concepts.' Whereas examples of the former may include meaningless geometric shapes or patterns, the latter must be identifiable as related to their underlying abstract concepts (e.g., symbols). Studies have previously found that images of concrete things are better remembered than abstract pictures (Bellhouse-King & Standing, 2007; Smith et al., 1990; Vogt & Magnussen, 2005). This makes sense from a dual coding perspective because the abstract material is less readily labelled and therefore less likely than a concrete image to evoke a verbal code in memory. Symbols, on the other hand, may represent an intermediate category of 'pictures representing abstract concepts,' such that the image could be labelled but the underlying associated concept is still abstract. Nevertheless, if dual coding of such stimuli is important for memory benefits, then memory for symbols should be equivalent to that for pictures of concrete stimuli such as objects because both should evoke imaginal and verbal representations. Here, this hypothesis was tested directly.

2.4.2 Method

Participants. An a priori power analysis was conducted using G*Power software (v. 3.1.9.7; Faul et al., 2007), targeting the smallest effect of interest in this study: a small- to medium-sized within-subject effect for the Symbols vs. Pictures pairwise comparison (d = 0.2, α = .05, power = 80%, two-tailed) as measured by free recall performance. This analysis indicated a minimum target sample size of N = 199. This target sample size would also provide adequate power to detect a small overall main effect of condition (target $N = 163$; $f = 0.10$, $\alpha = .05$, power = 80%, two-tailed). A small effect of Pictures vs. Symbols was targeted because there is no literature on the subject to date, and the hypothesized effect could be small if variation in the quality or fidelity of the image in memory is the only difference between these conditions.

Once again, participants were recruited using the Prolific data collection website. Filters were applied to allow participants to sign-up for the study only if they (1) had not taken part in one of our other studies on memory for sports logos, (2) were between the ages of 18 and 64, (3) were living in Canada or the USA, (4) were fluent in English, and (5) had normal or corrected-to-normal vision. Eligible participants self-selected to join the study. A built-in balancing service was used to ensure equivalent numbers of male and female participants in this particular experiment. A total of 241 participants took part in a single 25-minute session in exchange for $4.12 USD.

From this initial sample, participants' data were removed in sequential steps if they (1) duplicated responses ($n = 0$), (2) had self-reported non-ideal conditions (e.g., distractions) while completing the experiment ($n = 10$), (3) took less than 5 minutes to complete the study ($n = 0$), (4) took more than 40 minutes to complete the study ($n = 8$), or (5) were ±3 *SDs* away from the mean of the remaining participants for study duration ($n = 5$). These exclusions resulted in a sample of 218 participants. *R* statistical software was used to exclude participants whose performance was ±3 *SD*s away from the mean in their group on any memory performance metric (hits, false alarms, or accuracy; $n = 0$). The final sample of 218 participants used in the statistical analyses was 50.46% female, with ages ranging from 18 to 64 (*M* = 33.4, *SD* = 11.8).

Materials. Picture stimuli were 50 separate line drawings in black ink, representing small, everyday objects, presented on white backgrounds (see Appendix B); all were sourced from the International Picture Naming Project (IPNP) database (Szekely et al., 2004). To select images of common, easily identifiable objects from the IPNP, the following sequential sorting procedure was implemented: (1) only images from the ‘small artefacts’ category were selected, then (2) images were sorted from the least to the most number of alternative object names (etype), then (3) images were sorted from the most to the least percentage name agreement (elex1), then (4) images were sorted by highest to lowest CELEX frequency (efreq), and finally (5) any ‘outdated’ images (e.g., an old radio) or vague images (including ones that were hard to discern or that contained multiple objects; e.g., a table set with a fork, knife, napkin, and plate) were removed. After this sorting procedure, the top 50 images were retained for use in this study. Pictures were then re-sized to match the dimensions of the symbol stimuli (110 X 116 pixels). The word-based encoding stimuli were the word counterparts of the images sourced from the IPNP. The symbols stimuli were the same set of 50 items as had been used in Experiment 1.

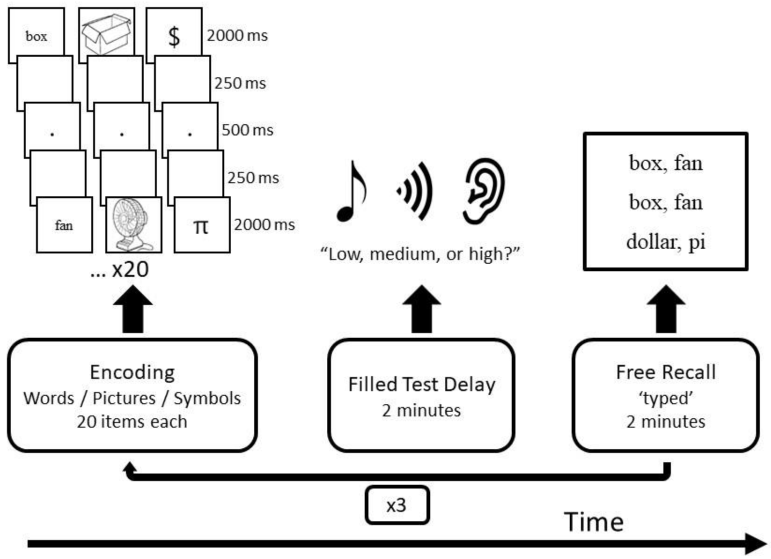
Procedure. The procedure generally followed that of Experiment 3, except for the following changes: (1) a third condition was added that included pictures of objects, (2) word stimuli were now labels of the selected pictured objects, (3) the symbols familiarity rating task and the Mill Hill Vocabulary Scale were not administered, and (4) there was now a final attention check question that asked participants whether they completed the experiment in ‘ideal’ conditions (they were not distracted, tried their best, etc.). Finally, this experiment switched back to a within-subject design.

Participants studied and were subsequently tested on three types of stimuli: pictures, words, and symbols. Stimuli were studied and tested in blocks, with block order randomized for each participant. During the picture and the word study phases, each participant was randomly assigned one of two 25-item stimulus sets. These two sets were counterbalanced such that no participant was shown the picture *and* word versions of the same item from the IPNP. To match the pictures and words blocks, symbols stimuli were randomly sorted into two lists of 25 items (only one of which was randomly assigned to each participant). During each encoding block, each participant was presented with 20 items, selected randomly from these master lists.

Following the tone classification filler task (as in Experiments 1-3), participants completed a free recall memory test for two minutes. Here, they were told to type out as many words, picture labels, or symbol labels as they could remember from the preceding study phase. Participants were encouraged to type out a physical descriptor of the remembered picture or symbol if they were unaware of its label. The procedures and materials for this study were approved by the Office of Research Ethics at the University of Waterloo (project #41594). All data, pre-registrations, analysis code, experiment programs, and other materials are listed on the Open Science Framework (OSF; https://osf.io/e53z4/).

Figure 8

Experiment 4: Procedure



Note. 20 items (words, pictures, or symbols) were presented at encoding in each of the three study-test cycles. Participants were never shown the same item in both word (e.g., fan) and picture (e.g., picture of a fan) format.

2.4.3 Results

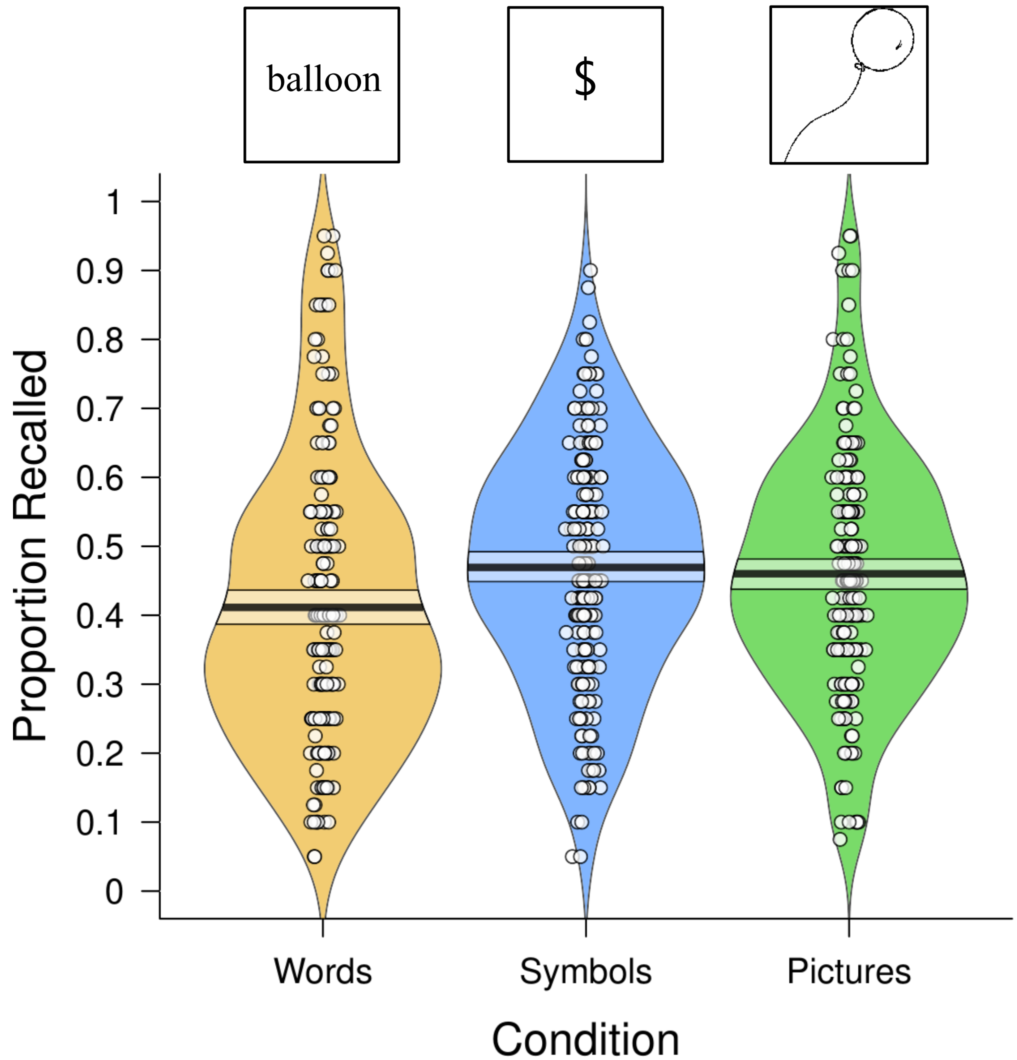
Memory Performance. Two naïve research assistants scored all recall data, determining whether each response was an intrusion (code = 0), a close response (code = 1), or an exact response (code = 2). As a reminder, in this experiment, participants typed their recall. As such, labels and physical descriptions of symbols and pictures were considered exact responses whereas synonyms were considered close responses. Then, three related metrics were tabulated based on these scores. First, a set of 'lenient' scores was calculated whereby all codes of 1 and 2 were counted the same as correctly recalled items. Next, a set of ‘strict’ scores was formed whereby only codes of 2 were counted as correctly recalled items. Finally, a weighted scheme was created whereby close and exact codes maintained their values (1 and 2, respectively), but to accommodate this the denominator was changed to 40 (rather than 20). The lenient scoring criteria was used throughout the results reported here, but additional analyses confirmed that the other weighting schemes made no difference to the overall pattern of effects.

To examine memory performance across conditions, a one-way repeated-measures ANOVA was conducted using the rstatix package for *R*, with Condition as the independent variable with three levels (Symbols, Words, and Pictures) and proportion of items recalled as the dependent measure. This analysis revealed a significant main effect of Group (Figure 9), $F(2, 651) = 6.38$, $p = .002$, $\eta_p^2 = .019$, $BF_{10} = 7.37$, with strong Bayesian support for the alternative model.[13] Paired-samples *t*-tests showed that recall performance was higher in the Symbols condition and in the Pictures condition than in the Words condition ($t(217) = 4.04$, $p < .001$, $d = 0.27$, CI_{95} [0.14, 0.41], $BF_{10} = 1.77e+02$, and $t(217) = 3.90$, $p < .001$, $d = 0.26$, CI_{95} [0.13, 0.40], $BF_{10} = 1.06e+02$, respectively), but that memory in the Symbols and Pictures conditions did not differ, $t(217) = 0.72$, $p = .475$, $d = 0.05$, CI_{95} [-0.08, 0.18], $BF_{01} = 10.25$, with strong Bayesian evidence for the null model in this final comparison.

[13] This analysis was also conducted with the pre-registered target sample size of 199 participants by randomly selecting data from the full data set. The pattern of results was identical.

Figure 9

Experiment 4: Recall of Words, Symbols, and Pictures



Note. Errors bars = 95% confidence intervals. Examples of encoding stimuli for each trial type are provided above their data points.

2.4.4 Discussion

The purpose of Experiment 4 was to examine whether memory is comparable for symbols and pictures. The prediction was that if symbols and pictures share similar underlying representation formats, then their memory performance should be equivalent, with both being superior to words. The results supported this prediction: Symbols and pictures both were better remembered than words and recall in the former two conditions was practically identical. That symbols were remembered better than words, but equally well to pictures of concrete objects, is suggestive of an all-or-none boost for images in memory. That is, there was no further enhancement to memory as a result of higher concreteness for pictures of objects relative to symbols representing abstract concepts. This implies that symbols were already serving the function of concretizing their associated concepts.

That pictures of abstract content (e.g., random shapes and/or patterns) are often remembered worse than pictures of concrete content (e.g., scenes or objects; Smith et al., 1990) speaks to the importance of being able to identify and label the content within a picture. That said, the results of this study, in general, are also consistent with both conceptual (Hamilton & Geraci, 2006; Nelson et al., 1977) and physical (Ensor et al., 2019b; Mintzer & Snodgrass, 1999) distinctiveness accounts of picture superiority in memory, accounts that will be discussed further in the General Discussion.

Chapter 3 – The Influence of Meaning-Based Familiarity on Memory for Symbols

3.1 Experiments 5A and 5B – Memory for Sports Logos[14]

3.1.1 Introduction

How we encode, store, and retrieve words and pictures in memory has intrigued researchers since the dawn of experimental psychology (Bergstrom, 1893; Moore, 1919; Mulhall, 1915). By the late 1800s, researchers began reporting better memory for objects relative to words (Calkins, 1898; Kirkpatrick, 1894). It was not until the 1960s, however, that research began to systematically compare memory for images and words (e.g., Shepard, 1967). By the early 1970s, Paivio (1971, 1991) and his colleagues had repeatedly demonstrated that pictures are better remembered than words, a finding termed the *picture superiority effect* (PSE). On this basis, Paivio proposed his influential dual coding theory—that pictures are remembered better because they often are represented in memory by two distinct codes—verbal and image—whereas words tend to evoke only the verbal code.

Examination of 'picture superiority' has continued over the decades (Ensor et al., 2019b; McDaniel & Pressley, 1987), yet researchers still often disagree on the underlying

---

[14] The work presented in Experiment 5 has been published:

Roberts, B. R. T., Fernandes, M. A., & MacLeod, C. M. (2023). Memory for symbolic images: Findings from sports team logos. *Journal of Applied Research in Memory and Cognition.* https://doi.org/10.1037/mac0000079 mechanism (e.g., Amrhein et al., 2002). Popular alternative accounts suggest (a) that pictures are *conceptually* more distinctive than words because they provide semantic characteristics that are diagnostic of a given item or semantic elaboration (Hamilton & Geraci, 2006; Nelson et al., 1977), akin to a levels-of-processing effect (Craik & Lockhart, 1972), and/or (b) that pictures are more *physically* distinctive because they vary more in appearance, especially relative to words that consist of the same recycled letters (Ensor et al., 2019b; Mintzer & Snodgrass, 1999). These two ideas are rarely mutually exclusive: Nelson's (1979) sensory-semantic model, for example, theorizes contributions from each type of distinctiveness, stemming from 'visual features,' 'phonemic features,' and 'meaning features.' Thus, pictures are held to be better remembered due to a general distinctiveness heuristic (Hunt & McDaniel, 1993).

Although both the distinctiveness account and the dual coding account are rooted in inherent differences between words and pictures, others have posited that the PSE arises from congruence between encoding and retrieval processes. Because pictures are usually more likely than words to access meaning during encoding (Craik & Lockhart, 1972), memory is greater on common retrieval tests, such as free recall and recognition, because these are theorized to assess primarily conceptual information (Jacoby, 1983; Roediger & Blaxton, 1987; Weldon & Roediger, 1987). Thus, because these retrieval tests echo the processing done at encoding, there is greater transfer-appropriate processing (Morris et al., 1977) for pictures.

One further explanation for the PSE is that the set size may be smaller for pictures than for words, making pictures easier to search through and select from during a memory test (Nelson et al., 1985a, 1985b; Nelson & McEvoy, 1979). This account relies on the assumption that, because pictures typically depict concrete things, there likely exists less imaged content than words in the world (words can represent abstract concepts as well, making their set size much larger). At the individual level, it has long been argued that people have more prior experience reading words than viewing images (see, e.g., Cattell, 1886). Thus, insofar as words and images constitute separate stimulus categories, there should be fewer images in memory and therefore reduced interference.

Not surprisingly, the PSE has interested the advertising world. Childers and Houston (1984) examined whether including images would influence consumer memory for brands and products. Participants were shown advertisements from a Yellow Pages phone book and were asked to rate them on dimensions including their physical appearance and meaning. Critically, the advertisements contained both images and words or only words. Participants' memory for these advertisements—both immediately and after a 2-day delay—was superior for those that included images.

The Childers and Houston (1984) study did not address the influence of integrating text into an advertisement with a product image beyond memory for a product image alone. According to dual coding theory, adding a label to an image should confer little benefit on memory because a verbal representation is likely already encoded with the image (Paivio & Csapo, 1973). In accord with this prediction, a study conducted without brand images, but comparing conditions of words, pictures, and pictures-plus-words, found that memory for the latter two conditions did not differ, with both superior to words alone (Maisto & Queen, 1992). My research sought to bridge these findings to investigate whether logos are better remembered than their brand name counterparts, and whether adding name labels would further enhance memorability.

My broad goal was to determine whether memory for logos—specifically, the logos of North American professional sports teams—both with and without integrated verbal labels, would be superior to memory for the labels alone. Would the PSE generalize to graphic symbols for abstract concepts, like those representing brands? Are logos, such as for the Toronto Blue Jays professional baseball team (Figure 10), processed similarly to images of concrete objects that are usually the focus of study in the PSE literature? My study is the first to directly assess whether picture superiority extends to brand logos. It is also among the few to provide evidence for whether symbolic representations of concepts (here, sports teams) operate using the same mechanisms that are proposed to underlie picture superiority for images of concrete items.

Two principal hypotheses were tested by comparing three conditions: Words (team names alone), Symbols (team logos without names), and Combined (team logos with integrated names). These stimuli can be seen as sharing equivalent semantics (i.e., all professional North American sports teams). My critical prediction based on dual coding theory is that, due to their combination of verbal and image representations, sports logos with and without integrated team names should be better remembered than their written (verbal only) counterparts. Because pictures are thought to already engage both verbal and image codes (Paivio, 1971), I also predicted no additional benefit from provision of team names embedded within the logos (i.e., Symbols = Combined).

Would my first hypothesis hold true for everyone or only for those very familiar with North American sports? I reasoned that people with high sports familiarity would already have in memory a verbal label corresponding to every logo. In contrast, people with low sports familiarity would not have ready access to the meanings of the logos, limiting the benefit conferred by presenting the sports teams in logo format compared to verbal format. Consequently, I administered a series of questions to gather data on participants' self-rated familiarity with and knowledge of sports, the goal being to compare memory benefits at different levels of this measure (for more details on this self-rating measure, see the Method). By examining whether familiarity with the stimuli affects encoding, I am among the first to determine whether relative expertise influences how pictures are studied and later remembered.

