Journal of Applied Research in Memory and Cognition

Memory for Symbolic Images: Findings From Sports Team Logos

Brady R. T. Roberts, Myra A. Fernandes, and Colin M. MacLeod

CITATION

Roberts, B. R. T., Fernandes, M. A., & MacLeod, C. M. (2023, January 19). Memory for Symbolic Images: Findings From Sports Team Logos. *Journal of Applied Research in Memory and Cognition*. Advance online publication. https://dx.doi.org/10.1037/mac0000079

AMERICAN PSYCHOLOGICAL ASSOCIATION SARMAC Journal of Applied Research in Memory and Cognition

ISSN: 2211-3681

https://doi.org/10.1037/mac0000079

EMPIRICAL ARTICLE

Memory for Symbolic Images: Findings From Sports Team Logos

Brady R. T. Roberts, Myra A. Fernandes, and Colin M. MacLeod
Department of Psychology, University of Waterloo, Canada

OPEN DATA OPEN MATERIALS PREREGISTERED+

Pictures typically are better remembered than words—the *picture superiority effect* (PSE). An obvious yet understudied application of picture superiority is to advertising. We compared memorability of names of professional sports teams presented in three encoding conditions: team names only, team logos without names, and team logos with integrated names. Results of Experiment 1A provided the first evidence of an intact PSE for graphic symbols representing abstract concepts. This effect was, however, influenced by familiarity with the to-be-remembered stimuli. Experiment 1B highlighted the role of expertise in memory for logos: When tested on team names, the magnitude of the benefit for the logos-only group depended on whether participants knew what the logos represented. These experiments emphasize familiarity as an undervalued factor influencing memory for pictures. We suggest that logos, when featured in advertisements, should be accompanied by text labels to maximize memorability, especially for those unfamiliar with the brand.

General Audience Summary

It has long been known that pictures are better remembered than words. Would graphic symbols representing abstract concepts also be better remembered than their word counterparts? To address this question, we studied memory for North American sports team logos. We found that sports team logos—with or without team name labels—were better remembered than their corresponding verbal labels. The benefit conferred by the graphic format was evident only in people who had preexisting familiarity with the team names. Overall, our work reveals that picture superiority in memory can be extended to graphic symbols, and that familiarity plays a role in determining their later memory. We suggest that advertisers make use of integrated labels in their logos to ensure that unfamiliar consumers are still able to associate logos with company names, ultimately improving brand memory.

Keywords: visual memory, sports logos, symbols, advertising

Supplemental materials: https://doi.org/10.1037/mac0000079.supp

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 colleagues had repeatedly

---

Brady R. T. Roberts https://orcid.org/0000-0002-2228-8805
Myra A. Fernandes https://orcid.org/0000-0002-1467-0342
Colin M. MacLeod https://orcid.org/0000-0002-8350-7362

The authors thank Batul Karimjee for their assistance in preparing materials for this study. This research was supported by a Natural Sciences and Engineering Research Council (NSERC) of Canada postgraduate scholarship to Brady R. T. Roberts and by NSERC Discovery Grants to Colin M. MacLeod (Grant A7459) and Myra A. Fernandes (Grant 2020-03917).

Brady R. T. Roberts had the idea to test memory for symbolic images and played the lead role in conceptualization, data curation, formal analysis, methodology, project administration, software, visualization, and writing of original draft and revisions. Myra A. Fernandes suggested using sport team logos to equate set size and to explore applicability of the research. Colin M. MacLeod suggested the verbal recognition test in Experiment 1B. Myra A. Fernandes and Colin M. MacLeod offered guidance throughout the project and participated in the preparation of the article.

The data are available at https://osf.io/xqsdz.

The experimental materials are available at https://osf.io/xqsdz.

The preregistered design and analysis plan are accessible at https://osf.io/uab64 and https://osf.io/hq3cm.