Finally, I examined whether the PSE would hold even for images from a finite set matched to a set of words. Recall that some (Nelson et al., 1985a, 1985b) have suggested that the PSE arises only because pictures represent a smaller set than words. Earlier work in this dissertation explored memory for 'everyday symbols' (e.g., !@#$%), demonstrating that symbols (e.g., $) are better remembered than their word counterparts (e.g., 'dollar'; Roberts et al., in press). A possible criticism of this work, however, is that symbols could be easier to remember simply because they form a smaller set than words, consequently reducing memory search time, the potential for interference, or both.

Using sports logos was therefore partially intended to find a finite set of symbolic stimuli with one-to-one mappings between symbols and their labels. I saw sports teams as ideal, given their cultural significance and the fact that they represent a ‘closed set’: In the four major North American sports leagues—the NFL (football), NHL (hockey), NBA (basketball), and MLB (baseball)—there currently are exactly 124 teams. This study is therefore interesting both for advertising purposes and for extension of the picture superiority literature. Moreover, it is informative in eliminating set size as a confounding factor when studying the encoding processes of other types of symbols that may not constitute a neat set.

Most graphic symbols (save for Peircean icons) would be difficult to interpret if their meanings were unknown to the viewer. For example, trying to remember ‘$’ would be more difficult if one could only encode its visuospatial properties. What would make things easier, of course, would be if the viewer also knew that the symbol represented the concept of ‘dollar’; symbols are most effective when an individual knows what they *mean*. If knowledge of a symbol’s meaning gives way to its verbal label (i.e., its word counterpart), it stands to reason that understanding what a symbol conveys is a sufficient (albeit not necessary) way to evoke an associated verbal code for the item in memory. That is, viewing a symbol provides an image code in memory but knowing what it means helps to provide the appropriate verbal label. Therefore, to gain dual codes in memory (along the lines of dual coding theory), it helps to know what a symbol means when trying to encode it. As a result, I predicted that familiarity with a symbol should track with one's knowledge of what the symbol means and therefore the likelihood that it will be remembered later.

In this experiment, I presented the logos or names of sports teams in a study phase and then, after a short delay, assessed memory for the studied teams on an old/new recognition test containing all of the teams. Importantly, I ran the study twice. In Experiment 5A, the recognition test stimuli matched the format seen during encoding. In Experiment 5B, the recognition test consisted of team names in text, regardless of whether they were encoded as words, symbols, or in a combined format. By doing so, I could examine differences in retrieval processes and better match prior research on the picture superiority effect. In both experiments, my major predictions were (1) that sports logos would be better remembered than team names (even equating for set size and semantic content), (2) that relative to the Symbols group, the addition of team name labels in the 'Combined' format would confer no further memory benefit, but (3) that familiarity could moderate the observed outcomes.

3.1.2 Method

Participants

An a priori power analysis was conducted using G*Power software (v. 3.1.9.6; Faul et al., 2007), targeting a medium-sized between-subjects omnibus main effect with three groups ($f = 0.25$, $\alpha = .05$) for memory accuracy (hit rate minus false alarm rate). This indicated required sample sizes of 53 or 69 participants in each of the three groups to achieve 80% or 90% statistical power, respectively. Accordingly, I aimed to collect a minimum of 53 participants per group (*N* = 159 total), with the ideal target sample size set higher at 69 per group (*N* = 207 total) in each experiment.

Initial samples consisted of data from 251 participants in Experiment 5A and 250 in Experiment 5B, all recruited via the online crowd sourcing platform, Prolific (www.prolific.co). Built-in pre-screening options on Prolific were used to permit participation only of those who had declared themselves to be residents of Canada or the USA. Participants were also required to have self-declared normal or corrected-to-normal vision, to be fluent in English, and to be between the ages of 18 and 64.

From these initial samples, participants were removed if they (1) had corrupted or incomplete data files (Experiment 5A: *n* = 8, Experiment 5B: *n* = 19), (2) took more than 30 minutes to complete the study (Experiment 5A: *n* = 1, Experiment 5B: *n* = 6), (3) took less than 5 minutes to complete the study (Experiment 5A: *n* = 1, Experiment 5B: *n* = 0), or (4) were greater than ±3 *SDs* away from the mean of remaining participants for study duration (Experiment 5A: *n* = 8, Experiment 5B: *n* = 5). Statistical analyses proceeded with 233 valid data files in Experiment 5A and 220 in Experiment 5B. From these samples, participants' data were removed if they were statistical outliers (±3 *SDs*) on any one of the metrics of memory performance (hits, false alarms, accuracy, or *d*-prime) as calculated within each experiment (Experiment 5A: *n* = 3, Experiment 5B: *n* = 3).

For Experiment 5A, the final sample used in formal statistical analyses consisted of 230 participants (49.13% female), ranging in ages from 18 to 64 (*M* = 32.38, *SD* = 10.96), split across three groups. For Experiment 5B, the final sample consisted of 217 participants (75.45% female, 4 preferred not to declare their sex), ranging in ages from 18 to 64 (*M* =

26.17, *SD* = 7.69), also split across three groups. The target sample size was therefore exceeded, ensuring adequate statistical power in both experiments[15]. Participation in each experiment took approximately 12 minutes and participants were paid the equivalent of $2.53 CAD for their time.

Materials

All materials were derived from professional North American sports teams playing in one of the four major leagues: NFL (football, n = 32), NHL (hockey, n = 32), NBA (basketball, n = 30), and MLB (baseball, n = 30). This ensured that the three types of stimuli—team names, team logos, and team logos plus names—came from equal and finite set sizes. There were 124 triples of stimuli generated in total (see Appendix C).

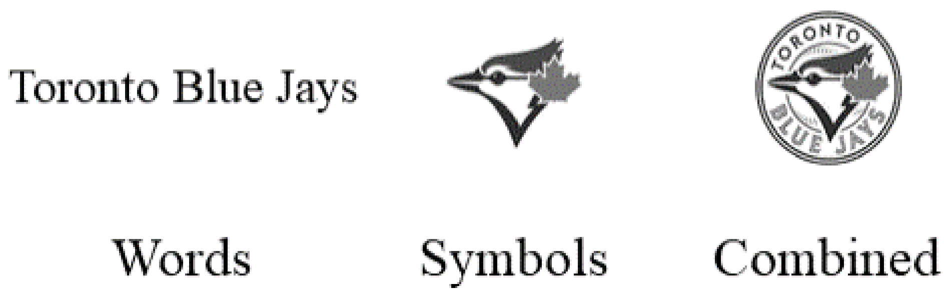
For the Words stimuli, full team names were presented at the center of the screen in Times New Roman size 48 black font, centered on a white background. Symbols were sourced from various online websites using the most up-to-date logos not containing a sports team's name or home city. When the most current logo for a team did contain one of these two elements, I sought a version with the text removed or I used slightly older logos that did not contain text (letters were permissible when unavoidable, e.g., the Pittsburgh Pirates logo). Stimuli in the Combined condition were required to contain a sports team's name, home city, or both, again with a preference for the most up-to-date logos. To equate colors between logos and better match stimuli used in previous studies of picture superiority as well as my

[15] Statistical analyses were also conducted with the pre-registered target sample size of 69 participants per group, per experiment by randomly selecting data from the full data set. The pattern of results was identical.

prior studies, all logo-type stimuli (Symbols and Combined) were then resized to be 110 X 116 pixels before being converted to greyscale with a white background (Figure 10).

Figure 10

Sample Stimuli Presented in the Words, Symbols, and Combined Groups



Note. I thank the MLB for granting permission to use the Toronto Blue Jays logos in this figure. *Major League Baseball trademarks and copyrights are used with permission of Major League Baseball. Visit MLB.com*

Procedure

Eligible participants self-selected to participate in the study via the Prolific data collection platform. After informed consent was provided, participants were randomly placed into one of the three stimuli groups before proceeding to complete the experiment on their personal computers. The experiment was built and hosted on the Qualtrics (www.qualtrics.com) survey-building website.

Prior to the study phase, participants were told that they would see either words or symbols presented one at a time on the screen and were instructed to try to remember as many as they could for a later memory test. Participants were then presented with either 62

Words, 62 Symbols, or 62 Combined stimuli sequentially in the center of the screen, randomly drawn from the complete set of 124 items. Each study trial consisted of a target stimulus shown for 2 s, followed by a blank screen for 250 ms, a fixation point for 500 ms, and finally another blank screen for 250 ms.

Next, participants completed a filled-delay task. They were instructed to press play on a media control bar to listen to a tone and then to respond by clicking ‘low,’ ‘medium,’ or ‘high,’ depending on the tone’s pitch. Examples of each pitch were provided in the task instructions. Participants were told to complete as many tone classification trials as they could before time elapsed and the page advanced, which occurred after two minutes. This task was included to guard against potential ceiling effects by eliminating recency and by minimizing rehearsal.

Following this interpolated task, participants completed an old/new recognition test for the items seen during the study phase. In a random order, all 62 target items from the encoding phase were presented, mixed with all of the remaining 62 items from the full set serving as lures. For Experiment 5A, test stimuli exactly matched those studied at encoding (except that, in the Words group, font size was reduced to 32 pt.). For Experiment 5B, all test stimuli were presented as team names in plain text, matching the retrieval test format of the Words group in Experiment 5A. Only text-based stimuli were used on the recognition test in Experiment 5B with the intention of limiting processing differences between conditions to the encoding phase.

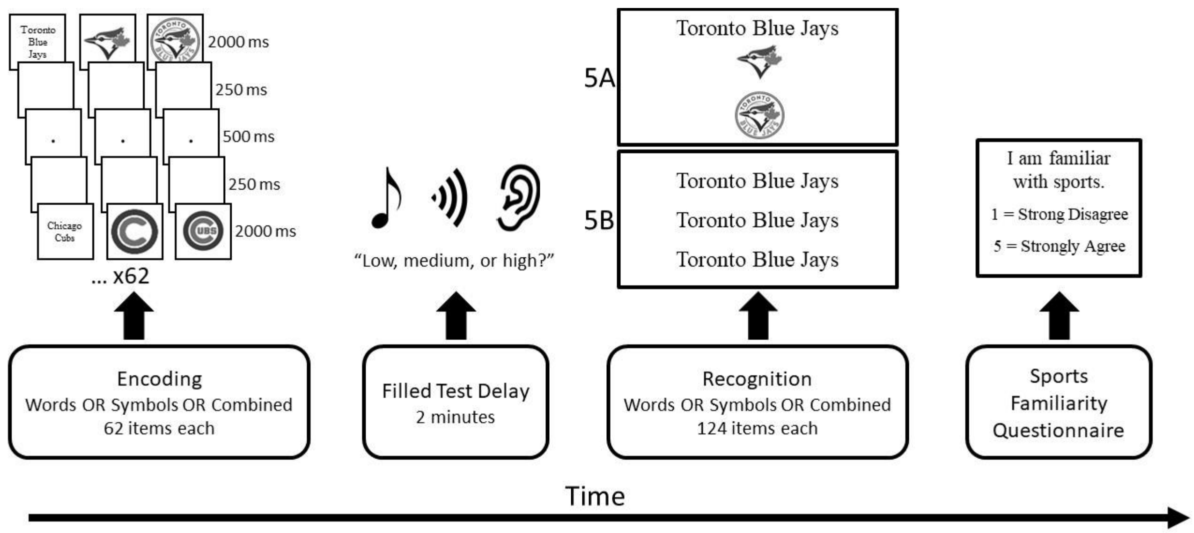
Participants were instructed to press the 'm' key to indicate that an item was 'old' (seen during study) or the 'n' key to indicate than an item was 'new' (not seen during study). For the Symbols and Combined groups in Experiment 5B, participants were additionally instructed to designate items as 'old' if the team name presented on-screen matched a previously seen team logo. Test items were presented one at a time in the center of the screen, advancing immediately to the next item following a participant's key response (i.e., the test was self-paced). Participants were instructed to respond as quickly and as accurately as possible.

Finally, following the recognition test, a series of questions probed participant familiarity with North American sports. This sports familiarity questionnaire asked whether they agreed or disagreed with statements such as: They watched sports often, they watched each type of sports league often, whether their self-rated knowledge of at least one major sports league was higher than average, and whether they were familiar (had pre-existing familiarity) with the sports teams presented in the study. Ratings were on a 1-5 scale from 'Strongly Disagree' to 'Strongly Agree,' with 'Neither Agree nor Disagree' as the middle option. The questionnaire also asked how often they watched sports in a typical week. Following this questionnaire, demographic information was collected, and a feedback letter was provided.

The procedures and materials for this study were approved by the Office of Research Ethics at the University of Waterloo (ORE #41594). All materials, experiment files, data, statistical analysis code, and pre-registration for these experiments can be found on the Open Science Framework (OSF; https://osf.io/xqsdz/).

Figure 11

Experiment 5: Procedure



Note. 62 items (team names, logos, or logos with text) were presented at encoding to each participant, depending on condition assignment. Recognition test stimuli were either congruent with encoding stimuli (Experiment 5A) or were text-based (Experiments 5B).

3.1.3 Results

Overall Memory Performance

For each experiment, a one-way[16] Welch-adjusted[17] between-subjects ANOVA was conducted with Group (Words, Symbols, Combined) as the only independent variable (Figure 12). The dependent measure was memory sensitivity (*d'*; d-prime) [18].

In Experiment 5A, there was a significant effect of Group[19], $F_{\text{Welch}}(2, 145.77) = 12.71$, $p < .001$, $\eta_p^2 = .15$, CI_{95} [0.07, 1.00], $BF_{10} = 4{,}091$. Games-Howell pairwise comparisons revealed that memory sensitivity was higher in the Symbols group than in the Combined group, ($p = .040$, $BF_{10} = 2.49$, $d = -0.40$, CI_{95} [-0.73, -0.07])[20], both of which showed higher memory sensitivity than the Words group ($p < .001$, $BF_{10} = 12{,}773$, $d = 0.80$, CI_{95} [0.47, 1.13], and $p = .030$, $BF_{10} = 4.03$, $d = 0.42$, CI_{95} [0.10, 0.74], respectively).

---

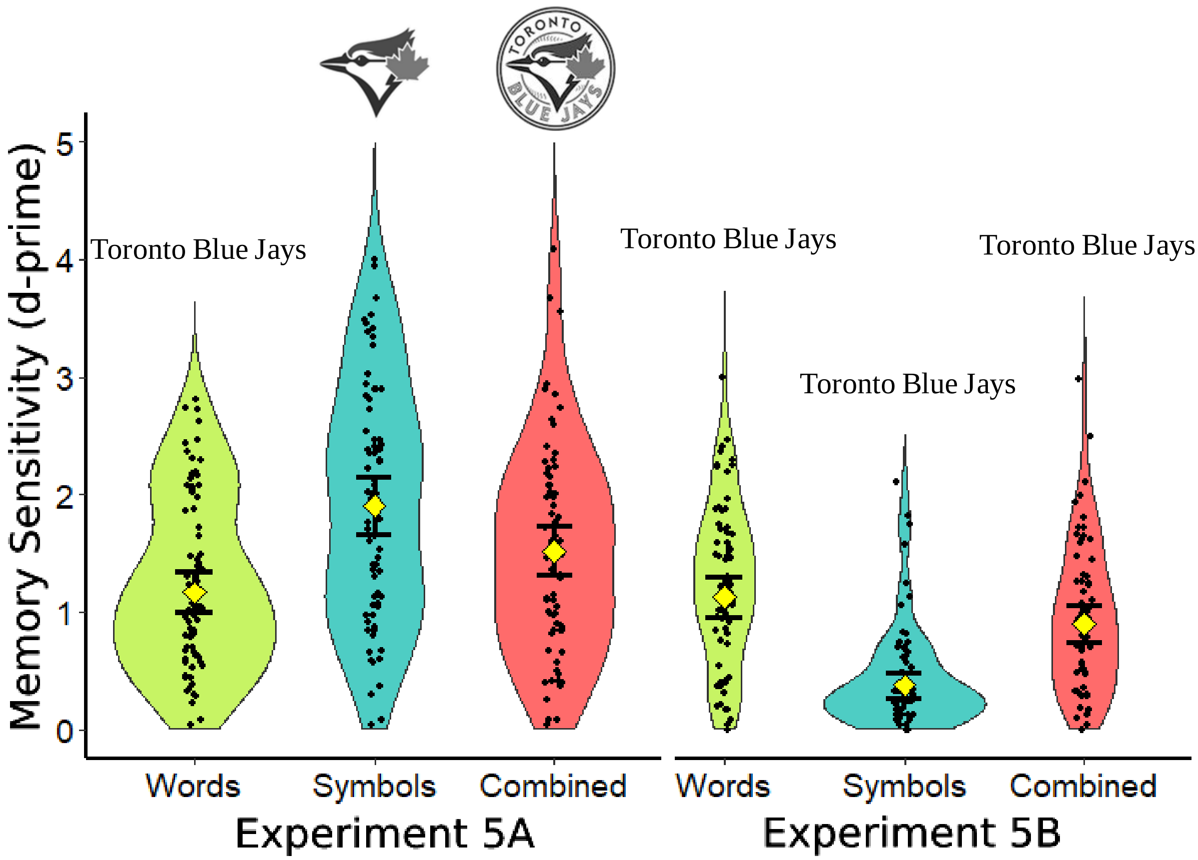
[16] I initially conducted a 2 (Experiment: 5A, 5B) x 3 (Group: Words, Logos, Combined) between-subjects ANOVA for memory sensitivity, which confirmed both significant main effects and the interaction, thus prompting separate one-way ANOVAs for each experiment.
[17] Welch's tests were used for all ANOVAs and Games-Howell tests were used for all pairwise comparisons in Experiments 5A and 5B. This was due to a violation of homogeneity of variance (via Levene's test) for the measure of memory sensitivity (*d'*; d-prime) in both Experiment 5A, $F(2, 227) = 4.46$, $p = .013$, and Experiment 5B, $F(2, 214) = 9.45$, $p < .001$, as well as when considering unequal group sizes in both experiments. In cases where homogeneity of variance was not violated, these tests—which are both based on the Welch–Satterthwaite adjustment to degrees of freedom—are still valid and recommended (Delacre et al., 2017, 2020; Ruxton, 2006). For consistency, I therefore used these tests throughout Experiments 5A and 5B.
[18] Throughout this dissertation, all *d'* and *c* values were formed using the *psycho* (Makowski, 2021) package for *R* and have been corrected using the log-linear rule for interpretation of extreme hit rate and false alarm rate values (Hautus, 1995).
[19] One-way ANOVAs, based separately on hit rate, false alarm rate, and accuracy, are presented in Appendix D.
[20] When pairwise comparison tests were performed on accuracy, this difference was no longer significant ($p = .065$).

In Experiment 5B, the effect of Group was also significant[21], $F_{Welch}(2, 138.16) = 31.17$, $p < .001$, $\eta_p^2 = .31$, CI_{95} [0.21, 1.00], $BF_{10} = 109,760,517$. Pairwise comparisons revealed a different pattern of results, however, such that accuracy was higher in the Combined group and in the Words group than in the Symbols group ($p < .001$, $BF_{10} = 53,830$, $d = 0.92$, CI_{95} [0.57, 1.27], and $p < .001$, $BF_{10} = 127,311,417$, $d = -1.17$, CI_{95} [-1.52, -0.82], respectively), but that the Combined group and the Words group did not differ ($p = .130$, $BF_{01} = 1.04$, $d = -0.32$, CI_{95} [-0.65, 0.01]).

[21] Because the percentage of female participants in Experiment 5A was 49% but in Experiment 5B it was 75%, I re-conducted all analyses with a trimmed sample in the latter study that matched the sex ratio of the former. The pattern of results was identical to that reported here.

Figure 12

Experiment 5: Memory Sensitivity Across Words, Symbols, and Combined Groups



Note. Error Bars = 95% confidence intervals. Examples of recognition test stimuli for each group are provided above their data points.

Memory as a Function of Sports Familiarity

Comparing Upper and Lower Quartiles

To assess whether memory performance was influenced by familiarity with sports, two one-way between-subjects ANOVAs were conducted with Group (Words, Symbols, Combined) as the independent variable, again using memory sensitivity as the dependent measure. In Experiment 5A, for those in the top quartile of sports familiarity, the three groups did not differ significantly (Table 2)[22], $F_{\text{Welch}}(2, 30.07) = 0.82$, $p = .450$, $\eta_p^2 = .05$, CI_{95} [0.00, 1.00], $BF_{01} = 4.02$. For those in the bottom quartile, however, there was a significant effect of Group, $F_{\text{Welch}}(2, 39.24) = 8.79$, $p < .001$, $\eta_p^2 = .31$, CI_{95} [0.11, 1.00], $BF_{10} = 286.56$. Games-Howell pairwise comparisons tests revealed that in the bottom quartile of participants, the Symbols group performed better than both the Combined group and the Words group ($p = .002$, $BF_{10} = 224.49$, $d = -1.28$, CI_{95} [-1.95, -0.60] and $p = .009$, $BF_{10} = 44.29$, $d = 1.03$, CI_{95} [0.39, 1.67], respectively), and that the latter two did not differ ($p = .490$, $BF_{01} = 1.72$, $d = -0.35$, CI_{95} [-0.95, 0.25]).

The main effect of Group in Experiment 5B was not statistically significant for participants in the top quartile of sports familiarity, $F_{\text{Welch}}(2, 27.54) = 1.76$, $p = .190$, $\eta_p^2 = .11$, CI_{95} [0.00, 1.00], $BF_{01} = 2.06$, whereas it was significant for participants in the bottom quartile, $F_{\text{Welch}}(2, 33.40) = 21.69$, $p < .001$, $\eta_p^2 = .57$, CI_{95} [0.36, 1.00], $BF_{10} = 2{,}009$. Relative to the finding for Experiment 5A, Games-Howell pairwise comparisons demonstrated the opposite pattern of results in the bottom quartile of participants: The Symbols group performed *worse* than both the Combined group and the Words group ($p = .003$, $BF_{10} = 50.69$, $d = 1.22$, CI_{95} [0.52, 1.90] and $p < .001$, $BF_{10} = 9{,}847$, $d = -1.70$, CI_{95} [-2.43, -0.95],

[22] Sports familiarity scores were calculated as the grand total after summing a given participant's responses on the questionnaire, with higher scores indicating greater sports familiarity and/or knowledge. Scores on this scale ranged from 9 to 55. The lowest quartile of sports familiarity scores encompassed scores 9 to 11 and 9 to 10 in Experiments 5A and 5B, respectively. The top quartile of sports familiarity scores encompassed scores 36 to 55 and 29 to 55 in Experiments 5A and 5B, respectively. Thus, the range of scores was greater in the top quartile than in the bottom quartile.

respectively), whereas the Combined group and the Words group did not differ (p = .250, BF_{01} = 1.18, d = -0.49, CI_{95} [-1.09, 0.11]).

Table 2

Experiments 5A and 5B: Descriptive Statistics for All Memory Performance Metrics as a Function of Experiment, Group, and Sports Familiarity Quartile

Experiment	Group	Sports Familiarity Quartile	*n*	Accuracy		Hit Rate		False Alarm Rate		*d'*		*c*	
				M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
5A	Words	Overall (1-4)	84	0.39	0.24	0.73	0.14	0.34	0.21	1.17	0.79	-0.05	0.42
		1st	24	0.43	0.19	0.71	0.17	0.28	0.16	1.30	0.65	0.05	0.46
		4th	12	0.45	0.25	0.79	0.10	0.34	0.23	1.80	0.86	-0.11	0.44
	Combined	Overall (1-4)	72	0.50	0.24	0.72	0.16	0.23	0.17	1.52	0.88	0.15	0.36
		1st	20	0.35	0.23	0.62	0.17	0.27	0.19	1.06	0.73	0.23	0.42
		4th	21	0.56	0.23	0.78	0.14	0.22	0.15	1.68	0.86	0.05	0.29
	Symbols	Overall (1-4)	74	0.59	0.26	0.78	0.16	0.19	0.16	1.90	1.03	0.15	0.33
		1st	21	0.66	0.22	0.81	0.13	0.15	0.16	2.13	0.93	0.17	0.33
		4th	22	0.54	0.33	0.76	0.19	0.22	0.20	1.83	1.23	0.15	0.41
5B	Words	Overall (1-4)	77	0.38	0.25	0.68	0.15	0.29	0.19	1.13	0.77	0.10	0.37
		1st	23	0.34	0.22	0.65	0.15	0.31	0.19	1.00	0.68	0.11	0.40
		4th	14	0.36	0.30	0.66	0.17	0.30	0.20	1.06	0.90	0.34	0.07
	Combined	Overall (1-4)	69	0.31	0.21	0.60	0.15	0.29	0.18	0.90	0.65	0.21	0.39
		1st	21	0.24	0.21	0.56	0.15	0.32	0.19	0.68	0.60	0.23	0.37
		4th	13	0.40	0.25	0.69	0.17	0.29	0.23	1.25	0.86	0.15	0.50
	Symbols	Overall (1-4)	71	0.14	0.17	0.59	0.15	0.45	0.15	0.38	0.48	-0.02	0.32
		1st	21	0.05	0.09	0.55	0.12	0.49	0.13	0.21	0.05	0.30	0.06
		4th	24	0.24	0.28	0.69	0.15	0.45	0.19	0.66	0.82	-0.18	0.31

Note. The 'overall' row includes all participants from that group, regardless of sports familiarity score. 1st quartile = lowest 25% of participants on sports familiarity score across all groups within each experiment, 4th quartile = highest 25%.

Correlation Between Sports Familiarity and Memory Performance

Finally, I examined Pearson correlations between total sports familiarity scores and all of the metrics of memory performance (hit rate, false alarm rate, and memory sensitivity). For Experiment 5A, collapsed across groups, there was a significant relation between sports familiarity score and hit rate, $r(228) = .16$, $p = .014$ (all other $ps \geq .104$). When performing correlation analyses split by group, I found that sports familiarity correlated significantly with hit rate in the Combined group, $r(70) = .31$, $p = .009$. All other correlations were non-significant ($ps \geq .090$).

In Experiment 5B, sports familiarity was found to correlate significantly with memory sensitivity, $r(215) = .21$, $p = .002$, and with hit rate, $r(215) = .21$, $p = .002$, when collapsed across groups. The correlation between sports familiarity and false alarm rate was, however, non-significant ($p = .275$). Breaking these correlations down to the group level, sports familiarity scores were correlated significantly with memory sensitivity in the Combined group, $r(67) = .30$, $p = .011$, and in the Symbols group, $r(69) = .45$, $p < .001$, as well as with hit rate in the same two groups ($r(67) = .27$, $p = .023$, and $r(69) = .34$, $p = .004$, respectively). All other correlations were non-significant ($ps \geq .162$).

3.1.4 Discussion

In the present set of two related experiments, I compared memory for professional sports teams presented in three different ways during encoding: team names only, team logos without names, and team logos with integrated names. I aimed first to determine whether the picture superiority effect (PSE) generalized to graphic symbols representing abstract concepts. Second, I wanted to determine whether the PSE would hold even for images sampled from a finite corpus matching the set size of words (i.e., team names). This novel approach also allowed examination of the effects of familiarity on picture encoding. Given that the symbols I used possess high cultural significance, some participants were very familiar with them and some were not. Thus, I could establish whether expertise plays a role when encoding and later retrieving pictures from memory.

My first and second major hypotheses were that, in accord with the findings of Maisto and Queen (1992) and aligning with both dual coding theory and the physical distinctiveness account, the Symbols group and the Combined group would exhibit better memory relative to the Words group. Moreover, because pictures routinely promote verbal encoding, the Symbols group and the Combined group were not expected to differ. I left open the possibility, however, that familiarity with sports could moderate these predictions.

Experiment 5A indicated that memory sensitivity was indeed better when sports logos were studied either with (i.e., Combined) or without (i.e., Symbols) integrated team names, relative to when only team names were provided (i.e., Words). However, that the Symbols group performed better than the Combined group ran contrary to my second major hypothesis—and to dual coding theory. Perhaps some participants in the Symbols group benefited from additional semantic elaboration at encoding by creating descriptors for the logos, as shown in previous PSE research (Slamecka & Graf, 1978; Weldon & Roediger, 1987). The sports familiarity analyses provide evidence for this possibility: The Symbols group demonstrated better recognition by participants more likely to rely on self-generated descriptive labels for the logos (i.e., those with low familiarity) but not by participants who already knew the labels (i.e., those with high familiarity). Thus, Experiment 5A demonstrated that picture superiority does generalize to graphic symbols that represent abstract concepts, even for stimuli sourced from a finite set with semantic information held constant. The findings of Experiment 5A do, however, provide little support for the role of familiarity in memory for symbols. That the group with lowest familiarity saw the greatest memory for symbols is directly in opposition to my predictions derived from dual coding theory. It seems that when symbols are encoded and retrieved in the same format (as was the case here) familiarity is not a determinant of memory; rather, when unfamiliar with the stimuli, participants may be able to improve memory for symbols by using elaborative encoding.

Recognition Memory Performance for Text-Based Retrieval Stimuli

The results of Experiment 5B require a more nuanced explanation. I reasoned a prioi that adding a verbal label in the Combined group would not augment performance beyond that in the Symbols group: Dual coding theory holds that verbal labels are routinely encoded with images. Of course, applying accurate verbal labels requires familiarity with the imaged content. In Experiment 5B, the PSE was sharply reduced in participants who lacked knowledge of the sports teams: When all test stimuli were presented in word format, I observed a performance advantage for the Words and Combined groups relative to the Symbols group, presumably because these two groups both provided team names, which consequently matched the stimulus format on the recognition test. That performance in the Words group and the Combined group was equivalent in Experiment 5B (even with picture encoding in the latter case) seems to conflict with a dual coding account and instead could be a result of transfer appropriate processing (Morris et al., 1977), or encoding specificity (Tulving & Thomson, 1973) for words.

The Symbols stimuli—lacking team name labels—likely were particularly difficult to identify in Experiment 5B by all but the most dedicated sports fans. Indeed, high familiarity participants in the Symbols group fared much better: The poor performance of low familiarity participants likely drove the overall decline in performance for this group. In fact, low familiarity participants in this group were essentially at chance (mean accuracy = 5%), suggesting that they simply did not know the team name-logo association. Sports familiarity would play the largest role in this condition because the memory test contained team names not provided at encoding for the Symbols group. This pattern was borne out in the correlational analyses: The highest correlation reported here (r = .45) was between sports familiarity and memory sensitivity in the Symbols group of Experiment 5B.

The drastic drop in performance for the Symbols condition in Experiment 5B (where recognition was text-based) suggests an important caveat to the PSE: One *must* be able to easily associate the correct verbal label for an image when the studied content is later only available for recognition by its proper label. There apparently also are boundaries to the PSE when familiarity is too high: All three groups in the top familiarity quartiles performed similarly on the recognition tests in both experiments. Paivio (1971) would predict that the PSE would be reduced or eliminated when words are likely to evoke image representations in memory. Typically, concrete words are more likely than abstract words to evoke imagery; my study demonstrates an exception to this proposal when familiarity is high.

High familiarity with the stimuli may have caused participants to spontaneously image related content, even though the studied word represented a concept (i.e., a sports team), not a concrete object. Devoted sports fans in the top quartiles likely imaged at least the associated logo, and perhaps more (e.g., uniforms, home stadium, etc.). I speculate that the magnitude of the picture superiority effect in memory may depend on familiarity being 'just right': The participant must be able to readily retrieve the meaning and corresponding verbal label for an image if the recognition test is text-based, but they must not be too familiar with the concepts that the words represent, lest the words automatically elicit mental imagery for associated content, effectively equating to encoding of pictures.

Implications for Logo and Symbol Memory

It should be noted that participants often demonstrate poor memory for highly familiar visual stimuli, perhaps due to inattentional amnesia (Wolfe, 1999). As illustrations, impoverished visual memory has been reported for national flags (Blake & Castel, 2019), the locations of safety equipment (e.g., fire extinguishers; Castel et al., 2012), buildings (Murphy & Castel, 2021), coins (Marmie & Healy, 2004; Nickerson & Adams, 1979), letters of the alphabet (Wong et al., 2018), and most pertinently here—brand logos (e.g., the Apple logo; Blake et al., 2015). The key conclusion is that while visual memory can sometimes be poor for well-known stimuli, familiarity with a logo or picture remains critical in determining precisely what the imaged content represents. Therefore, while the minutiae of brand logos may go unnoticed due to overexposure, the intended effect of enhanced brand memory likely occurs with high degrees of familiarity.

The finding that set size did not drive enhanced memory for symbols aligns with past research on word recall reporting that category set size is inconsequential when a category cue is available at study and test (Nelson & McEvoy, 1979), as was true here. Indeed, the findings of Experiment 3 of this dissertation demonstrate that even restricting the set size of the words to a single category of concrete nouns (e.g., produce) does not undermine superior memory for symbols (e.g., !@#$%). Furthermore, because sports team logos change over time but team names usually do not, the set size of logos actually could be *larger* than that of written team names. If so, this would further confirm that a smaller set size for symbols than for words is unlikely to be driving superior memory here or in my other experiments of memory for symbols.

Henderson and Cote (1998) noted that logos add value to a company only if (1) consumers remember seeing the logo, and (2) the logo reminds consumers of the brand name. The logo-brand name association is therefore paramount to the overall goal of improving brand memory; Experiment 5B supports this claim. Creating ads with images increases associative memory for the brand names (Barrett, 1985), attracts visual attention (Rihn et al., 2019), and helps consumers narrow their interpretation (van Riel & van den Ban, 2001), but my results indicate a boundary condition in explicit memory associations when familiarity is low and retrieval is text-based. I agree, then, that effective logos must serve as cues to remind consumers of brand names. Integrating a brand name into a logo is an effective way to foster this association. Furthermore, logos with integrated text provide two forms of encoding support, ensuring that consumers can recognize brands no matter the format in which they are encountered later (i.e., text-less logos or brand names).

Conclusion

In conclusion, sports team logos were better remembered than the text of a team name. Adding an integrated team name label did not further enhance memory. However, memory for team names associated with label-less logos suffered when familiarity with sports was low and retrieval was text-based, likely because participants were not familiar with the pre-existing logo-name associations. Results from participants with high familiarity suggest that the PSE can be eliminated if brand familiarity is already high. Consequently, I suggest that, whenever possible, advertisers include matching logos in advertisements and on products. Moreover, because advertisers do not always know the format in which a consumer might later encounter their brand, advertisements containing logos with integrated verbal labels are likely the best way to maximize memorability in all contexts. Even people relatively unfamiliar with a logo could learn to connect a verbal label to it, thereby harnessing two forms of encoding support in memory to remember the brand more effectively.

3.2 Experiment 6 – The Intrinsic Memorability of Everyday Symbols

3.2.1 Introduction

Thus far, I have argued—largely based on dual coding theory—that it is possible symbols elicit image codes in memory which effectively serve to 'concretize' abstract concepts by providing visual referents. In other words, symbols allow for imagery-based processing of concepts that would otherwise invoke only verbal processing.

Symbols can only effectively communicate ideas when the viewer already knows, or can figure out, what they are intended to convey. For practical reasons, then, it is preferable for symbols to be highly standardized, and better yet for them to have intuitive interpretations. But what if someone has never seen a symbol before or cannot figure out what it means? In that case, the symbol would fail to elicit a verbal code in memory because the viewer would not have a verbal label ready to apply. Therefore, familiarity with 76ymboll is likely key to adding the second code.

In the case of pictures, on the other hand, researchers often use images of objects and therefore need not be concerned with how familiar a person is with the imaged content because the pictured objects typically are all highly familiar and easily labelable. My prediction is that memory for symbolic content, whether in word form or symbol form, should improve with greater familiarity. To access a verbal code, it helps to be able to readily retrieve the meaning and corresponding verbal label for a symbol.

The Influence of Familiarity

In the final experiment of this thesis, familiarity ratings were gathered from naïve participants for the same set of symbols that had been used in Experiments 1-4, the goal being to assess whether average familiarity with a symbol would correlate with its average memorability in the current experiments. Research in psycholinguistics has typically measured 'familiarity' in one of two distinct ways: 'meaning familiarity' and 'frequency of occurrence.' For instance, Balota et al. (2001, 2004) demonstrated that traditional measures of familiarity are strongly associated with meaning, whereas subjective estimates of frequency are more related to actual frequency of occurrence. Keeping with current psycholinguistic conceptualizations of familiarity, two distinct but related conceptual components were studied: meaning-based familiarity (Juhasz et al., 2015) and subjective frequency of occurrence (Balota et al., 2001). For the former measure, participants were asked to rate how well they know the meaning of the symbol; for the latter measure, they were asked how often they encounter it.

Based once again on dual coding theory, I reasoned that meaning-based familiarity with a symbol would increase the likelihood of it being dually encoded. The prediction was that high familiarity with a symbol would make labelling it easier and would increase the likelihood of that symbol eliciting an additional verbal code in memory. The logic follows along the lines of my earlier investigation of sports logos in Experiment 5: A participant that is highly familiar with sports would be able to generate the correct verbal label for a label-less logo (i.e., the 'Symbols' condition in that experiment). Higher rates of dual coding were expected and therefore a positive correlation was predicted between symbol familiarity and memory performance for symbols. As a result, this analysis included data from all participants who encoded symbols in the four preceding experiments.

Although my main predictions are centered in meaning-based familiarity, it is also possible that subjective estimates of frequency could provide a window into how reliably a symbol can be identified or imaged. For instance, the '#' symbol can be found on the top row of most English QWERTY keyboards and is encountered frequently on social media sites, but its meaning could be relatively ambiguous to participants, especially when taken out of context. It is in cases such as these that one would expect subjective frequency estimates to correlate positively with memory performance because, while the 'true' meaning of the symbol is somewhat variable, it is still frequently encountered and consequently a participant—without knowing the meaning of the symbol—could very well be able to label it (e.g., 'hashtag') or spontaneously image it when provided with its word counterpart. By determining whether familiarity and/or frequency correlate with memory performance, one can obtain a glimpse into the underlying mechanisms driving encoding of symbols insofar as verbal labelling or understanding are required for dual coding.

The Influence of Inherent Memorability

Recent work has demonstrated that all stimuli have some degree of inherent memorability that can be separated from the cognitive and neural signatures of attention, priming, and low-level perception (Bainbridge et al., 2013, 2017). To investigate whether the symbols in the set of stimuli used across the first four experiments in this dissertation were better remembered simply because of their intrinsic memorability properties, the freely available ResMem neural network (Needell & Bainbridge, 2022) was enlisted to assign memorability scores to each of the symbols. ResMem is a deep residual neural network that was built upon two existing models—ResNet-152 (Khosla et al., 2015) and AlexNet (Krizhevsky et al., 2012)—before being re-trained for the purpose of optimizing predictions of memorability. The model was trained to provide memorability estimates by using the LaMem (Khosla et al., 2015) and MemCat (Goetschalckx & Wagemans, 2019) databases, each of which contains thousands of colour images with memory performance metrics from real humans completing a continuous recognition task.