Correspondence concerning this article should be addressed to Brady R. T. Roberts, Department of Psychology, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L 3G1, Canada. Email: bradyrtroberts@gmail.com

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 for decades (e.g., Ensor et al., 2019; McDaniel & Pressley, 1987), yet researchers still often disagree on the underlying mechanism (e.g., Amrhein et al., 2002). Popular alternative accounts suggest (a) that pictures are *conceptually* more distinctive than words because they elicit greater 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., 2019; 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 encoding 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, as 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, Bajo, & Casanueva, 1985; Nelson, Cañas, et al., 1985; 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 any potential contribution of integrating text into a product image, beyond memory for a product image alone. According to the 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). Our research sought to bridge these findings to investigate whether logos are better remembered than their brand name counterparts, and whether adding integrated labels would further enhance memorability.

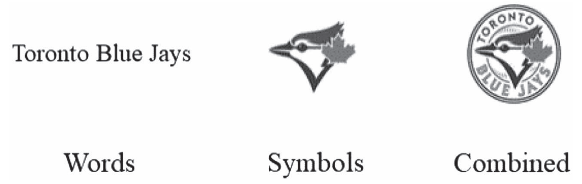
Our broad goal was to determine whether memory for logos—specifically, the logos of North American professional sports teams—both with and without integrated verbal labels, would be superior to memory for the labels alone. Would the PSE generalize to graphic symbols for abstract concepts, like those representing brands? Are logos, such as for the Toronto Blue Jays professional baseball team (see Figure 1), processed similarly to images of concrete objects that are usually the focus of study in the PSE literature? Our 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 abstract concepts (here, sports teams) operate using the same mechanisms proposed to underlie picture superiority for images of concrete items.

We tested two principal hypotheses 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). Our 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), we also predicted no additional benefit from provision of team names embedded within the logos (i.e., symbols = combined).

Would our first hypothesis hold true for everyone or only for those very familiar with North American sports? We reasoned that people with high sports familiarity may already have in memory a verbal label corresponding to most sports logos. 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, we administered a "sports familiarity questionnaire" and compared memory benefits at different levels of this measure. By examining whether familiarity with the stimuli affects encoding,

Figure 1

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



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

we are among the first to determine whether relative expertise influences how pictures are studied and later remembered.

Finally, we examined whether the PSE would hold even for images from a finite set matched to a set of words. Recall that some (Nelson, Bajo, & Casanueva, 1985; Nelson, Cañas, et al., 1985) have suggested that the PSE arises only because pictures represent a smaller set than words. Related work in our laboratory exploring memory for "everyday symbols" (e.g., !@#$%) demonstrates that symbols (e.g., $) are better remembered than their word counterparts (e.g., "dollar"; Roberts et al., 2022). A possible criticism of our ongoing 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. We 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 National Football League (football), National Hockey League (hockey), National Basketball Association (basketball), and Major League Baseball (baseball)—currently, there are exactly 124 teams. Our study is therefore interesting both for advertising purposes and for extension of the picture superiority literature to symbolic images more broadly. 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.

In the present study, we 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, we ran the study twice. In Experiment 1A, the recognition test stimuli matched the format seen during encoding. In Experiment 1B, 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, we could examine differences in retrieval processes and better match other PSE research. In both experiments, our major predictions were that (a) sports logos would be better remembered than team names (even when equating for set size and semantic content), (b) relative to the symbols group, the addition of team name labels in the "combined" format would confer no further memory benefit, and (c) familiarity could moderate the observed outcomes.

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 − 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, we aimed to collect a minimum of 53 participants per group ($N = 159$ total), with our ideal target sample size set higher at 69 per group ($N = 207$ total) in each experiment.