ResNet-152 was originally trained for the purpose of classifying images into semantic categories using the ImageNet (Deng et al., 2009) dataset of over 14 million pictures. AlexNet, on the other hand, was trained using 1.2 million images to extract features of low-level perceptual attributes in order to classify pictures (Krizhevsky et al., 2012). As a result of the combination of these two networks, the new ResMem network that was used here makes use of both semantic and perceptual features of the image to generate memorability predictions (Needell & Bainbridge, 2022).

Although ResMem has not been explicitly trained on or validated with symbols, it has been shown to work well with homogenous visual stimuli from the same semantic category that make heavy use of black and white shading (the Food Folio dataset; Lloyd et al., 2020). In brief, ResMem takes into account perceptual features (such as lines and curves) as well as conceptual features of the image (such as category) to provide a single value, ranging from 0 to 1, that indicates the item's estimated likelihood of being remembered by a real person. Unfortunately, as is the case with most neural networks, ResMem is quite opaque regarding how it determines memorability. These models are 'black boxes' such that we cannot be sure which factors are preferentially being used to predict memorability (Joshi et al., 2021). The creators of this neural network, however, have suggested that semantic features may contribute more variance to memorability scores than do perceptual features in the ResMem network (Needell & Bainbridge, 2022).

The use of ResMem in exploratory analyses was motivated by the question of whether symbols have intrinsic properties that make them particularly memorable, and whether these properties are separable from participants' ratings of familiarity and frequency of occurrence. From the previous experiments, it was expected that the everyday symbols used here would be classified as highly memorable, and that—based on previous work on memorability (e.g., Bainbridge et al., 2017)—the inherent memorability of these symbols would contribute unique variance to their memorability, apart from that contributed by familiarity and frequency.

3.2.2 Method

Participants. For this experiment, no a priori power analysis was conducted as there was no targeted effect size of interest. Instead, the target sample size was set to match the largest sample size that had been collected in the previous experiments (target N = 315). This target sample size also substantially exceeds the 30-100 participant ratings per item that are often found in psycholinguistic norming research with rating scales (e.g., Balota et al., 2001).

Once again, participants were recruited from Prolific. Several restrictions were set to match the sample characteristics of Experiments 1-4 in this dissertation. Participants could sign up for the study if they (1) had not taken part in one of our earlier studies on memory for symbols or sports logos, (2) were between the ages of 18 and 26, (3) lived in Canada or the USA, (4) were fluent in English, (5) had normal or corrected-to-normal vision, and (6) were currently undergraduate students. Finally, an attempt was made to roughly match the average percent of female participants across the preceding experiments (75.89%). Eligible participants self-selected to join the study. A total of 350 participants took part in this single 9-minute session in exchange for $2.13 USD.

From this initial sample, participants' data were removed in sequential steps if they (1) were duplicate responses ($n = 0$), (2) took less than 3 minutes to complete the study ($n = 0$), (3) took more than 40 minutes to complete the study ($n = 0$), (4) were ±3 *SDs* away from the mean of remaining participants for study duration ($n = 8$), or (5) had self-reported non-ideal conditions for their participation in the experiment ($n = 5$). These exclusions resulted in a final sample of 337 participants who were entered into the statistical analyses, consisting of 76.26% females, with ages ranging from 18 to 26 ($M = 20.9$, $SD = 1.8$).

Materials. The full set of 80 symbols used in Experiments 2 and 3 (Figure 1) was used here. As in the preceding studies, symbols were sized at 110 X 116 pixels using Segoe UI Symbol black font, centered on a white background. Participants rated each symbol using two different conceptualizations of familiarity with an item—one for meaning-based familiarity and one for frequency of occurrence—each with a different 7-point scale.

The measure of meaning-based familiarity that was used, including the response options and instructions, was adapted from Juhasz et al. (2015). Response options ranged from (1) Very Unfamiliar to (7) Very Familiar. Intermediate response options were unlabeled. The instructions for this scale were as follows:

> “Please provide a rating between 1 (very unfamiliar) to 7 (very familiar). If you feel you know the meaning of the symbol and use it frequently, then give it a high rating on this scale. For example, Jim has known the symbol * (asterisk) since he was a child, uses the symbol frequently, and if asked could easily tell anyone what it is. He should give this symbol a very high rating. If the symbol is not familiar at all, you do not know its meaning, or you are not sure whether it is a symbol or not, then give it a low rating. For example, Jim has never encountered the symbol № (numero) and has no sense of whether it is a symbol or not, or for what it means. He should give this symbol a very low rating. If the item falls somewhere in the middle of these two extremes, where you have some familiarity with the symbol, then give it a rating in the middle of the scale.”

The measure of subjective frequency that was used, including the response options and instructions, was adapted from Balota et al. (2001). Responses once again were made on a 7-point scale. The instructions for this scale were as follows:

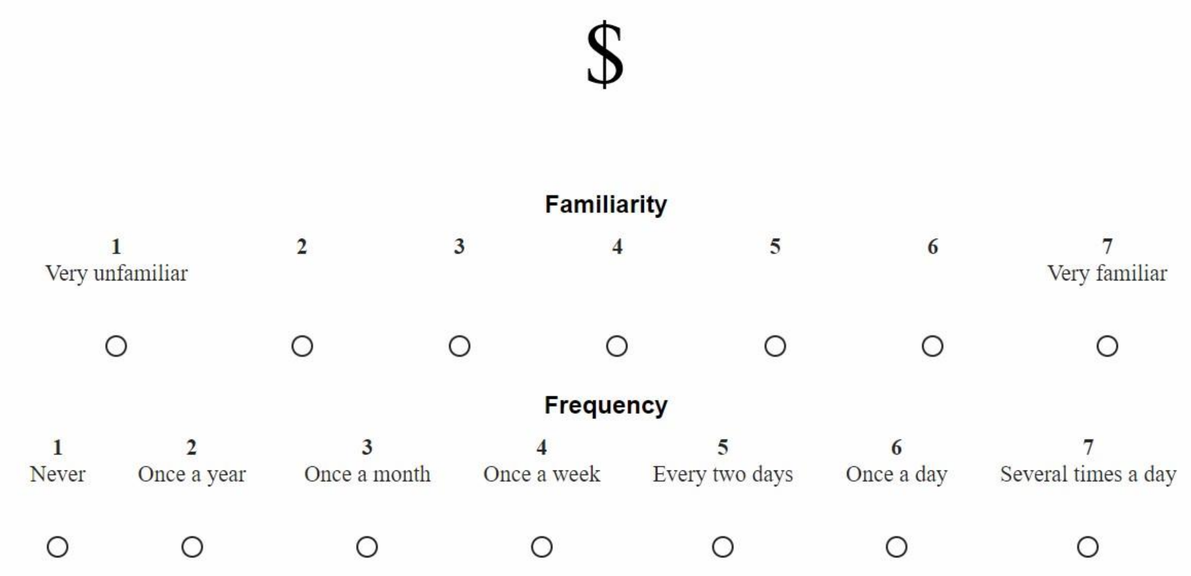
> “Symbols differ in how commonly or frequently they have been encountered. Some symbols are encountered very frequently, whereas other symbols are encountered infrequently. The purpose of this scale is to rate each symbol with respect to the frequency you encounter it. You should base your ratings according to the following 7-point scale: 1 = never, 2 = once a year, 3 = once a month, 4 = once a week, 5 = every two days, 6 = once a day, 7 = several times a day.”

Procedure. Following informed consent, participants were told that they would be rating a series of symbols, one at a time, on two different dimensions (familiarity and frequency). Participants were then provided with definitions of each scale prior to the rating task.

During the rating task, a symbol was presented in the center of the screen with each of the two scales below it, spanning horizontally across the screen (Figure 13). Participants responded by clicking on one of the radio buttons of each scale, at which point the experiment program automatically and immediately proceeded to the next rating trial. Following presentation of all 80 symbols, participants were asked whether their participation occurred under ‘ideal’ conditions (as in Experiment 4). They were then provided with a detailed feedback letter. The procedures and materials for this study were approved by the Office of Research Ethics at the University of Waterloo (project #41594). All data, pre-registrations, analysis code, experiment programs, and other materials are listed on the Open Science Framework (OSF; https://osf.io/e53z4/).

Figure 13

Familiarity and Frequency Rating Scales Used in Experiment 6



3.2.3 Results

Correlations with Average Familiarity Ratings. To determine whether memory performance for symbols correlated with average ratings of symbol meaning-based familiarity and subjective frequency, memory data from the four preceding experiments were aggregated. For each item, average proportions of free recall for symbols (in Experiments 1 and 4) and of old/new recognition hit rate for symbols (in Experiments 2 and 3) were collected.

In Experiments 2 and 3, half of the symbols were randomly presented at encoding. Because of this, hit rate for those experiments was calculated as number of hits divided by the number of times that the symbol was presented at encoding. In the end, memory performance data were aggregated from 579 participants across Experiments 1-4. Every symbol was rated on meaning-based familiarity and subjective frequency by each of the 337 naïve participants in the current experiment.

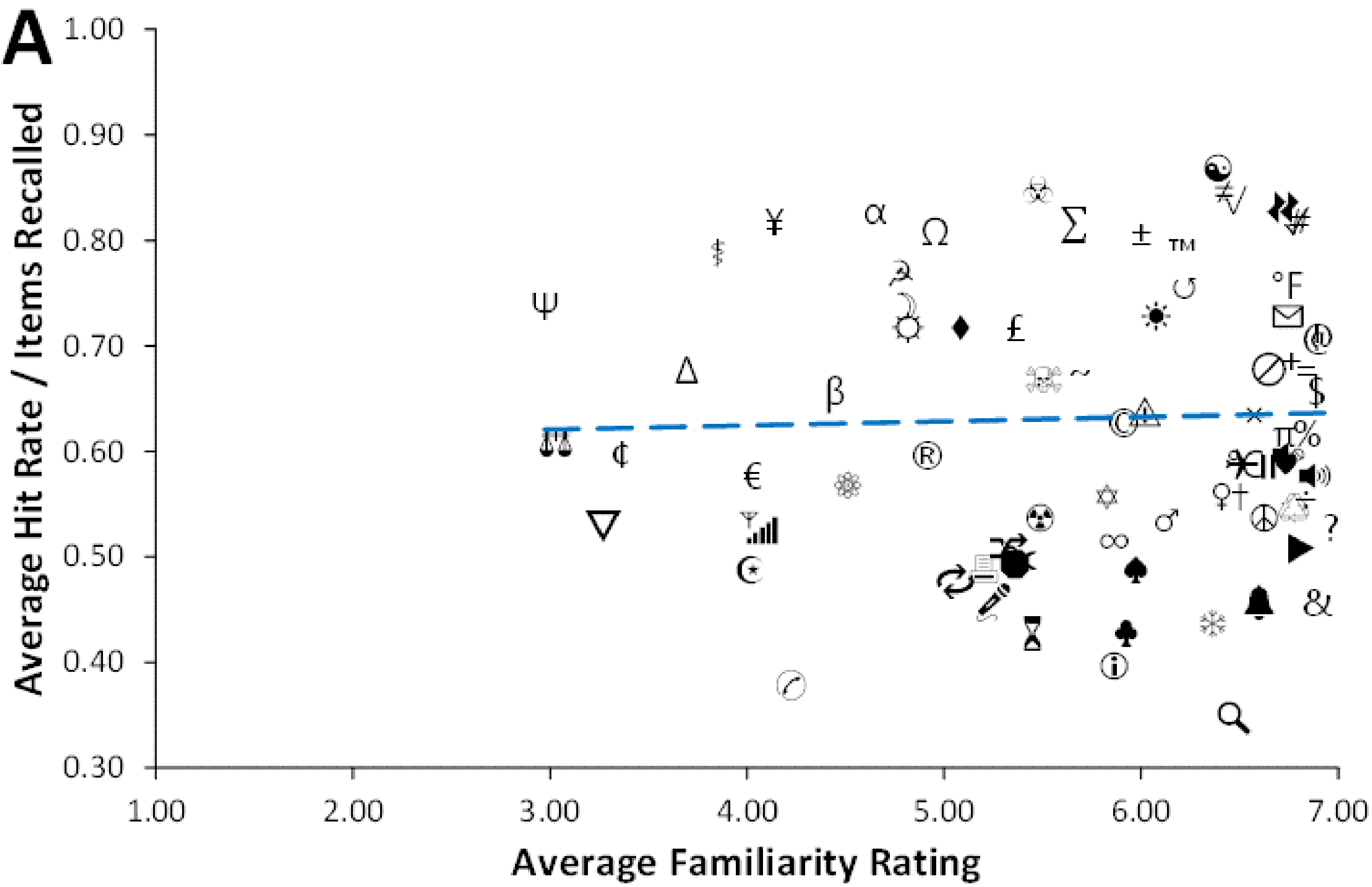
A series of Pearson correlations was conducted to compare memory performance, average ratings of meaning-based familiarity, and average ratings of subjective frequency for each of the 80 symbols. As expected, familiarity ratings for the current set of symbols along a 7-point scale were quite high (M = 5.74, SD = 1.42); frequency ratings were more moderate (M = 4.06, SD = 1.42). Average memory performance for symbols did not correlate with average familiarity ratings of symbols, $r(78) = .03$, $p = .776$ (Figure 14A), or with frequency of encountering symbols, $r(78) = .01$, $p = .945$.

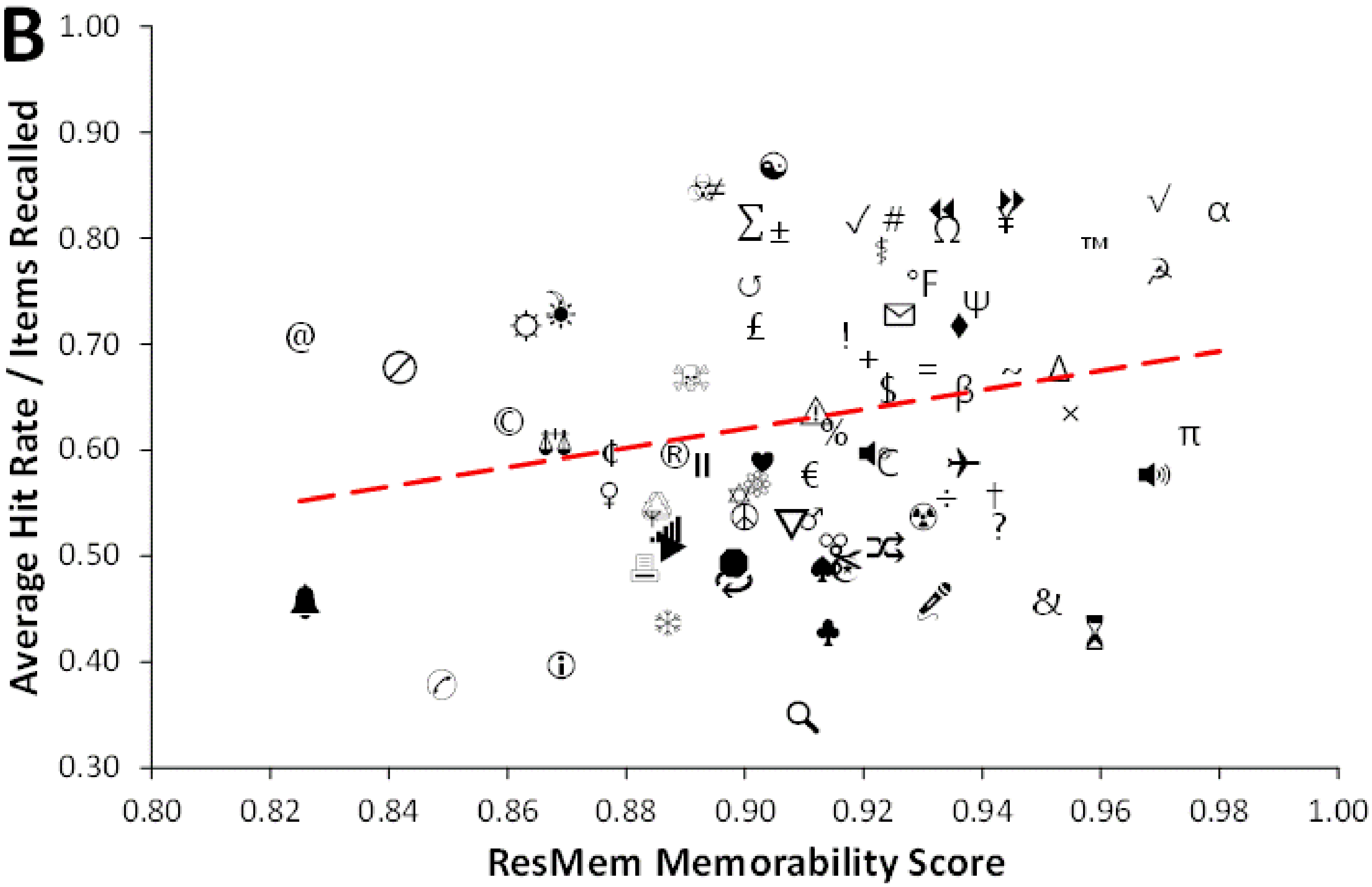
Predictors of Memory for Symbols. Submitting the full 80-item set of symbols to ResMem indicated that the stimuli that were used in the current experiments were highly memorable, with scores ranging from 0.83 to 0.98 (M = 0.91, SD = 0.03, on a scale from 0 to 1). It makes intuitive sense that scores would be rather high: Everyday symbols often are designed to be simple in terms of their low-level visual features, to be easily identifiable, and to be pervasive in visual communication media. Memorability estimates for symbols provided by ResMem, rather unsurprisingly, correlated positively with memory performance for symbols in Experiments 1-4, $r(78) = .23$, $p = .038$ (Figure 14B). When word versions of the same symbols set were input into ResMem, each item received a score of exactly 0.694, suggesting that the model could not differentiate memorability of words.

Finally, a hierarchical linear regression was conducted to determine whether the intrinsic memorability scores assigned to the symbols by ResMem could predict memory performance over and above contributions stemming from familiarity ratings or subjective frequency estimates. First, familiarity and frequency ratings for the set of symbols were entered into the model; second, scores from the ResMem network for the same symbols were added. Results indicated that the second model predicted significantly more variance than the first, $\Delta R^2 = .05$, $F(1, 76) = 4.17$, $p = .045$, suggesting that intrinsic memorability properties were still able to predict memory performance beyond familiarity ratings and frequency estimates. This analysis highlights that the inherent memorability of symbols, stemming from their perceptual qualities, predicts later memory beyond the combined contributions of one's personal familiarity with the symbols and the frequency with which these symbols are encountered.

Figure 14

Memory Performance Scores by Familiarity Rating and ResMem Score





Note. Memory performance on the y-axes of panels A and B, as well as on the x-axis of panel B, were on scales from 0 to 1 but are truncated here for the sake of clarity. Linear trendlines are included in each panel.

3.2.4 Discussion

This sixth and final experiment sought to assess whether familiarity, frequency, or inherent memorability influence memory for symbols. A dual coding explanation for picture superiority in memory rests entirely on the notion that images are more likely to be spontaneously labelled than words are to be imaged. This is also often touted as the reason that concreteness effects occur in memory: Concrete words are more easily imaged than abstract words. Based on this idea, the current experiment tested an extrapolation from dual coding theory—that as a person's familiarity with a symbol increases, so too should their likelihood of producing a mental label for that symbol. Hence, higher ratings of familiarity were expected to correlate with higher likelihoods of dual coding and therefore better memory performance for symbols.

My predictions, however, were not supported in this study: The correlation between familiarity and memory performance collapsed across the four preceding experiments was not statistically significant. Subjective estimates of frequency also failed to exhibit a significant relation with memory performance, suggesting that the number of encounters with a given symbol is perhaps inconsequential to memory. These findings are at odds with extrapolations from dual coding theory as they indicate that knowing what is being imaged, so as to obtain a verbal memory code, is not a determinant of later memory for graphic symbols. One might predict that without this knowledge, trying to remember a meaningless shape would prove difficult, much like the studies of abstract pictures mentioned previously. This prediction, however, was not supported in the present case. Overall, my data suggest that dual coding theory is perhaps limited to the extent that it cannot explain memory for graphic stimuli as a function of familiarity.

The use of ResMem here was motivated by the question of whether symbols have intrinsic properties that make them highly memorable. Using this neural network allowed me, via computational means, to quantify how memorable the stimulus set was and to determine whether the intrinsic memorability of the stimuli was separable from participants' ratings of familiarity and frequency of those same stimuli. As one might expect, the common everyday symbols used here were determined to be quite memorable by ResMem. The results of regression analyses suggest that this high inherent memorability contributes to memory performance beyond participants' familiarity with the symbols and their frequency of occurrence. Despite the analyses based on ResMem scores having been purely exploratory, these findings are in line with past literature showing that memorability is a perceptually-based phenomenon that is largely separable from any previous experience with the item (Bainbridge et al., 2017).

Chapter 4 – General Discussion

This dissertation investigated the cognitive mechanisms underlying representations of graphic symbols in memory. The major predictions—based on parallels drawn between symbols and pictures—were that symbols would be remembered better than words and that they would be remembered as well as pictures. These predictions, although rooted in Paivio's (1969, 1971) dual coding theory of memory, would have been quite similar if one were to begin instead with a distinctiveness account (Mintzer & Snodgrass, 1999). The core hypothesis is that graphic symbols behave like pictures in terms of their memorability: They serve as convenient ways to 'image' abstract words without relying on other concrete referents. Given this, memory for graphic symbols should match that of pictures, even if the pictures are of concrete (rather than abstract) objects. Hence, encoding of symbols should eliminate the often-observed performance decrement for abstract words that likely stems from their low imageability.

The first four experiments consistently demonstrated that graphic symbols are indeed better remembered than words. Experiment 1, using a within-subject design, demonstrated the basic finding that recall of symbols was greater than recall of their word counterparts. Experiment 2 showed that the effect generalized to a between-subjects design and to recognition testing. Experiment 3 showed that the symbol superiority effect (better memory for symbols than words) is robust, regardless of concreteness and set size of the word comparators. Experiment 4 pitted symbols, words, and pictures directly against each other and found that, even when pictures and words were of concrete objects, memory was still better for symbols than for words and was on par with memory for pictures.

In Experiment 5, a two-part study tested whether memory would differ for sport logos with and without integrated text relative to sports team names. This was done to test whether an explicit verbal code integrated into a logo would aid memory over and above the label thought to be spontaneously generated when viewing a recognizable logo. In the end, encoding logos led to memory superior to that for team names, and integrating text into a logo produced no further benefit to memory beyond text-less logos, but familiarity did moderate this relation.

Experiment 6 combined memory data from the first four experiments with new human ratings of familiarity from naïve participants and memorability scores derived using the ResMem deep neural network (Needell & Bainbridge, 2022). This experiment demonstrated that symbol superiority in memory is likely determined in part by the memorable visual characteristics that symbols possess, but surprisingly not by their familiarity. Next, an account for the present findings based in dual coding theory and distinctiveness explanations is presented before discussing other possible factors contributing to superior memory for symbols.

4.1 Dual Coding vs. Distinctiveness Explanations

From the outset, I reasoned that symbols may be processed like pictures. My prediction was originally couched in a dual coding explanation whereby symbols serve to concretize abstract words by providing otherwise absent visual referents. Although my logic was originally derived from Paivio's (1971) dual coding theory, distinctiveness accounts (Hamilton & Geraci, 2006; Mintzer & Snodgrass, 1999) of picture superiority are equally well supported by the data. According to more recent work (e.g., Ensor et al., 2019b), however, it could also be the case that whether dual coding or distinctiveness mechanisms are at play depends on the type of retrieval test used. Next, I consider this possibility in light of the present data.

Recall that Paivio's (1969, 1971) dual coding theory of memory posits that pictures are better remembered than words because pictures benefit from two codes in memory—verbal and image—whereas words rely on only a verbal code. Paivio's theory rests entirely on the notion that people are more likely to spontaneously label images than they are to spontaneously create mental images of the referents of words. Moreover, words that are more easily imageable will be more likely to be imagined automatically, thus increasing their probability of dual coding (and hence of better memory; Paivio et al., 1994). In this study, I predicted that symbols would be better remembered than words because, like pictures, symbols would elicit dual codes in memory. That is, upon being viewed, a symbol would first provide an image and thereafter would likely be spontaneously labelled. As a result, memory should be better for symbols relative to words, but memory performance for symbols and pictures should be equivalent.

Across the first four experiments, consistent support was found for these predictions. Graphic symbols were better remembered relative to (1) their word counterparts, (2) concrete nouns from a constrained set, and (3) concrete nouns representing highly imageable objects. The fourth experiment also confirmed that memory performance for symbols representing abstract concepts was equivalent to that for pictures of concrete objects. These results align precisely with dual coding theory for three primary reasons. First, memory was superior for symbols than for their abstract word counterparts, suggesting that—with semantics held constant—symbols were providing something additional to improve memory. What could this additional factor be? In line with dual coding theory, symbols are provided an extra image record in memory, as is the case for pictures. That Experiment 2 pointed to the encoding phase as the locus of the symbol superiority effect in memory also suggests that improved encoding—rather than eased retrieval—drives the benefit of symbols. The notion of enhanced encoding of symbols aligns with the prediction that they elicit dual codes in memory upon being viewed during the study phase. Finally, because symbols were as well remembered as line drawings of objects, the cognitive processes used to remember them likely are similar.

What is more, there is evidence that dual coding occurred in Experiments 1 and 4 in which a free recall test was used. Ensor et al. (2019b) suggested that the picture superiority effect might manifest due to dual coding on tests of free recall, as verbal reporting of recalled images would require access to the verbal label of the studied item (the 'logogen pathway', as Paivio, 1991 called it). To illustrate this point, consider that, in Experiments 1 and 4, similar sets of symbols were used but the words being compared differed. Experiment 1 used abstract words that were counterbalanced with the symbols set, whereas Experiment 4 used concrete words that were counterbalanced with the images used in that experiment. Comparing the effect size for symbol superiority in memory for each experiment, it was clearly larger in Experiment 1 ($d = 0.77$) than in Experiment 4 ($d = 0.27$). As shown in Table

1, this difference in memory outcomes was due to average word recall performance rising in Experiment 4 (from 0.32 to 0.41) whereas symbol recall performance was relatively unchanged in the two experiments (from 0.44 to 0.47). It stands to reason that the effect size in Experiment 4 was reduced as a result of word comparators now being concrete (rather than abstract, as was the case in Experiment 1). Paivio and Csapo (1969) would predict precisely this finding based on dual coding theory: concrete words are more likely than abstract ones to elicit spontaneous imagery, resulting in better memory (i.e., the concreteness effect). Thus, there is evidence indicating that dual coding was at play in my experiments where free recall testing was used.

Recall that the two main types of distinctiveness accounts concern conceptual (Hamilton & Geraci, 2006; Nelson et al., 1977) and physical (Ensor et al., 2019b; Mintzer & Snodgrass, 1999) aspects of an item. With respect to the conceptual version, symbols could be better remembered because, relative to words, there are fewer possible lures (i.e., distracting items) available for symbols. Intuitively, it is important that symbols be unique both in form and in meaning, lest their semantic or visual spaces become so overcrowded that their ability to represent general abstract concepts becomes diluted or ambiguous. For example, the poison symbol (☠) represents *all* poisons and consequently fails to distinguish important information such as which type of poison is being referenced, its current physical state of matter, or how exactly it is harmful. Some consumer product warning systems use combinations of symbols to circumvent this problem. For example, certain countries use a household hazardous waste warning system that combines common warning symbols (e.g., the poison symbol; ☠) but nests them within upside-down triangles, diamonds, or octagons to denote increasing levels of danger or states of matter. In this way, symbols can be combined with other symbols and/or shapes to enhance their specificity when required.

Nonetheless, symbols most often are used to convey information quickly and universally, and in so doing they are intentionally limited in their specificity and therefore in the number of semantic neighbors that they have. This absence of plausible lures could be what leads to conceptual distinctiveness, which in turn may drive the observed memory benefits for symbols. Indeed, even while conceptually broad there is little overlap between symbols. Continuing with the example of the poison symbol (☠), the closest plausible lures would perhaps be the biohazard (☣) and radiation (☢) symbols, as all three refer to dangerous materials that should not be mishandled. However, two critical aspects are apparent: (1) The symbols are all visually distinct from each other, and (2) they still have distinct underlying meanings even if they are potentially confusable with each other. Therefore, if one knows the meaning of the poison symbol but not of the biohazard or radiation symbols, the universality of the former should imply that the latter two symbols are not thought to *also* represent poison. Therefore, even in cases of semantic neighbours, the ubiquity of a symbol often means it is conceptually distinct from others.

This conceptual distinctiveness affords convenience, which could be a driving factor for the prevalent use of symbols in communication. For example, to express admiration, a simple heart symbol (♥) can be used rather than a lengthy statement. In fact, modern digital keyboards on cellphones often offer users suggestions of symbols that are relevant to the content that they are typing. Consequently, the conceptual distinctiveness of symbols offers a quick, convenient way to express complex abstract ideas. When insufficient or when more nuance is required, a symbol can easily be combined with written communication.

Consider as well that most symbols have few related items while words often have many, further exacerbating differences in conceptual distinctiveness. Returning to an example from the beginning of this dissertation, the word ‘play’ (in the context of watching television) has many related words, such as ‘begin,’ ‘start,’ ‘commence,’ etc., while the symbol for ‘play’ (▶) is semantically distinct from other symbols (i.e., no other symbol is used for the same purpose). Therefore, the symbol for play would be considered to possess high conceptual distinctiveness (Hamilton & Geraci, 2006) relative to its word counterpart.

The same symbols that lack semantic competitors could also be more distinct physically than their word counterparts. As noted previously, words in the English language are constructed from the same set of 26 letters, recycled repeatedly. So, while the physical form of words can vary depending on their underlying letters, length, or case, they are rather similar to each other in appearance. Symbols are instead an open-ended medium that can range from a single dot (i.e., the period symbol, ‘.’) to complex shapes that could be difficult to draw by hand (e.g., the biohazard symbol, ‘☣’). Hence not only are symbols more distinct than words semantically but they also vary more perceptually, making them easier to identify and distinguish from each other.

It seems plausible that if physical distinctiveness was driving memory performance in Experiment 4, then the picture stimuli should have elicited the best performance. While the symbols and pictures used in that experiment were sized identically, and both were black and white, the symbols typically shared many simple geometric shapes and had solid fills whereas the images—with their multiple lines, shapes, shading, and sometimes distinct sub-parts—were arguably more physically distinct from one another. And yet, performance was practically identical between symbols and pictures in Experiment 4. This result is consistent with the idea from Ensor et al. (2019b) that, because Experiment 4 used a free recall test, dual coding—rather than differences in physical distinctiveness—was likely driving the superior memory performance for symbols and pictures relative to words.

The results of Experiments 2 and 3, on the other hand, provide evidence more consistent with a distinctiveness explanation. In these two experiments, memory performance was measured using old/new recognition tests. Consequently, Ensor et al. (2019b) would contend that the picture superiority effect should result from increased distinctiveness for images rather than from dual coding. This is because, in these cases, access to an image code in memory is not required to determine whether the item was studied, as it provided for the participant during the memory decision. Therefore, the concreteness of the underlying concept should no longer affect memory in the case of recognition tests because there is no benefit of an additional image code for concrete words. This is precisely what I found: In comparing recognition performance for symbols vs. words in Experiments 2 and 3, the effect size in each case was identical (both d = 0.44), despite the word comparators being abstract and concrete, respectively. Thus, there was little evidence of dual coding at play in these particular experiments where recognition testing was used, suggesting that a different mechanism (physical distinctiveness, perhaps) was the main determinant of memory. Indeed, the intrinsically memorable visual properties of symbols highlighted in Experiment 6 further promote the idea that symbols may benefit memory as a result of physical distinctiveness.

Taken together, evidence from the present set of experiments is consistent with the idea that both dual coding and distinctiveness could be driving memory performance, and that which contributes most depends on the type of retrieval test used. Therefore, while this work was motivated by and is consistent with Paivio's dual coding theory, other explanations are indeed viable as well. Although the current data do not allow us to disentangle theories of picture superiority, two novel theoretical contributions emerged: (1) The picture superiority effect in memory extends to symbols, even though people may intuitively hesitate to classify symbols as images, and (2) that the concreteness of the underlying concept that symbols and pictures represent does not alter picture superiority in memory. This effect could be due to symbols offering unique visual referents for abstract words that are otherwise unlikely to be spontaneously imaged.

It is worth noting that although symbols can provide visual referents for abstract concepts, that does not necessarily mean that semantic activation of symbols is limited to the visual modality. Whereas Paivio's work involved image and verbal representations, other theories suggest that semantic processing may be distributed across discrete cortical networks. For instance, work by Martin and Chao (2001) contends that category-specific brain networks are activated similarly when a picture of an object or a word representing that object is presented. Related work on 'perceptual symbols theory' takes an embodied approach to cognition, suggesting that meaning is distributed not just across the brain but throughout the entirety of the body (Barsalou, 1999). Therefore, although symbols may provide a visual stimulus, it is possible that the semantic activation brought about by that stimulus could extend beyond the neural architecture underlying vision alone.

While my a priori reasoning was that symbols could be eliciting an image code for improved retention, for a verbal code to be provided one must also be familiar enough with a symbol to identify and label its imaged content. On top of that possibility, symbols also contain quite simple and recognizable visual features which could contribute to improved discriminability. As a result, both high-level familiarity and low-level visual attributes could play key roles in the memorability of symbols. Next, I summarize the evidence for this claim.

4.2 Implications for Studies of Familiarity and Memorability

Influence of Familiarity and Memorability

Other studies certainly have investigated memory for abstract visuospatial stimuli (e.g., Fernandes & Guild, 2009) but, in those studies, the items used were novel shapes or patterns with no meanings. Consequently, familiarity could have played little role in those studies. Because previous studies of images have investigated only the extremes of familiar stimuli by using pictures of easily recognizable objects (e.g., Paivio & Csapo, 1973) or semantically void patterns (e.g., Smith et al., 1990), no prior investigations have explored the influence of familiarity with regard to memory for pictures.

This study provided the first investigation of familiarity's influence on memory for picture-like stimuli. My prediction, based in dual coding theory, was that familiarity would correlate positively with memory for symbols. In the end, this prediction was not supported by the data: Familiarity was not consistently related to memory performance using within-subject comparisons in Experiments 1-3; using aggregate data with familiarity ratings from naïve participants in Experiment 6 confirmed the same result. Nonetheless, the results for familiarity reported here are important to consider for theories of picture superiority because—insofar as symbols are pictures—dual coding theory may be insufficient to explain the lack of a predicted relation between familiarity and memory for imaged content. The missing link between familiarity and memory for symbols suggests that knowing what a symbol *means* is not necessary to gain a memory benefit.

It may also be that symbols are simply efficient vehicles for conveying concepts, thanks to their visual properties. After all, symbols are often compact, use simple shapes and lines, and do not need color to be interpreted. To investigate this possibility, Experiment 6 made use of ResMem (Needell & Bainbridge, 2022), a newly built residual neural network that is capable of providing for any image a memorability score that corresponds to the likelihood of a person remembering it. The set of symbols used here received high memorability scores (M = 0.91 on a 0 to 1 scale). As expected, these memorability scores correlated significantly with memory performance for the current set of symbols. . Hierarchical regression showed that, even when accounting for familiarity and frequency, memorability scores still significantly predicted memory performance. Therefore, it could be the case that symbols contain inherent visual properties that boost memory, apart from any influence of the observer's knowledge or personal experiences. Although research on memorability of symbols is still in its early days, there are many parallels to be drawn to existing concepts of visual imagery and distinctiveness accounts of memory more generally.

Reconciling Picture Superiority with Inherent Memorability

Dual coding and distinctiveness accounts of picture superiority share at least one critical parallel: the concept of an image trace from dual coding theory and the notion of distinctiveness for pictures. Both accounts posit that there is something inherently special about the visual nature of images. In dual coding theory, pictures are thought to elicit an ‘image code’ and a ‘verbal code.’ In distinctiveness accounts, pictures are thought to be visually more distinct or diagnostic than words. But what about the low-level visual characteristics of images in relation to their memorability? Might image codes, visual distinctiveness, and intrinsic memorability represent different perspectives on the same mechanism?

Although the favorable visual aspects of symbols (and pictures) may be captured by the notion of an image trace, of physical distinctiveness, and of inherent memorability, each of these has yet another feature in common: meaning. In dual coding theory, this is the verbal trace, in distinctiveness accounts it is conceptual processing, and in memorability it is high-level semantic information such as category. Thus, in each case, encoding of low-level visual features is inherently tied to high-level semantic conceptualizations. One might assume that visual features drive higher-level abstractions of meaning—after all, one must perceive something before it can be identified and interpreted. There has been work, however, showing that top-down meaning-based processes can affect bottom-up perception in a visual search task (Lupyan & Spivey, 2008).

Emerging conceptualizations of memorability have also suggested a critical role for semantic information in determining later memory (e.g., Koch et al., 2020). For instance, a recent study by Kramer et al. (2022) gathered over a million human ratings of memorability and found that whereas semantic and visual dimensions accounted for a combined 35% of the variance in the inherent memorability of images, the vast majority of that variance (31%) was due to semantic predictors alone. This finding also aligns well with the relative contributions of perceptual and semantic features thought to underlie memorability predictions from the ResMem neural network (Needell & Bainbridge, 2022). Thus, the memorability of a graphic symbol may depend little on its low-level visual features—such as whether it is comprised of curved or straight lines. It is likely, however, that there exists interplay between the distinct visual and semantic facets represented in each individual symbol.

4.3 Implications for Models of Memory

From Atkinson and Shiffrin's (1968) multi-store model of memory, to Tulving's (1972) monohierarchical multimemory systems model, to Baddeley and Hitch's (1974) working memory model, there has been no shortage of grand cognitive theories attempting to encompass the onslaught of primary research that this field has generated. Unfortunately for cognitive modelers, I am only adding to the pile. I can, however, point the well-intentioned modeler in the right direction. To discuss implications of the present work on our understanding of memory more broadly, I turn to a contemporary model of memory: the Scale Invariant Memory and Perceptual Learning (SIMPLE) model (Neath & Brown, 2006).

In essence, the SIMPLE model of memory suggests that an item is discriminated as 'old' during retrieval if it sufficiently matches the stored representation more than other items in memory (i.e., if it can overcome competing interference). SIMPLE is a local distinctiveness model because it allows for near and far items to affect the distinctiveness of a given item differentially. 'Closer' items on any dimension (e.g., visual or semantic) will cluster together in recall, but may also cause interference that leads to forgetting. The 'dimensions' can be anything, but always include the temporal dimension (i.e., when the item was encoded). Therefore, it stands to reason that having distinct representations of items in memory will aid performance.