Our initial samples consisted of data from 251 participants in Experiment 1A and 250 in Experiment 1B, all recruited via the online crowdsourcing platform, Prolific (www.prolific.co). Built-in prescreening options on Prolific were used to permit participation only of those who had declared themselves to be residents of Canada or the United States. 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 (a) had corrupted or incomplete data files (Experiment 1A: $n = 8$; Experiment 1B: $n = 19$), (b) took more than 30 min to complete the study (Experiment 1A: $n = 1$; Experiment 1B: $n = 6$), (c) took less than 5 min to complete the study (Experiment 1A: $n = 1$; Experiment 1B: $n = 0$), or (d) were greater than ±3 *SD* away from the mean of remaining participants for study duration (Experiment 1A: $n = 8$; Experiment 1B: $n = 5$). We therefore entered statistical analyses with 233 valid data files in Experiment 1A and 220 in Experiment 1B. From these samples, participants were filtered out if they were statistical outliers (±3 *SD*) on any one of our metrics of memory performance (hits, false alarms, accuracy, or d-prime) as calculated within each experiment (Experiment 1A: $n = 3$; Experiment 1B: $n = 3$).

For Experiment 1A, the final sample used in formal statistical analyses consisted of 230 participants (49.13% female), ranging in age from 18 to 64 ($M = 32.38$, $SD = 10.96$), split across three groups. For Experiment 1B, the final sample consisted of 217 participants (75.45% female, four preferred not to declare their sex), ranging in age from 18 to 64 ($M = 26.17$, $SD = 7.69$), also split across three groups. We therefore exceeded our target sample sizes, ensuring adequate statistical power in both experiments.[1] Participation in each experiment took approximately 12 min, and participants were paid the equivalent of $2.53 Canadian Dollars for their time.

Materials

All materials were derived from professional North American sports teams playing in one of the four major leagues: National Football League (football, $n = 32$), National Hockey League (hockey, $n = 32$), National Basketball Association (basketball, $n = 30$), and Major League Baseball (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.

For the words stimuli, full team names were presented at the center of the screen in Times New Roman size 48 black font 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, we sought a version with the text removed or we 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, all logo-type stimuli (symbols and combined) were then

---

[1] We preregistered a target sample size of $n = 69$ per group, per experiment, but ended up with more than this after data cleaning. To ensure the highest statistical power possible, we did not trim data in the analyses presented here. However, when the data were randomly trimmed to our original target sample size in each experiment—consistent with our preregistered method—the patterns of results were identical to those presented here.

resized to be 110 × 116 pixels before being converted to gray scale with a white background (see Figure 1).

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 2 min. 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 1A, test stimuli matched those studied at encoding (except that, in the words group, font size was reduced to 32 pt.). For Experiment 1B, all test stimuli were presented as team names in plain text, matching the retrieval test format of the words group in Experiment 1A. We used only text-based stimuli on the recognition test in Experiment 1B with the intention of limiting processing difference 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 1B, 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, if their self-rated knowledge of at least one major sports league was higher than average, and if they were familiar (had preexisting 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 No. 41594). All materials, experiment files, data, statistical analysis code, and our preregistration for these experiments can be found on the Open Science Framework (https://osf.io/xqsdz/).

Results

Overall Memory Performance

For each experiment, we conducted one-way[2] Welch-adjusted[3] between-subjects analyses of variance (ANOVAs) with group (words, symbols, combined) as the only independent variable (see Figure 2). The dependent measure was memory sensitivity (d'; d-prime).[4]

In Experiment 1A, there was a significant effect of group,[5] $F_{\text{Welch}}(2, 145.77) = 12.71$, $p < .001$, $\eta_p^2 = .15$, 95% CI [0.07, 1.00], $\text{BF}_{10} = 4{,}091$.[6] Games-Howell pairwise comparisons revealed that memory sensitivity was higher in the symbols group than in the combined group ($p = .040$, $\text{BF}_{10} = 2.49$, $d = -0.40$, 95% CI [−0.73, −0.07]),[7] both of which showed higher memory sensitivity than the words group ($p < .001$, $\text{BF}_{10} = 12{,}773$, $d = 0.80$, 95% CI [0.47, 1.13] and $p = .030$, $\text{BF}_{10} = 4.03$, $d = 0.42$, 95% CI [0.10, 0.74], respectively).