Indeed, one of the four main claims that SIMPLE makes is that memory works as a function of local distinctiveness. That is, performance on most memory tasks is determined by the level of interference from 'near' psychological constructs. Perhaps, then, both physical and conceptual distinctiveness attributes provide improved memory for symbols. As I argued previously, symbols are visually distinct from each other (relative to words, at least) and they also share few semantic neighbours. Thus, insofar as memory operates in accordance with the SIMPLE model, perhaps retrieval of symbols benefits as a result of distinct visual and semantic traces in memory (instituted at encoding), which provide high signal-to-noise ratios that aid in discrimination from competing items during retrieval.

That symbols elicit better memory than their semantically equated word counterparts suggests that it is the *visual* attributes of symbol that make them potent. For instance, take the word 'danger'. It has multiple anagrams that share the exact same letters (and therefore, also the same visual attributes): gander, garden, grande, and ranged. So, while none of these anagrams may crowd the word 'danger' conceptually, they do so physically. The symbol for danger (⚠), on the other hand, is quite unique in its visual characteristics. Surely one can think of other triangle-based symbols, but there is no other symbol that shares both a triangle and an exclamation point.

As mentioned at the beginning of this dissertation, symbols may also benefit from their relatively sparse semantic neighborhoods. Previously I had used the word ‘play’ as an example of a case whereby the word has many readily available substitutes, such as ‘begin,’ ‘start,’ ‘commence,’ while the symbol for ‘play’ (▶) has no substitute. As such, the symbol for play would benefit from high conceptual distinctiveness, while the word ‘play’ would not. I see this as a fair example but let me argue another point by returning to the just mentioned example of ‘danger’ (⚠).

In the case of the word danger, there are, of course, numerous available substitutes such as ‘peril’, ‘threat’, and ‘risk’. When considering the conceptual distinctiveness for the danger symbol, in this case there now are *also* many semantic neighbors: the symbols for poison (☠), biohazards (☣), radioactivity (☢), just to name a few. The difference driving memory performance for the danger symbol above that of the word danger is then perhaps left mostly to its visual features. Symbols are often highly distinctive along both their physical and conceptual dimensions, so when one of these dimensions fails to be distinctive, memory may be able to rely on the other dimension. Distinctive attributes of symbols—both along their conceptual and physical dimensions—therefore fit with the SIMPLE model of memory insofar as these dimensions serve as local distinctiveness characteristics that make symbols stand out memory. In short, I speculate that symbols are better remembered as a result of their reduced crowding along the visual and semantic ‘dimensions’, supporting the general idea that SIMPLE forwards: Local distinctiveness drives memory response accuracy.

Finally, the SIMPLE model assumes that the process of forgetting works as a function of interference alone; trace decay is assumed not to occur at all. Essentially, if one were to encode a list of items and then a delay were to ensue, those items will become less ‘distinct’ temporally because they cluster together ‘far’ away in time. As items get farther away in time, they become less distinct, therefore making discrimination of those far items more difficult, especially as new experiences constantly flood the senses. This process of continual, additive interference for old material in favor of new material causes ‘forgetting’ without any need for the notion of trace decay. Indeed, because SIMPLE assumes an interference-based account of forgetting, symbols should persist longer in memory as a result of their predicted isolation from competing information. As I argue above, a proponent of the SIMPLE model would predict that with fewer neighbors in semantic and visual space comes reduced interference and therefore better long-term retention. Future work could test this prediction empirically by comparing memory for symbols that are low on distinctiveness, as determined by their closely linked conceptual neighbours and visual traits (e.g., ▶, ⏩,⏪, 🔀), relative to those that are more conceptually and visually isolated (e.g., ⚠, 🔍, ?, 🔊).

4.4 Limitations and Future Directions

This study, being the first to investigate memory for graphic symbols, has several limitations. By my count, there are three minor limitations and one more moderate limitation. First, the experiment comparing pictures and symbols could be driven by item-selection effects. Second, the analyses of inherent stimulus memorability were purely exploratory. Third, I only ever presented symbols in isolation, one at a time. And finally, the cultural specificity of various symbols necessarily limits the generalizability of my findings.

While I spent much of this dissertation relating symbols to pictures theoretically, I made only one direct comparison to pictures—Experiment 4—and semantics were not equated between symbols and images in that experiment. Of course, as I have argued, it is necessarily difficult to come up with images for abstract words (e.g., peace) without imaging something else (e.g., white doves or olive branches). That is precisely where symbols are useful. Nonetheless, that symbols and pictures were not semantically equated means that there was a potential confound that was unique to Experiments 3 and 4: the possibility of item-selection effects. In a typical picture superiority effect study, pictures (e.g., a picture of a dog) are compared with corresponding words (e.g., the word 'dog'). By definition, however, there are no pictures that are semantically identical to symbols because the former depict physical things whereas the latter depict abstract concepts. This makes counterbalancing symbols and pictures impossible. It is conceivable, therefore, that because I could not counterbalance content between symbols and pictures, I may have inadvertently chosen a set of images that lead to particularly poor memory or, conversely, I may have selected a set of symbols that are especially memorable. A similar criticism could be levied against the fruit/vegetable words used in Experiment 3.

To combat this limitation, I opted to use concrete nouns and their exact associated images to give pictures and words the best chance of being remembered. Yet, I still found the predicted pattern of results such that memory performance for symbols and pictures was practically identical and that both were superior to words. In prior work that compared memory for words and sounds, Ensor et al. (2019a) ran into a similar item-selection issue. Following a series of experiments, they concluded that item-selection effects were not likely biasing their results. Given the strong arguments against item-selection effects presented by Ensor et al. (2019a) when in a similar predicament, and that I chose stimuli that should bolster picture and word memory, I do not believe item selection effects were the main determinant of symbols superiority in memory found here.

A second minor limitation is that my analyses based on memorability were purely exploratory, limiting the ability to draw strong conclusions from a priori reasoning. In addition, the ResMem neural network that I used had not been pre-trained on symbols, so the scores provided for my stimuli may not be accurate. As noted previously, however, ResMem has been trained on highly visually similar images with black and white shading. Nonetheless, my findings offer a first look into the intrinsic memorability of symbols and provide insight into the factors underlying picture superiority in memory. Further confirmatory research is required to solidify the conclusions drawn here.

Another minor limitation is that the generalizability of my findings is limited to symbols presented in isolation. While this is likely the naturalistic environment for many of the symbols in my set (e.g., ☣), many other symbols are used in written communication (e.g., !, +, ?, $, %, &). Therefore, it is possible that improved memory for this latter class of symbols may not generalize when they appear in sentence contexts. It is my view that the applicability of the newfound symbol superiority should be applied directly to wherever symbols can be effectively used. Hence there is a wide avenue of future research available to be explored.

Finally, the largest limitation of this work is perhaps the limit that I must place on the generalizability of my findings. Symbols are derived from culture, so the specificity of that culture is of great importance. Consider that, even within highly similar cultures, symbols can be misaligned with humorous results: the word ‘thong’ in North America refers to a type of underwear, whereas in Australia the same word refers to a type of sandal. What’s more, something can start as one type of sign and evolve into another. For instance, the floppy disk (🖫) began as a reminder to the computer operator to insert a disk to save what they were working on. Because of its physical resemblance to the real thing, the floppy disk would have been considered to be an icon by Peirce’s definition. When floppy disks were still in recent memory but not used for saving, the floppy disk represented the old saving action while no longer referring to a physical object, and so it became an index. Nowadays, it can best be described as a symbol, as its physical and logical associations are no longer known to most users. Symbols can also become prominent rather quickly: The ‘hamburger’ icon (☰) is well-known for its use in modern mobile applications to denote a menu, yet it did not see widespread use until 2009 and therefore may be unfamiliar to older participants or to those lacking access to modern technologies like cell phones.

As a result of the cultural specificity of graphic symbols, the set of symbols used in the present study necessarily limits the generalizability of my findings. The symbols that I used were chosen precisely because they would be known to participants in Canada and the

US. Because symbols are often culturally specific, it is possible that my set of symbols would not lead to replicable effects in other cultures. For example, if the meanings of studied symbols were entirely unknown to participants in another culture, dual coding theory would predict that the memory boost seen for the symbols used here would disappear as a result of the lost verbal code. The findings of my current studies do, however, cast doubt on this prediction. Perhaps a more pertinent cultural consideration is that written languages vary in the degree to which they elicit verbal or visuospatial processing. Dual coding theory would predict that, when comparing symbols (e.g., ☠) to word counterparts written in character-based languages that elicit visuospatial processing, like Chinese (e.g., 毒; Fernandes et al., 2013), the memory benefit for symbols may be eliminated. In short, I speculate that words written in character-based languages may evoke dual memory representations, similar to those of graphic symbols.

Finally, my set of 'symbols' contained icons, indices, and symbols (by Peirce's 1867 definitions) and therefore it is possible that one or more of these sub-types is driving or constraining the effects reported here. For instance, it is possible that associations built on physical or logical connections (as are the cases in Peirce's 'icons' and 'indices', respectively) could behave in memory in a qualitatively different manner than culturally derived signs (i.e., Peirce's 'symbols').

Future Directions

Future studies of symbolic representations in memory (and in cognition more generally) should, in my view, pursue three main facets: applicability, diversity, and neural underpinnings. By applicability, I refer to the real-world use of symbols. By exploring this, one could determine whether symbol superiority is maintained when symbols are embedded in their respective naturalistic environments, such as in sentence contexts, and further, whether the contents of those sentences are also better remembered by association (e.g., John paid $5 for his lemonade vs. John paid 5 dollars for his lemonade). Emerging research in this area has suggested that emojis and brand logos embedded in sentences in a similar way slow reading times but do not diminish sentence comprehension (Cohn et al., 2018). Similarly, emojis placed in sentence-end positions can slow reading times if the emoji is incongruent with the semantic content of the preceding sentence (Barach et al., 2021).

Beyond memory research, it would also be important to consider the efficiency with which symbols convey information relative to other modes of visual communication like words or images. For instance, would the symbol for ‘stop’ (⯃) lead to faster response times than the word ‘stop’ or an image of someone halted at an intersection? While this is just one brief example, there is vast potential for scientific explorations that yield real-world developments. Relatedly, the type of association a symbol has with its referent may impact its treatment by cognition.

In his 1997 book, biological anthropologist Terrance Deacon presented an account of the co-evolution of human language and the brain. In it, he reasoned that human language is unique insofar as we have learned to use symbols in the absence of more literal referents.

> “Stone and symbolic tools, which were initially acquired with the aid of flexible ape-learning abilities, ultimately turned the tables on their users and forced them to adapt to a new niche opened by these technologies. Rather than being just useful tricks, these behavioral prostheses for obtaining food and organizing social behaviors became indispensable elements in a new adaptive complex. The origin of

"humanness" can be defined as that point in our evolution where these tools became the principal source of selection on our bodies and brains. It is the diagnostic trait of *Homo symbolicus*". (Deacon, 1997, p. 345)

Deacon's account constitutes a hierarchy of communication media that, he argues, tracks with evolutionary increases in brain mass. Essentially, he claims that to master symbolic communication, we must first have mastered indexical communication. And to master indexical communication, we must first have mastered iconic communication.

Take the example of a primitive group of humans that had just survived their first attack by a bear. They quickly learn the 'iconic' association that any bear they see in the future should be considered dangerous. Icons in this case could eventually even take the form of a painting of a bear outside caves where they had been spotted. Then, with more time and experience, these early humans could develop an 'indexical' association between the bear and its paw prints; they should be worried if they see tracks around camp. Finally, they develop a 'symbolic' association between a culturally derived sign and the danger that the bear poses. This could be in the form of a written word, a drawn symbol, or even shouting the primitive word for 'bear' to warn others that danger is nearby. It is precisely these logical and material differences in representational formats that I am referring to as the diversity of symbols.

While numerous types of symbols have developed naturally over millennia, and have been outlined by Peirce among others, most have not been explored in terms of their potential differential effects on cognition. Could it be that icons, as arguably the oldest forms of symbols—and with their reference to real-world objects—provide the strongest memory benefit? Could it be that indices, with their A-B logical associations, lead to quicker interpretation? What about Peircean symbols and their culturally derived connotations: Could variations in interpretation affect their cognitive representations? By exploring the influence that different types of symbols and their contexts have on cognition, we can eventually learn to develop safer traffic signs, more effective advertisements, or even increasingly efficient communication systems. Crucially, the various levels of abstraction present across types of symbolic associations could inform related work, such as that on numerical cognition, where it is hotly debating whether there is a unitary centre in the brain for processing abstract representations.

While low-level perceptual characteristics of images undoubtedly contribute to their memorability (Bainbridge et al., 2013), recent work has suggested that high-level semantic features may be even more influential (Needell & Bainbridge, 2022). In future experiments, therefore, I will explore the route that symbolic meaning takes in the brain. Participants will study a list of common symbols or their associated words/labels for a later memory test while in an fMRI scanner. During this phase, participants will be asked to press a button if they understand what each symbol or word means, promoting engagement with the meaning of the stimuli. Critically, the symbols will vary in the degree to which they exhibit physical (e.g., ✈), logical (e.g., 🔄), or artificial (e.g., ▶) associations with their referents. Afterward, a multivoxel pattern similarity analysis could be used to compare neural instantiations of symbol- and word-based depictions of the same abstract concepts.

My work presented here suggests that symbols and words may share semantic representations but differ in degree of visualization. If my hypothesis is correct, a similarity analysis should reveal that words and symbols elicit activation overlap in semantic association areas like the hippocampus (Binder et al., 2009), ventromedial prefrontal cortex, and angular gyrus (Gilboa & Marlatte, 2017), indicative of integration with existing schemas. Symbols, however, should also evoke activity in medial temporal sites linked to visual identification in the ventral ('what') visual stream—especially if they share physical or logical associations with their referents (Goodale & Milner, 1992).

This planned experiment will not only indicate whether words and symbols are similar in activation of meaning, but will also permit delineation of the different *types* of associations (physical, logical, or artificial) that abstract concepts can have with their visual referents and whether these semantic associations influence later memory performance. Symbols are known to participants and highly integrated in memory, yet they differ in shape, meaning, and association type. As a result, these efforts will offer unique insights into how perceptual attributes and semantic associations support memorability more broadly. Through planned future work, we will come to better understand how the creation of one of the oldest forms of communication—visual symbols—comes to pass and, more generally, how neural mechanisms support symbolic representations and intrinsic memorability in the brain.

As this line develops, I can answer other theoretically motivated questions: Do rich visual features like colour interact with existing schemas to alter memorability of images? For example, would displaying the peace symbol (☮) in a negative valence colour like red as opposed to white change its cognitive representation? Real-world applications are equally abundant (e.g., optimizing traffic signs, brand logos, and workplace safety). Much like the rapid adoption and standardization of the biohazard symbol (☣) following a 1960s study

(Baldwin & Runkle, 1967), this work will serve as the basis for newly created symbols to follow.

4.5 Conclusions

This dissertation has sought to address how common symbols are processed in human cognition. The major hypothesis is that symbols serve to concretize abstract concepts. As a result, it was predicted that memory for symbols should be superior to memory for words. As well, if symbols truly are akin to pictures, then memory for these two types of stimuli should be equivalent. Across six experiments, these predictions were confirmed: Symbols and logos were indeed better remembered than words and memory for them did not differ from that for pictures. These findings remained stable in the face of changes to experiment setting, study design, retrieval test type, and even the nature of word-based comparators. Overall, this dissertation was the first to show that symbols and logos may be processed in memory as 'mini pictures', garnering all the encoding benefits thought to underlie superior memory for full-size images as well.

A final experiment showed that the memorable properties that symbols possess predicted performance beyond ratings of familiarity and frequency. Thus, I conclude that graphic symbols likely are processed distinctly from words but that there seem to be moderating effects of the visual properties inherent in the design of symbols. A major conclusion of this dissertation is the powerful nature of image-based encoding that most graphic symbols share. Future work on symbols will elucidate factors underlying memory, cognitive faculties for abstract representation, and optimal techniques for efficient visual communication. While philosophers have debated the definition and use of symbols for centuries, modern psychology allows us to truly disentangle how symbols are represented in the mind.

References

Amrhein, P. C., McDaniel, M. A., & Waddill, P. (2002). Revisiting the picture-superiority effect in symbolic comparisons: Do pictures provide privileged access? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(5), 843–857. https://doi.org/10.1037/0278-7393.28.5.843

Ansari, D., Garcia, N., Lucas, E., Hamon, K., & Dhital, B. (2005). Neural correlates of symbolic number processing in children and adults. *NeuroReport*, *16*(16), 1769–1773. https://doi.org/10.1097/01.wnr.0000183905.23396.f1

Atkin, A. (2013). Peirce's theory of signs. *The Stanford encyclopedia of philosophy* (Summer 2013 Edition). https://plato.stanford.edu/archives/sum2013/entries/peirce-semiotics/

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In *The psychology of learning and motivation* (2nd ed., pp. 89–195). Academic Press.

Baddeley, A. D., & Hitch, G. (1974). Working Memory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 8, pp. 47–89). Academic Press. https://doi.org/10.1016/S0079-7421(08)60452-1

Bainbridge, W. A. (2019). Memorability: How what we see influences what we remember. In *Psychology of Learning and Motivation* (Vol. 70, pp. 1–27). Elsevier. https://doi.org/10.1016/bs.plm.2019.02.001

Bainbridge, W. A. (2020). The resiliency of image memorability: A predictor of memory separate from attention and priming. *Neuropsychologia*, *141*. https://doi.org/10.1016/j.neuropsychologia.2020.107408

Bainbridge, W. A. (2021). *Memorability: Reconceptualizing Memory as a Visual Attribute*. PsyArXiv. https://doi.org/10.31234/osf.io/ckv92

Bainbridge, W. A. (2022). Shared Memories Driven by the Intrinsic Memorability of Items. In B. Ionescu, W. A. Bainbridge, & N. Murray (Eds.), *Human Perception of Visual Information* (pp. 183–206). Springer International Publishing. https://doi.org/10.1007/978-3-030-81465-6_7

Bainbridge, W. A., Dilks, D. D., & Oliva, A. (2017). Memorability: A stimulus-driven perceptual neural signature distinctive from memory. *NeuroImage*, *149*, 141–152. https://doi.org/10.1016/j.neuroimage.2017.01.063

Bainbridge, W. A., Isola, P., & Oliva, A. (2013). The intrinsic memorability of face photographs. *Journal of Experimental Psychology: General*, *142*(4), 1323–1334. https://doi.org/10.1037/a0033872

Bainbridge, W. A., & Rissman, J. (2018). Dissociating neural markers of stimulus memorability and subjective recognition during episodic retrieval. *Scientific Reports*, *8*(1), 8679. https://doi.org/10.1038/s41598-018-26467-5

Baldwin, C. L., & Runkle, R. S. (1967). Biohazards symbol: Development of a biological hazards warning signal. *Science*, *158*(3798), 264–265. http://www.jstor.org/stable/1722446

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology. General*, *133*(2), 283–316. https://doi.org/10.1037/0096-3445.133.2.283

Balota, D. A., Pilotti, M., & Cortese, M. J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory & Cognition*, *29*(4), 639–647. https://doi.org/10.3758/BF03200465

Barach, E., Feldman, L. B., & Sheridan, H. (2021). Are emojis processed like words?: Eye movements reveal the time course of semantic processing for emojified text. *Psychonomic Bulletin & Review*, *28*(3), 978–991. https://doi.org/10.3758/s13423-020-01864-y

Barrett, T. J. (1985). Interactive imagery and recall of advertisements: Interactive imagery, noninteractive imagery, and printed text. *Psychological Reports*, *56*(3), 922–922. https://doi.org/10.2466/pr0.1985.56.3.922

Barsalou L. W. (1999). Perceptual symbol systems. *The behavioral and brain sciences*, *22*(4), 577–660. https://doi.org/10.1017/s0140525x99002149

Bellhouse-King, M. W., & Standing, L. G. (2007). Recognition memory for concrete, regular abstract, and diverse abstract pictures. *Perceptual and Motor Skills*, *104*(3), 758–762. https://doi.org/10.2466/pms.104.3.758-762

Bergstrom, J. A. (1893). Experiments upon physiological memory by means of the interference of associations. *The American Journal of Psychology*, *5*(3), 356. https://doi.org/10.2307/1410999

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, *15*(11), 527–536. https://doi.org/10.1016/j.tics.2011.10.001

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767–2796. https://doi.org/10.1093/cercor/bhp055

Bjork, R. A. (1975). Retrieval as a memory modifier: An interpretation of negative recency and related phenomena. In R. L. Solso (Ed.), *Information Processing and Cognition: The Loyola Symposium* (pp. 123–144). Lawrence Erlbaum.