In Experiment 1B, the effect of group was also significant,[8] $F_{\text{Welch}}(2, 138.16) = 31.17$, $p < .001$, $\eta_p^2 = .31$, 95% CI [0.21,

---

[2] We initially conducted a 2 (experiment: 1A, 1B) × 3 (group: words, symbols, 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.

[3] Welch’s tests were used for all ANOVAs, and Games-Howell tests were used for all pairwise comparisons due to a violation of homogeneity of variance (via Levene’s test) for our measure of memory sensitivity (d'; d-prime) in both Experiment 1A, $F(2, 227) = 4.46$, $p = .013$, and Experiment 1B, $F(2, 214) = 9.45$, $p < .001$, 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, we therefore used these tests throughout.

[4] 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).

[5] One-way ANOVAs, based separately on hit rate, false alarm rate, and accuracy, are presented in the Supplemental Materials.

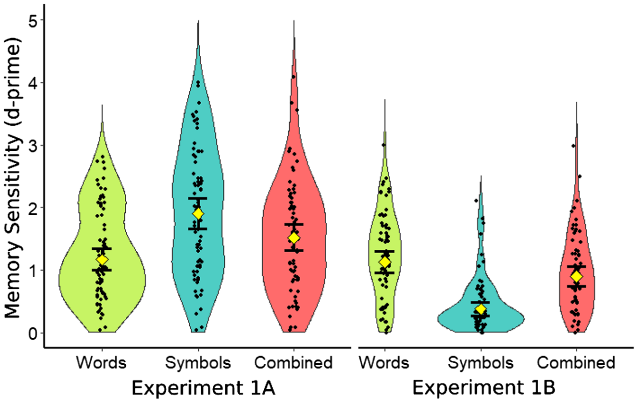
[6] Throughout this article, Bayes factors were calculated using the *BayesFactor* (Morey et al., 2011) package for *R*, enlisting a default Jeffreys–Zellner–Siow (JZS) prior with a Cauchy distribution (center = 0, $r = 0.707$). This package compares the fit of various linear models. In the present case, Bayes factors for the alternative (BF_{10}) are in comparison to null models with subject-level random error. 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 null hypothesis significance testing analyses (i.e., based on a $p < .05$ criterion).

[7] When pairwise comparison tests were performed on accuracy, this difference was no longer statistically significant ($p = .065$).

[8] Because the percentage of female participants in Experiment 1A was 49%, but in Experiment 1B, it was 75%, we reconducted 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 2

Memory Sensitivity Across Groups in Each Experiment



Note. Error bars = 95% confidence intervals. See the online article for the color version of this figure.

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 ($p < .001$, $BF_{10} = 53{,}830$, $d = 0.92$, 95% CI [0.57, 1.27] and $p < .001$, $BF_{10} = 127{,}311{,}417$, $d = -1.17$, 95% CI [−1.52, −0.82], respectively), but that the combined group and the words group did not differ ($p = .130$, $BF_{01} = 1.04$, $d = -0.32$, 95% CI [−0.65, 0.01]).

Memory as a Function of Sports Familiarity

Comparing Upper and Lower Quartiles

To assess whether memory performance was influenced by familiarity with sports, we performed 2 one-way between-subjects ANOVAs with group (words, symbols, combined) as the independent variable, again using memory sensitivity as the dependent measure. In Experiment 1A, for those in the top quartile of sports familiarity, the three groups did not differ significantly (see Supplemental Table 1),[9] $F_{\text{Welch}}(2, 30.07) = 0.82$, $p = .450$, $\eta_p^2 = .05$, 95% CI [0.00, 1.00], $BF_{01} = 4.02$. For those in the bottom quartile, however, there was a significant effect of group, $F_{\text{Welch}}(2, 39.24) = 8.79$, $p < .001$, $\eta_p^2 = .31$, 95% CI [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 ($p = .002$, $BF_{10} = 224.49$, $d = -1.28$, 95% CI [−1.95, −0.60] and $p = .009$, $BF_{10} = 44.29$, $d = 1.03$, 95% CI [0.39, 1.67], respectively), and that the latter two did not differ ($p = .490$, $BF_{01} = 1.72$, $d = -0.35$, 95% CI [−0.95, 0.25]).