Blake, A. B., & Castel, A. D. (2019). Memory and availability-biased metacognitive illusions for flags of varying familiarity. *Memory & Cognition*, *47*(2), 365–382. https://doi.org/10.3758/s13421-018-0872-y

Blake, A. B., Nazarian, M., & Castel, A. D. (2015). The apple of the mind's eye: Everyday attention, metamemory, and reconstructive memory for the apple logo. *Quarterly Journal of Experimental Psychology*, *68*(5), 858–865. https://doi.org/10.1080/17470218.2014.1002798

Brady, T. F., & Bainbridge, W. A. (2022). *Visual Memory* (1st edition). Routledge.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

Calkins, M. W. (1898). Short studies in memory and in association from the Wellesly College psychological laboratory. *Psychological Review*, *5*(5), 451–462. https://doi.org/10.1037/h0071176

Carreiras, M., Armstrong, B. C., Perea, M., & Frost, R. (2014). The what, when, where, and how of visual word recognition. *Trends in Cognitive Sciences*, *18*(2), 90–98. https://doi.org/10.1016/j.tics.2013.11.005

Castel, A. D., Vendetti, M., & Holyoak, K. J. (2012). Fire drill: Inattentional blindness and amnesia for the location of fire extinguishers. *Attention, Perception, & Psychophysics*, *74*(7), 1391–1396. https://doi.org/10.3758/s13414-012-0355-3

Cattell, J. M. (1886). The time it takes to see and name objects. *Mind*, *os-XI*(41), 63–65. https://doi.org/10.1093/MIND/OS-XI.41.63

Childers, T. L., & Houston, M. J. (1984). Conditions for a picture-superiority effect on consumer memory. *Journal of Consumer Research*, *11*(2), 643–654. https://doi.org/10.1086/209001

Cohen Kadosh, R., Cohen Kadosh, K., Schuhmann, T., Kaas, A., Goebel, R., Henik, A., & Sack, A. T. (2007). Virtual dyscalculia induced by parietal-lobe TMS impairs automatic magnitude processing. *Current Biology: CB*, *17*(8), 689–693. https://doi.org/10.1016/j.cub.2007.02.056

Cohen Kadosh, R., Lammertyn, J., & Izard, V. (2008). Are numbers special? An overview of chronometric, neuroimaging, developmental and comparative studies of magnitude representation. *Progress in Neurobiology*, *84*(2), 132–147. https://doi.org/10.1016/j.pneurobio.2007.11.001

Cohn, N., Roijackers, T., Schaap, R., & Engelen, J. (2018). Are emoji a poor substitute for words? Sentence processing with emoji substitutions. *Proceedings of the 40th Annual*

Conference of the Cognitive Science Society (pp. 1524–1529). Cognitive Science Society. http://www.cognitivesciencesociety.org/conference/cogsci-2018/

Coltheart, V. (1993). Effects of phonological similarity and concurrent irrelevant articulation on short-term-memory recall of repeated and novel word lists. *Memory & Cognition*, *21*(4), 539–545. https://doi.org/10.3758/BF03197185

Coppens, L., Verkoeijen, P., & Rikers, R. (2011). Learning Adinkra symbols: The effect of testing. *Journal of Cognitive Psychology*, *23*(3), 351–357. https://doi.org/10.1080/20445911.2011.507188

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

da Costa e Silva, T. (2019). *The logic of design process: Invention and discovery in light of the semiotics of Charles S. Peirce*. Transcript-Verlag.

Deacon, T. (1997). *The symbolic species*. W.W. Norton & Co.

Delacre, M., Lakens, D., & Leys, C. (2017). Why psychologists should by default use Welch's t-test instead of Student's t-test. *International Review of Social Psychology*, *30*(1), 92–101. https://doi.org/10.5334/irsp.82

Delacre, M., Leys, C., Mora, Y. L., & Lakens, D. (2020). Taking parametric assumptions seriously: Arguments for the use of Welch's F-test instead of the classical F-test in one-way ANOVA. *International Review of Social Psychology*, *32*(1), 1–12. https://doi.org/10.5334/IRSP.198

Deloache, J. S. (2004). Becoming symbol-minded. *Trends in Cognitive Sciences*, *8*(2), 66–70. https://doi.org/10.1016/j.tics.2003.12.004

Deng, J., Li, L.-J., Li, K., Li, F. F., Dong, W., Socher, R., Fei-Fei, L., & Kai Li. (2009). ImageNet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248–255. https://doi.org/10.1109/cvpr.2009.5206848

Diester, I., & Nieder, A. (2007). Semantic associations between signs and numerical categories in the prefrontal cortex. *PLOS Biology*, *5*(11), e294. https://doi.org/10.1371/journal.pbio.0050294

Ding, J., Liu, W., & Yang, Y. (2017). The influence of concreteness of concepts on the integration of novel words into the semantic network. *Frontiers in Psychology*, *8*. https://www.frontiersin.org/article/10.3389/fpsyg.2017.02111

Ensor, T. M., Bancroft, T. D., & Hockley, W. E. (2019a). Listening to the picture-superiority effect: Evidence for the conceptual-distinctiveness account of picture superiority in recognition. *Experimental Psychology*, *66*(2), 134-153. http://dx.doi.org/10.1027/1618-3169/a000437

Ensor, T. M., Surprenant, A. M., & Neath, I. (2019b). Increasing word distinctiveness eliminates the picture superiority effect in recognition: Evidence for the physical-distinctiveness account. *Memory & Cognition*, *47*(1), 182–193. https://doi.org/10.3758/s13421-018-0858-9

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

Fernandes, M., & Guild, E. (2009). Process-specific interference effects during recognition of spatial patterns and words. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *63*(1), 24–32. https://doi.org/10.1037/a0012870

Fliessbach, K., Weis, S., Klaver, P., Elger, C. E., & Weber, B. (2006). The effect of word concreteness on recognition memory. *NeuroImage*, *32*(3), 1413–1421. https://doi.org/10.1016/j.neuroimage.2006.06.007

Goetschalckx, L., Moors, P., & Wagemans, J. (2018). Image memorability across longer time intervals. *Memory*, *26*(5), 581–588. https://doi.org/10.1080/09658211.2017.1383435

Goetschalckx, L., & Wagemans, J. (2019). MemCat: A new category-based image set quantified on memorability. *PeerJ, 7*(e8169). https://doi.org/10.31234/osf.io/64xfa

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*(1), 20–25. https://doi.org/10.1016/0166-2236(92)90344-8

Goodman, N. (1976). *Languages of Art: An Approach to a Theory of Symbols*. Hackett Publishing.

Hamilton, M., & Geraci, L. (2006). The picture superiority effect in conceptual implicit memory: A conceptual distinctiveness hypothesis. *The American Journal of Psychology*, *119*(1), 1. https://doi.org/10.2307/20445315

Hamilton, M., & Rajaram, S. (2001). The concreteness effect in implicit and explicit memory tests. *Journal of Memory and Language*, *44*(1), 96–117. https://doi.org/10.1006/jmla.2000.2749

Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of d′. *Behavior Research Methods, Instruments, & Computers*, *27*(1), 46–51. https://doi.org/10.3758/BF03203619

Henderson, P. W., & Cote, J. A. (1998). Guidelines for selecting or modifying logos. *Journal of Marketing*, *62*(2), 14–30. https://doi.org/10.2307/1252158

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.

Hulme, C., Roodenrys, S., Schweickert, R., Brown, G. D. A., Martin, S., & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a redintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(5), 1217–1232. https://doi.org/10.1037/0278-7393.23.5.1217

Hunt, R. R. (2013). Precision in memory through distinctive processing. *Current Directions in Psychological Science*, *22*(1), 10–15. https://doi.org/10.1177/0963721412463228

Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory and Language*, *32*(4), 421–445. https://doi.org/10.1006/jmla.1993.1023

Huttenlocher, J., & Higgins, E. T. (1978). Issues in the study of symbolic development. In W. A. Collins (Ed.), *Minnesota symposia on child psychology: Vol. Vol. 11* (pp. 98–138).

Erlbaum. https://www.taylorfrancis.com/chapters/mono/10.4324/9780203780961-8/issues-study-symbolic-development-collins

Isola, P., Jianxiong, X., Parikh, D., Torralba, A., & Oliva, A. (2014). What makes a photograph memorable? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *36*(7), 1469–1482. https://doi.org/10.1109/TPAMI.2013.200

Isola, P., Xiao, J., Torralba, A., & Oliva, A. (2011). What makes an image memorable? *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 145–152.

Jacoby, L. L. (1983). Perceptual enhancement: Persistent effects of an experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(1), 21–38. https://doi.org/10.1037/0278-7393.9.1.21

Jessen, F., Heun, R., Erb, M., Granath, D. O., Klose, U., Papassotiropoulos, A., & Grodd, W. (2000). The concreteness effect: Evidence for dual coding and context availability. *Brain and Language*, *74*(1), 103–112. https://doi.org/10.1006/brln.2000.2340

Joshi, G., Walambe, R., & Kotecha, K. (2021). A review on explainability in multimodal deep neural nets. *IEEE Access*, *9*, 59800–59821. https://doi.org/10.1109/access.2021.3070212

Juhasz, B. J., Lai, Y.-H., & Woodcock, M. L. (2015). A database of 629 English compound words: Ratings of familiarity, lexeme meaning dominance, semantic transparency, age of acquisition, imageability, and sensory experience. *Behavior Research Methods*, *47*(4), 1004–1019. https://doi.org/10.3758/s13428-014-0523-6

Kassambara, A. (2021). *rstatix: Pipe-friendly framework for basic statistical tests*.

Khanna, M. M., & Cortese, M. J. (2021). How well imageability, concreteness, perceptual strength, and action strength predict recognition memory, lexical decision, and reading aloud performance. *Memory (Hove, England)*, *29*(5), 622–636. https://doi.org/10.1080/09658211.2021.1924789

Khosla, A., Raju, A. S., Torralba, A., & Oliva, A. (2015). Understanding and predicting image memorability at a large scale. *2015 IEEE International Conference on Computer Vision (ICCV)*, 2390–2398. https://doi.org/10.1109/ICCV.2015.275

Kirkpatrick, E. A. (1894). An experimental study of memory. *Psychological Review*, *1*, 602–609. https://doi.org/10.1037/h0068244

Koch, G. E., Akpan, E., & Coutanche, M. N. (2020). Image memorability is predicted by discriminability and similarity in different stages of a convolutional neural network. *Learning & Memory*, *27*(12), 503–509. https://doi.org/10.1101/lm.051649.120

Kramer, M. A., Hebart, M. N., Baker, C. I., & Bainbridge, W. A. (2022). *The Features Underlying the Memorability of Objects*. bioRxiv. https://doi.org/10.1101/2022.04.29.490104

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, *25*. https://proceedings.neurips.cc/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html

Kronbichler, M., Hutzler, F., Wimmer, H., Mair, A., Staffen, W., & Ladurner, G. (2004). The visual word form area and the frequency with which words are encountered: Evidence from a parametric fMRI study. *NeuroImage*, *21*(3), 946–953. https://doi.org/10.1016/j.neuroimage.2003.10.021

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge University Press. https://doi.org/10.1017/CBO9781139087759

Lloyd, E. C., Shehzad, Z., Schebendach, J., Bakkour, A., Xue, A. M., Assaf, N. F., Jilani, R., Walsh, B. T., Steinglass, J., & Foerde, K. (2020). Food folio by Columbia Center for Eating Disorders: A freely available food image database. *Frontiers in Psychology*, *11*. https://www.frontiersin.org/articles/10.3389/fpsyg.2020.585044

Lupyan, G., & Spivey, M. J. (2008). Perceptual processing is facilitated by ascribing meaning to novel stimuli. *Current Biology*, *18*(10), R410–R412. https://doi.org/10.1016/j.cub.2008.02.073

Lyons, I. M., & Ansari, D. (2009). The cerebral basis of mapping nonsymbolic numerical quantities onto abstract symbols: An fMRI training study. *Journal of Cognitive Neuroscience*, *21*(9), 1720–1735. https://doi.org/10.1162/jocn.2009.21124

Maisto, A. A., & Queen, D. E. (1992). Memory for pictorial information and the picture superiority effect. *Educational Gerontology*, *18*(2), 213–223. https://doi.org/10.1080/0360127920180207

Makowski, D. (2021). *psycho: Efficient and publishing-oriented workflow for psychological science*.

Marmie, W. R., & Healy, A. F. (2004). Memory for common objects: Brief intentional study is sufficient to overcome poor recall of US coin features. *Applied Cognitive Psychology*, *18*(4), 445–453. https://doi.org/10.1002/acp.994

Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: Structure and processes. *Current Opinion in Neurobiology*, *11*(2), 194–201. https://doi.org/10.1016/s0959-4388(00)00196-3

McDaniel, M. A., & Pressley, M. (1987). *Imagery and related mnemonic processes: Theories, individual differences, and applications* (M. A. McDaniel & M. Pressley, Eds.). Springer-Verlag Publishing. https://doi.org/10.1007/978-1-4612-4676-3

Mintzer, M. Z., & Snodgrass, J. G. (1999). The picture superiority effect: Support for the distinctiveness model. *American Journal of Psychology*, *112*(1), 113–146. https://doi.org/10.2307/1423627

Modrak, D. K. W. (2001). *Aristotle's theory of language and meaning*. Cambridge University Press.

Moore, T. V. (1919). Image and meaning in memory and perception. *Psychological Monographs*, *27*(2), i–296. https://doi.org/10.1037/H0093193

Morey, R. D., Rouder, J. N., Jamil, T., Urbanek, S., Forner, K., & Ly, A. (2011). *BayesFactor: Computation of Bayes factors for common design*.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*(5), 519–533. https://doi.org/10.1016/S0022-5371(77)80016-9

Mulhall, E. F. (1915). Experimental studies in recall and recognition. *The American Journal of Psychology*, *26*(2), 217. https://doi.org/10.2307/1413251

Murphy, D. H., & Castel, A. D. (2021). Tall towers: Schemas and illusions when perceiving and remembering a familiar building. *Applied Cognitive Psychology*, *35*(5), 1236–1246. https://doi.org/10.1002/acp.3855

Neath, I., & Brown, G. D. A. (2006). SIMPLE: Further applications of a local distinctiveness model of memory. *Psychology of Learning and Motivation - Advances in Research and Theory*, *46*, 201–243. https://doi.org/10.1016/S0079-7421(06)46006-0

Needell, C. D., & Bainbridge, W. A. (2022). Embracing new techniques in deep learning for estimating image memorability. *Computational Brain & Behavior*, *5*(2), 168–184. https://doi.org/10.1007/s42113-022-00126-5

Nelson, D. L. (1979). Remembering pictures and words: Appearance, significance, and name. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 45–75). Erlbaum. https://www.taylorfrancis.com/chapters/mono/10.4324/9781315796192-11/remembering-pictures-words-appearance-significance-name-laird-cermak-fergus-craik

Nelson, D. L., Bajo, M. T., & Casanueva, D. (1985a). Prior knowledge and memory: The influence of natural category size as a function of intention and distraction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(1), 94–105. https://doi.org/10.1037/0278-7393.11.1.94

Nelson, D. L., Cañas, J., Casanueva, D. M., & Castaño, D. (1985b). Prior knowledge and recognition. *The American Journal of Psychology*, *98*(3), 379–397. https://doi.org/10.2307/1422624

Nelson, D. L., & McEvoy, C. L. (1979). Encoding context and set size. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(3), 292–314. https://doi.org/10.1037/0278-7393.5.3.292

Nelson, D. L., Reed, V. S., & McEvoy, C. L. (1977). Learning to order pictures and words: A model of sensory and semantic encoding. *Journal of Experimental Psychology: Human Learning & Memory*, *3*(5), 485–497. https://doi.org/10.1037/0278-7393.3.5.485

Nickerson, R. S., & Adams, M. J. (1979). Long-term memory for a common object. *Cognitive Psychology*, *11*(3), 287–307. https://doi.org/10.1016/0010-0285(79)90013-6

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review*, *76*(3), 241–263. https://doi.org/10.1037/h0027272

Paivio, A. (1971). *Imagery and verbal processes*. Holt, Rinehart, and Winston.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Experimental Psychology*, *45*(3), 255–287. https://doi.org/10.1037/H0084295

Paivio, A., & Csapo, K. (1969). Concrete image and verbal memory codes. *Journal of Experimental Psychology*, *80*, 279–285. https://doi.org/10.1037/h0027273

Paivio, A., & Csapo, K. (1973). Picture superiority in free recall: Imagery or dual coding? *Cognitive Psychology*, *5*(2), 176–206. https://doi.org/10.1016/0010-0285(73)90032-7

Paivio, A., Walsh, M., & Bons, T. (1994). Concreteness effects on memory: When and why? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(5), 1196–1204. https://doi.org/10.1037/0278-7393.20.5.1196

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Poirier, M., & Saint-Aubin, J. (1995). Memory for related and unrelated words: Further evidence on the influence of semantic factors in immediate serial recall. *The Quarterly Journal of Experimental Psychology Section A*, *48*(2), 384–404. https://doi.org/10.1080/14640749508401396

Prada, M., Rodrigues, D., Silva, R. R., & Garrido, M. V. (2016). Lisbon Symbol Database (LSD): Subjective norms for 600 symbols. *Behavior Research Methods*, *48*(4), 1370–1382. https://doi.org/10.3758/S13428-015-0643-7

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.r-project.org/

Raven, J. C. (1958). *Guide to using the Mill Hill Vocabulary Scale with the Progressive Matrices Scales* (p. 64). H. K. Lewis & Co.

Reinke, K., Fernandes, M. A., Schwindt, G., O'Craven, K., & Grady, C. L. (2008). Functional specificity of the visual word form area: General activation for words and symbols but specific network activation for words. *Brain and Language*, *104*(2), 180–189. https://doi.org/10.1016/j.bandl.2007.04.006

Rihn, A., Wei, X., & Khachatryan, H. (2019). Text vs. Logo: Does eco-label format influence consumers' visual attention and willingness-to-pay for fruit plants? An experimental auction approach. *Journal of Behavioral and Experimental Economics*, *82*, 101452. https://doi.org/10.1016/j.socec.2019.101452

Roberts, B. R. T., Fernandes, M. A., & MacLeod, C. M. (in press). Memory for symbolic images: Findings from sports team logos. *Journal of Applied Research in Memory and Cognition*.

Roberts, B. R. T., MacLeod, C. M., & Fernandes, M. A. (in press). Symbol superiority: Why $ is better remembered than 'dollar'. *Cognition*. https://doi.org/10.1016/j.cognition.2023.105435

Roberts, B. R. T., & Wammes, J. D. (2021). Drawing and memory: Using visual production to alleviate concreteness effects. *Psychonomic Bulletin & Review*, *28*(1), 259–267. https://doi.org/10.3758/s13423-020-01804-w

Roediger, H. L., & Blaxton, T. A. (1987). Effects of varying modality, surface features, and retention interval on priming in word-fragment completion. *Memory & Cognition*, *15*(5), 379–388. https://doi.org/10.3758/BF03197728

Roediger, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, *1*(3), 181–210.