Corresponding analyses for Experiment 1B revealed a similar overall pattern. The main effect of group was not statistically significant for participants in the top quartile of sports familiarity, $F_{\text{Welch}}(2, 27.54) = 1.76$, $p = .190$, $\eta_p^2 = .11$, 95% CI [0.00, 1.00], $BF_{01} = 2.06$, whereas it was significant for participants in the bottom quartile, $F_{\text{Welch}}(2, 33.40) = 21.69$, $p < .001$, $\eta_p^2 = .57$, 95% CI [0.36, 1.00], $BF_{10} = 2{,}009$. Relative to the finding for Experiment 1A, 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 ($p = .003$, $BF_{10} = 50.69$, $d = 1.22$, 95% CI [0.52, 1.90] and $p < .001$, $BF_{10} = 9{,}847$, $d = -1.70$, 95% CI [−2.43, −0.95], respectively), whereas the combined group and the words group did not differ ($p = .250$, $BF_{01} = 1.18$, $d = -0.49$, 95% CI [−1.09, 0.11]).

Correlation Between Sports Familiarity and Memory Performance

Finally, we examined Pearson correlations between total sports familiarity scores and all of our metrics of memory performance (hit rate, false alarm rate, and memory sensitivity). For Experiment 1A, 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, we found that sports familiarity correlated significantly with hit rate in the combined group, $r(70) = .31$, $p = .009$. All other correlations were nonsignificant ($ps \geq .090$).

In Experiment 1B, 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, nonsignificant ($p = .275$). 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 hit rate in the same two groups, $r(67) = .27$, $p = .023$ and $r(69) = .34$, $p = .004$, respectively. All other correlations were nonsignificant ($ps \geq .162$).

General Discussion

We 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. We aimed first to determine whether the PSE generalized to graphic symbols representing abstract concepts. Second, we 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). Our novel approach also allowed examination of the effects of familiarity on picture encoding. Given that the symbols we used possess high cultural significance, some participants were very familiar with them and some were not. Thus, we could establish whether expertise plays a role when encoding and later retrieving pictures from memory.

Our first and second major hypotheses were that, in accord with the findings of Maisto and Queen (1992) and aligning with both the 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. We left open the possibility, however, that familiarity with sports could moderate these predictions.

---

[9] Note that possible sports familiarity scores ranged from 9 to 55. The lowest quartile of sports familiarity scores encompassed scores 9–11 and 9–10 in Experiments 1A and 1B, respectively. The top quartile of sports familiarity scores encompassed scores 36–55 and 29–55 in Experiments 1A and 1B, respectively. Thus, the range of scores was greater in the top quartile than the bottom.

Experiment 1A 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 our 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 1A 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. Moreover, when unfamiliar with the stimuli, participants can enhance memory for images using elaborative encoding.

The results of Experiment 1B require a more nuanced explanation. We reasoned *a priori* 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 1B, the PSE was sharply reduced in participants who lacked knowledge of the sports teams: When all test stimuli were presented in word format, we observed a performance advantage for the words and combined groups relative to the symbols group 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 1B (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 1B by all but 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 associations. Sports familiarity would play the largest role in this condition because the memory test contained team names not provided at encoding for that 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 1B.

The drastic drop in performance for the symbols condition in Experiment 1B (where recognition was text-based) suggests an important caveat to the PSE: One *must* be able to easily associate, or generate, the verbal label for an image when the studied content is later only available for recognition by its 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 more likely to evoke image representations in memory. Typically, concrete words are more likely than abstract words to evoke imagery; our study demonstrates an exception to this proposal. 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). We speculate that the magnitude of the PSE depends on familiarity being "just right": The participant must be able to readily retrieve the meaning and corresponding verbal label for an image if the test is word-based, but they must not be overly familiar with the concepts that the words represent, lest the words automatically elicit mental imagery for associated content, effectively equating to encoding of pictures.