Roodenrys, S., & Quinlan, P. T. (2000). The effects of stimulus set size and word frequency on verbal serial recall. *Memory*, *8*(2), 71–78. https://doi.org/10.1080/096582100387623

Ruxton, G. D. (2006). The unequal variance t-test is an underused alternative to Student's t-test and the Mann-Whitney U test. *Behavioral Ecology*, *17*(4), 688–690. https://doi.org/10.1093/beheco/ark016

Sakurai, Y., Mimura, I., & Mannen, T. (2008). Agraphia for kanji resulting from a left posterior middle temporal gyrus lesion. *Behavioural Neurology*, *19*(3), 93–106.

Saussure, F. de, & Harris, R. (1998). *Course in General Linguistics* (Reprint edition). Open Court.

Shakespeare, W. (2011). *Romeo and Juliet* (B.A. Mowat & P. Werstine, Eds). Simon & Schuster. https://www.simonandschuster.com/books/Romeo-and-Juliet/William-Shakespeare/Folger-Shakespeare-Library/9781451621709 (Original work published 1597)

Shepard, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior*, *6*(1), 156–163. https://doi.org/10.1016/S0022-5371(67)80067-7

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning & Memory*, *4*(6), 592–604. https://doi.org/10.1037/0278-7393.4.6.592

Smith, A. D., Park, D. C., Cherry, K., & Berkovsky, K. (1990). Age differences in memory for concrete and abstract pictures. *Journal of Gerontology*, *45*(5), P205–P210. https://doi.org/10.1093/geronj/45.5.P205

Sokolowski, H. M., Fias, W., Bosah Ononye, C., & Ansari, D. (2017). Are numbers grounded in a general magnitude processing system? A functional neuroimaging meta-analysis. *Neuropsychologia*, *105*, 50–69. https://doi.org/10.1016/j.neuropsychologia.2017.01.019

Stenberg, G., Radeborg, K., & Hedman, L. R. (1995). The picture superiority effect in a cross-modality recognition task. *Memory & Cognition*, *23*(4), 425–441. https://doi.org/10.3758/BF03197244

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., Lu, C. C., Pechmann, T., Pléh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., Orozco-Figueroa, A., … Bates, E. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language*, *51*(2), 247–250.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of Memory*. Academic Press.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 352–373. https://doi.org/10.1037/h0020071

van Riel, C. B. M., & van den Ban, A. (2001). The added value of corporate logos - An empirical study. *European Journal of Marketing*, *35*(3–4), 428–440. https://doi.org/10.1108/03090560110382093

Vinson, D., Jones, M., Sidhu, D. M., Lau-Zhu, A., Santiago, J., & Vigliocco, G. (2021). Iconicity emerges and is maintained in spoken language. *Journal of Experimental Psychology: General*, *150*(11), 2293–2308. https://doi.org/10.1037/xge0001024

Vogt, S., & Magnussen, S. (2005). Hemispheric specialization and recognition memory for abstract and realistic pictures: A comparison of painters and laymen. *Brain and Cognition*, *58*(3), 324–333. https://doi.org/10.1016/j.bandc.2005.03.003

Wammes, J. D., Meade, M. E., & Fernandes, M. A. (2016). The drawing effect: Evidence for reliable and robust memory benefits in free recall. *Quarterly Journal of Experimental Psychology*, *69*(9), 1752–1776. https://doi.org/10.1080/17470218.2015.1094494

Weldon, M. S., & Roediger, H. L. (1987). Altering retrieval demands reverses the picture superiority effect. *Memory & Cognition*, *15*(4), 269–280. https://doi.org/10.3758/BF03197030

Wolfe, J. M. (1999). Inattentional amnesia [Data set]. In *Fleeting memories* (pp. 71–94). MIT Press. https://doi.org/10.1037/e536982012-171

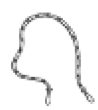
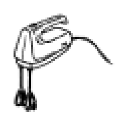
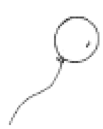
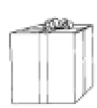
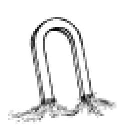
Wong, K., Wadee, F., Ellenblum, G., & McCloskey, M. (2018). The devil's in the g-tails: Deficient letter-shape knowledge and awareness despite massive visual experience. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(9), 1324–1335. https://doi.org/10.1037/xhp0000532

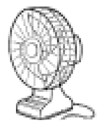
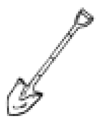
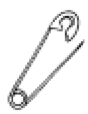
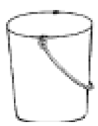
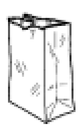
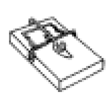
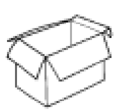
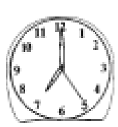
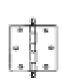
Xie, W., Bainbridge, W. A., Inati, S. K., Baker, C. I., & Zaghloul, K. A. (2020). Memorability of words in arbitrary verbal associations modulates memory retrieval in the anterior temporal lobe. *Nature Human Behaviour*, *4*(9), 937–948. https://doi.org/10.1038/s41562-020-0901-2

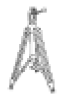
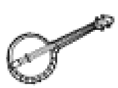
Appendix A – Fruit/Vegetable Word Stimuli Presented in Experiment 3

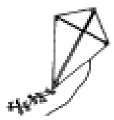
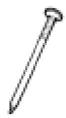
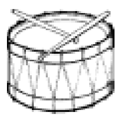
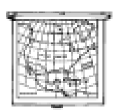
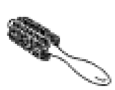
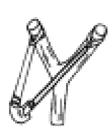
apple	cayenne pepper	kidney beans	rutabaga
artichoke	celery	leek	scallion
arugula	chard	lemongrass	seaweed
asparagus	chickpeas	lentils	shallot
avocado	chili pepper	lettuce	snap peas
banana pepper	chives	lima bean	soybean
banana squash	collard greens	mushroom	spinach
basil	cucumber	mustard greens	split peas
bean sprouts	dill	okra	squash
beet	eggplant	onion	sweet potato
bell pepper	fennel	parsley	sweetcorn
black beans	garlic	parsnip	tabasco pepper
bok choy	ginger	peas	thyme
broccoli	green beans	pickle	tomato
brussels sprouts	green onion	potato	turnip
butternut squash	habanero	pumpkin	wasabi
cabbage	horseradish	radish	water chestnut
capers	iceberg lettuce	red cabbage	watercress
carrot	jalapeño	rhubarb	white radish
cauliflower	kale	romaine	zucchini

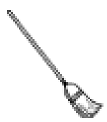
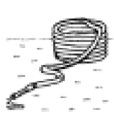
Appendix B – Picture and Word Stimuli Presented in Experiment 4

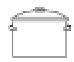
Stimuli List	Word	Image	# Alternative Labels	% Label Agreement	Frequency (CELEX)
A	bottle		3	0.90	4.76
A	gun		2	0.71	4.61
A	dress		1	1.00	4.48
A	jumprope		3	0.84	
A	mixer		6	0.39	1.10
A	balloon		1	1.00	1.95
A	present		5	0.67	2.89
A	magnet		2	0.96	1.39
A	globe		2	0.98	2.49

Stimuli List	Word	Image	# Alternative Labels	% Label Agreement	Frequency (CELEX)
A	puzzle		2	0.98	2.30
A	fan		2	0.98	2.89
A	shovel		1	1.00	1.61
A	safetypin		7	0.53	0.69
A	bucket		2	0.66	3.05
A	bag		4	0.84	4.39
A	mousetrap		3	0.65	0.69
A	box		1	1.00	4.64
A	clock		2	0.98	3.69
A	hinge		5	0.88	1.61

Stimuli List	Word	Image	# Alternative Labels	% Label Agreement	Frequency (CELEX)
A	tripod		6	0.79	0.69
A	teapot		8	0.44	1.61
A	fishingpole		6	0.53	
A	needle		5	0.91	2.83
A	violin		4	0.82	1.95
A	paperclip		4	0.81	
B	banjo		5	0.87	
B	comb		1	1.00	1.79
B	razor		3	0.94	2.30
B	telescope		2	0.98	2.20

Stimuli List	Word	Image	# Alternative Labels	% Label Agreement	Frequency (CELEX)
B	stethoscope		4	0.93	0.69
B	kite		1	1.00	1.79
B	nail		1	1.00	3.26
B	drum		2	0.80	2.83
B	tweezers		5	0.91	1.10
B	arrow		2	0.98	2.77
B	map		1	1.00	3.71
B	piggybank		4	0.94	
B	brush		3	0.84	2.89
B	slingshot		5	0.82	0.69

Stimuli List	Word	Image	# Alternative Labels	% Label Agreement	Frequency (CELEX)
B	mop		2	0.94	1.39
B	microscope		2	0.84	2.20
B	picture		4	0.83	5.17
B	glass		2	0.71	4.98
B	broom		1	1.00	2.20
B	microphone		5	0.90	2.20
B	bomb		5	0.90	3.74
B	vase		3	0.94	2.08
B	chain		1	1.00	3.89
B	hose		2	0.96	1.61

Stimuli List	**Word**	**Image**	**# Alternative Labels**	**% Label Agreement**	**Frequency (CELEX)**
B	pot		8	0.73	3.61

Appendix C – Sports Team Stimuli Presented in Experiments 5A and 5B

The content of Appendix C has been removed due to copyright restrictions. It was a list of sports logos stimuli used in Experiment 5A and 5B. Original sources of these logos are available from numerous online websites, but I obtained them primarily from https://www.sportslogos.net/.

Appendix D – Analyses of Hits, False Alarms, and Accuracy in Experiments 5A and 5B

Hit Rate and False Alarm Rate

In Experiment 5A, there was only a marginal main effect of Group when hit rate was the dependent measure, $F_{Welch}(2, 147.06) = 2.70$, $p = .070$, $\eta_p^2 = .04$, CI_{95} [0.00, 0.10], $BF_{01} = 1.68$, whereas false alarm rate differed markedly across groups, $F_{Welch}(2, 151.20) = 13.43$, $p < .001$, $\eta_p^2 = .15$, CI_{95} [0.06, 0.25], $BF_{10} = 17{,}984.08$. Games-Howell pairwise comparisons on false alarm rate showed that incorrect endorsements were lower in the Symbols group and in the Combined group than in the Words group ($p < .001$, $BF_{10} = 1{,}0525.97$, $d = -0.81$, CI_{95} [-1.13, -0.48], and $p = .001$, $BF_{10} = 76.74$, $d = -0.60$, CI_{95} [-0.92, -0.28], respectively), but that the Symbols group and the Combined group did not differ ($p = .345$, $BF_{01} = 2.30$, $d = 0.23$, CI_{95} [-0.09, 0.56]).

In Experiment 5B, the main effect of Group was significant for hit rate, $F_{Welch}(2, 142.10) = 7.36$, $p = .001$, $\eta_p^2 = .09$, CI_{95} [0.03, 1.00], $BF_{10} = 33.20$. Pairwise comparisons showed that hit rate was higher in the Words group than in the Combined group and the Symbols group ($p = .011$, $BF_{10} = 8.18$, $d = -0.48$, CI_{95} [-0.81, -0.15] and $p = .001$, $BF_{10} = 58.96$, $d = -0.59$, CI_{95} [-0.92, -0.26], respectively), whereas the Combined group and the Symbols group did not differ ($p = .819$, $BF_{01} = 4.67$, $d = 0.10$, CI_{95} [-0.23, 0.43]). The main effect of Group in Experiment 5B was also significant with false alarm rate as the dependent measure, $F_{Welch}(2, 141.11) = 21.44$, $p < .001$, $\eta_p^2 = .23$, CI_{95} [0.13, 1.00], $BF_{10} = 352{,}190$. Pairwise comparisons revealed that false alarms were more prevalent in the Symbols group than in the Combined group and the Words group ($p < .001$, $BF_{10} = 45{,}034.07$, $d = -0.91$,

CI_{95} [-1.26, -0.56] and $p < .001$, $BF_{10} = 91{,}146.08$, $d = 0.92$, CI_{95} [0.58, 1.25], respectively), while the Combined group and the Words group did not differ ($p = .993$, $BF_{01} = 5.58$, $d = 0.02$, CI_{95} [-0.31, 0.34]).

Accuracy

For each experiment, I conducted one-way[23] Welch-adjusted between-subjects ANOVAs with Group (Words, Symbols, Combined) as the factor (Figure A1). The dependent measure was memory accuracy (hit rate minus false alarm rate).

In Experiment 5A, there was a significant effect of Group, $F_{\text{Welch}}(2, 149.30) = 11.68$, $p < .001$, $\eta_p^2 = .14$, CI_{95} [0.04, 0.24], $BF_{10} = 1{,}704.82$. Games-Howell pairwise comparisons revealed that accuracy was higher in the Symbols group and in the Combined group than in the Words group ($p < .001$, $BF_{10} = 5{,}004.76$, $d = 0.77$, CI_{95} [0.44, 1.09], and $p = .029$, $BF_{10} = 3.61$, $d = 0.42$, CI_{95} [0.10, 0.73], respectively), but that the Symbols group and the Combined groups did not differ ($p = .065$, $BF_{01} = 0.56$, $d = -0.37$, CI_{95} [-0.70, -0.05])[24].

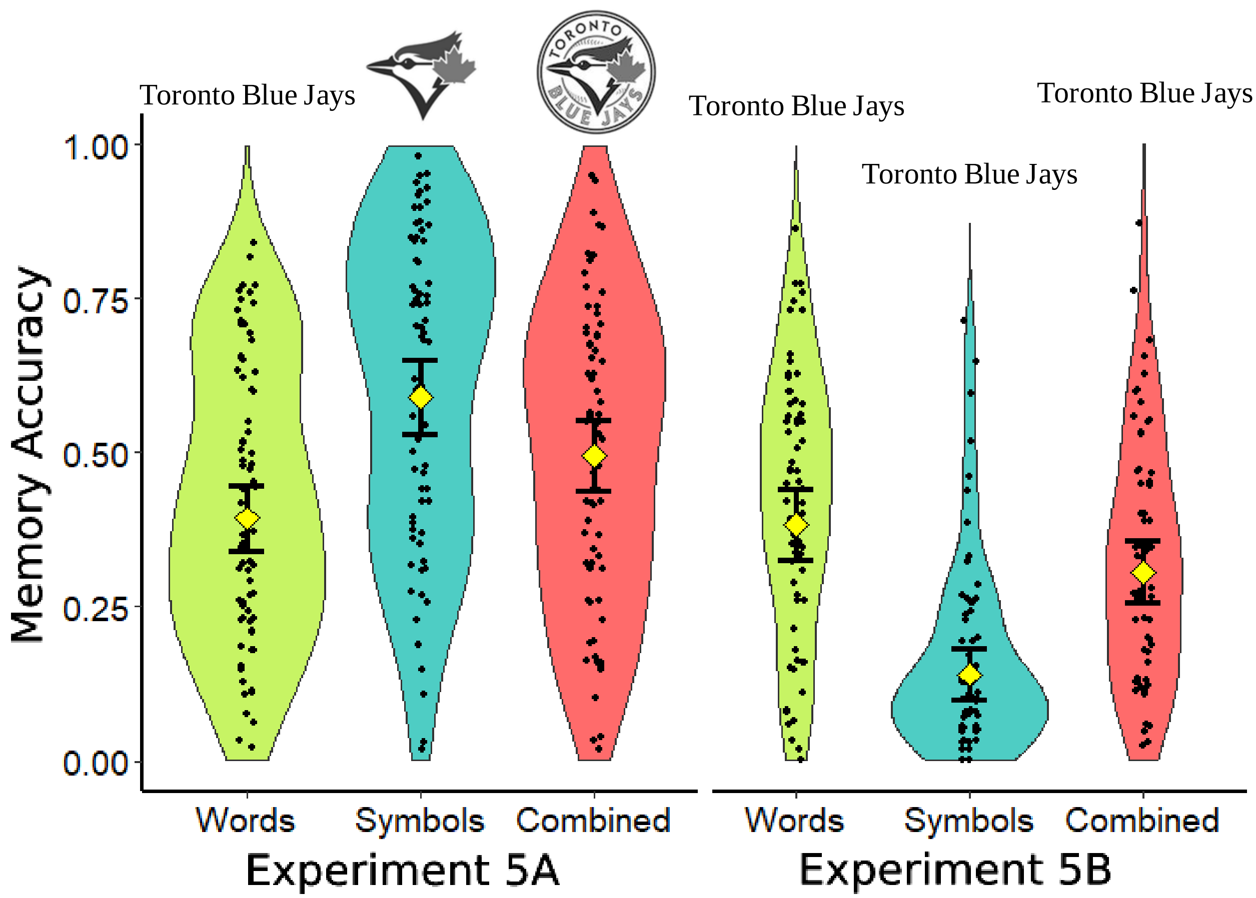
In Experiment 5B, the effect of Group was also significant, $F_{\text{Welch}}(2, 140.46) = 28.30$, $p < .001$, $\eta_p^2 = .29$, CI_{95} [0.18, 1.00], $BF_{10} = 58{,}481{,}398$. Pairwise comparisons revealed a different pattern of results, however, such that accuracy was higher in the Combined group and in the Words group than in the Symbols group ($p < .001$, $BF_{10} = 20{,}626.31$, $d = 0.88$, CI_{95} [0.53, 1.23], and $p < .001$, $BF_{10} = 48{,}670{,}185$, $d = -1.14$, CI_{95} [-1.49, 0.79],

[23] I initially conducted a 2 (Experiment: 5A, 5B) x 3 (Group: Words, Logos, Combined) between-subjects ANOVA for memory accuracy, which confirmed both significant main effects and the interaction, thus prompting separate one-way ANOVAs for each experiment.
[24] When pairwise comparison tests were performed on *d'*, this difference was statistically significant ($p = .044$).

respectively), but that the Combined group and the Words group did not differ (p = .109, BF_{01} = 0.89, d = -0.33, CI_{95} [-0.66, -0.01]).

Figure A1

Memory Accuracy Across Groups in Each Experiment



Note. Error Bars = 95% confidence intervals. Examples of recognition test stimuli for each group are provided above their data points.

Appendix E – Symbols and Their Word Counterparts Presented in Experiments 1-4

Symbol	Word	Symbol	Word	Symbol	Word	Symbol	Word
♀	female	∇	yield	🔊	sound	♦	diamond
¢	cent	♂	male	÷	divide	⚖	justice
&	and	=	equals	$	dollar	⚕	medicine
🔍	search	%	percent	∞	infinity	☀	brightness
!	exclamation	⏸	pause	❄	cold	☽	night
♥	love	♣	clubs	β	beta	☼	day
✡	judaism	©	copyright	☢	radioactive	✓	correct
✆	telephone	↺	undo	✂	cut	™	trademark
🔀	shuffle	🎤	recording	⚛	atomic	Ψ	psi
🔇	mute	@	at	☠	poison	Ω	omega
⯃	stop	∑	sum	⚠	danger	🖂	email
♲	recycle	~	approximate	☪	islam	≠	unequal
♠	spade	®	registered	⌛	time	⏩	fast-forward
€	euro	¥	yen	℃	celsius	⏪	rewind
🛈	information	£	pound	×	multiply	🔁	repeat
☮	peace	℉	fahrenheit	+	add	Δ	delta
†	christianity	☣	biohazard	📶	reception	±	plus-minus
?	question	⊘	do not	π	pi	√	square root
🖶	print	☭	communism	🔔	notification	#	hashtag
▶	play	☯	yin yang	✈	airport	α	alpha