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 takeaway is that while visual memory may be poor for well-known stimuli, familiarity with a logo or picture remains critical in determining what the imaged content represents. Therefore, while the minutiae of brand logos themselves 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. Other findings in our laboratory demonstrate that even restricting the set size of the words to a single category (e.g., vegetables) does not undermine superior memory for symbols (e.g., !@#$%; Roberts et al., 2022). 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 our other studies of memory for symbols.

Henderson and Cote (1998) noted that logos add value to a company only if (a) consumers remember seeing the logo and (b) 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 1B 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 our results indicate a boundary condition in explicit memory associations when familiarity is low and retrieval is text-based. We 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, integrated text provides two forms of encoding support, ensuring that unacquainted consumers still can encode the brand name into memory.

In conclusion, sports team logos were better remembered than the text of a team name, and 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 preexisting logo–name associations. Results from participants with high familiarity suggest that the PSE can be eliminated if brand familiarity is already high. Consequently, we 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.

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

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

Bergstrom, J. A. (1893). Experiments upon physiological memory by means of the interference of associations. *The American Journal of Psychology*, *5*(3), 356–369. https://doi.org/10.2307/1410999

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: Human Experimental Psychology*, *68*(5), 858–865. https://doi.org/10.1080/17470218.2014.1002798

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

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*, *11*(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. *The Journal of Consumer Research*, *11*(2), 643–654. https://doi.org/10.1086/209001

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

Delacre, M., Lakens, D., & Leys, C. (2017). Why psychologists should by default use Welch's t-test instead of Student's t-test. *Revue Internationale de Psychologie Sociale*, *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. *Revue Internationale de Psychologie Sociale*, *32*(1), 1–12. https://doi.org/10.5334/irsp.198

Ensor, T. M., Surprenant, A. M., & Neath, I. (2019). 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. https://doi.org/10.3758/BF03193146

Hamilton, M., & Geraci, L. (2006). The picture superiority effect in conceptual implicit memory: A conceptual distinctiveness hypothesis. *The American Journal of Psychology*, *119*(1), 1–20. https://doi.org/10.2307/20445315

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/10.1177/002224299806200202

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

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

Kirkpatrick, E. A. (1894). An experimental study of memory. *Psychological Review*, *1*(6), 602–609. https://doi.org/10.1037/h0068244

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

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* (V. 0.6.1) [Computer software].

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

McDaniel, M. A., & Pressley, M. (1987). *Imagery and related mnemonic processes: Theories, individual differences, and applications*. 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. *The American Journal of Psychology*, *112*(1), 113–146. https://doi.org/10.2307/1423627

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* (V. 0.9.12-4.2) [Computer software].

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), 217–228. 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

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. 45–76). Lawrence Erlbaum.

Nelson, D. L., Bajo, M. T., & Casanueva, D. (1985). 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. (1985). 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 and 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. (1971). *Imagery and verbal processes*. Holt, Rinehart, and Winston.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45*(3), 255–287. https://doi.org/10.1037/h0084295

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

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*, Article 101452. https://doi.org/10.1016/j.socec.2019.101452

Roberts, B. R. T., MacLeod, C. M., & Fernandes, M. A. (2022). *Symbol superiority: Why $ is better remembered than 'dollar' [Manuscript submitted for publication].*

Roediger, H. L., III, & 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

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

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 and Memory*, *4*(6), 592–604. https://doi.org/10.1037/0278-7393.4.6.592

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

Weldon, M. S., & Roediger, H. L., III. (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. In I. V. Coltheart (Ed.), *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

Received November 9, 2021
Revision received June 24, 2022
Accepted October 9, 2022 ■