11-2010

Memory for pictures: Sometimes a picure is not worth a single word

Joyce M. Oates
American University

Lynne M. Reder
Carnegie Mellon University, cherdema@andrew.cmu.edu

Published In
Benjamin, A. S. (Ed.), Successful Remembering and Successful Forgetting: A Festschrift in Honor of Robert A. Bjork. New York: Psychology Press, 447- 462.

Memory for pictures: Sometimes a picture is not worth a single word

Joyce M. Oates[a] and Lynne M. Reder[b]

[a]American University

[b]Carnegie Mellon University

Mailing Address:

Department of Psychology

Carnegie Mellon University

Pittsburgh, PA 15213

e-mail: jo5436a@student.american.edu or reder@cmu.edu

Preparation of this chapter was supported by grant R01-MH52808 from the National Institute of Mental Health to the second author.

Words: 5,586

Memory for pictures: Sometimes a picture is not worth a single word

Although there are some phenomena in memory that are poorly understood, it is generally accepted that an item studied as a picture will be better remembered than an item studied as a word. Most of us are familiar with the expression, “a picture is worth a thousand words” and agree that memory for pictures is remarkable. Indeed in most situations, it would be wise to assume that pictures will be better remembered than words. One picture can contain enough information to convey many sentences. Children learn to follow stories in picture books before they are able to comprehend written text. Pictures are often universal and not restricted by knowledge of a specific language. For example, even if a traveler cannot read the local language, s/he can find the ladies’ or the men’s room in virtually any country by looking at picture signs.

Conventional wisdom amongst many memory theorists also holds that pictures are better remembered than words on recognition tests (e.g., Ally & Budson, 2007; Anderson, 2009; Brady, Konkle, Alvarez, & Oliva, 2008; Mintzer & Snodgrass, 1999; Nelson, Reed, & Walling, 1976; Shepard, 1967; Schacter, Israel, & Racine, 1999; Standing, 1973). Paivio’s Dual Code Theory (Paivio & Csapo, 1973) explains this seemingly ubiquitous phenomena as follows: when pictures are studied, they elicit their verbal label and thus two representations or codes are stored in memory. In contrast, words do not automatically elicit a picture, and thus they have a relatively impoverished memory representation. The redundant representation for pictures makes their retrieval or recognition more probable compared to stimuli studied as words. The explanation offered by Paivio implies that there are limits on the picture superiority effect; however memory experts sometimes only remember the general phenomenon and forget the theoretical explanation for its occurrence.[1]

In this chapter we will re-examine the conditions under which picture memory is superior to word memory. We will review evidence from the literature and present data of our own work that demonstrate when picture recognition is neither superior to, nor even comparable to word recognition. As part of the discussion of the attributes of situations that cause picture-inferiority, we will elaborate Paivio's Dual Code Theory. First, we review the conditions under which pictures are better remembered than words. Next we discuss the conditions that produce the picture-inferiority effect.

Pictures are better remembered than words.

The finding of better memory for pictures compared to words was reported as early as the 19th century (Kirkpatrick, 1894). Kirkpatrick demonstrated that real objects were better remembered than either written or spoken words both tested immediately, and at a 3-day delay. This picture superiority effect (PSE), as it has come to be called, is a robust phenomenon with numerous demonstrations of the basic finding that pictures are better recognized and recalled than their labels (e.g., Brady et al., 2008; Madigan, 1974, 1983; Nelson, Reed, & Walling, 1976; Nickerson 1965, 1968; Paivio, 1991; Paivio & Csapo, 1973; Paivio, Rogers, & Smythe, 1968; Shepard, 1967). In addition, the number of pictures that can be remembered is striking. Standing (1973) showed that people can remember thousands of unique pictures with great accuracy.

Brady et al. (2008) explored the limits of the finding reported by Standing by using a two alternative forced choice procedure to measure the nature of subjects' ability to discriminate targets from foils. They showed subjects 2,500 pictures of concrete objects, followed by a recognition test that compared three different types of foils that could be paired with the target: the "Novel" condition paired an old item with a new unrelated item; the "Exemplar" condition required a discrimination between an old item and a foil that shared the same concept (e.g. an "old" picture of *bread* paired with a "new" picture of *bread*), and the "State" condition where an old item was shown with another picture of the *same* item in a different "state" (e.g. a melon shown whole and a new picture depicting the *same* melon, half eaten). Discrimination accuracy was best (92.5%) when the foil was Novel but accuracy was surprisingly good in the Exemplar and State conditions as well (87.6 vs. 87.2%, respectively). In order to discriminate the studied picture from a semantically similar foil, subjects must have encoded and retained visual details of the stimulus such as color, shape, texture, etc.

The most popular and frequently cited explanation for PSE is Paivio's Dual Code Theory (Paivio, 1971; 1986; 1991; Paivio & Csapo, 1973; Paivio, Rogers, & Smythe, 1968). The theory's basic premise is that there are 2 codes for pictures, one that is visual (imagen) and one that is verbal (logogen; Morton, 1969) and only one for words (verbal). Pictures have perceptual information (i.e. colors, shapes etc…) and also verbal information (i.e. picture of a “dog”). Together the 2 codes increase the memory strength of pictures because there can usually be two ways to represent any one pictorial item. Pictures have a “naming” advantage since labels are often automatically elicited, whereas images for words are not generated without explicit instruction and additional mental effort. Paivio and Csapo demonstrated that when subjects are instructed to generate images when studying lists of pictures and words, words are remembered as accurately as pictures. This showed that words can benefit from dual code, but only with effortful mental imaging. Conversely, it has been reported that using a rapid presentation rate (5.3 items per second) of pictures and words eliminates PSE. This is because subjects are slower at naming a picture than reading a word (Fraisse, 1968), and the rapid presentation does not allow subjects enough time to generate a label, so pictures lose the dual code advantage.

Pictures are more perceptually rich than words, and this visual distinctiveness lends them an advantage in memory. To the extent that subjects also encode the stimulus as a verbal label, subjects have two codes for pictures: in addition to the perceptual features of the stimulus such as color, shape, and texture, subjects also store a verbal label (similar to the representation for a studied word), that enriches the memory trace and provides redundancy. Picture illustrations are included in textbooks because they corroborate text and are often more effective than text alone for problem-solving transfer (Mayer, 1984; 1989; 1993; Mayer, Steinhoff, Bower, & Mars, 1995). Words are visually sparse, as the letters are usually presented in one color (black) in a common font (e.g. Times New Roman, Helvetica). Not only is the visual image of a word impoverished, but the common font has been seen with thousands if not millions of other words and provides no visual distinctiveness. Unlike a picture that tends to be automatically labeled, words rarely stimulate the generation of an image (Paivio & Csapo, 1973; Kroll & Potter, 1984).

Consider the situation in which memory is measured using an old/new recognition paradigm. If a picture is tested as a word, the test word is likely to match the word that was automatically generated and stored. However, if a word is tested as a picture, it is unlikely that the test picture will match the one a subject generated (in the rare case that one was generated). If memory is measured using free recall, the Dual Code Theory of Paivio can explain why pictures produce more free recall. The two potential cues (visual and verbal) to remember a picture make accurate recall more probable. In summary, it seems logical that a picture of anything (e.g. a balloon) is more likely to be remembered than its label.

Sometimes words are better remembered than pictures.

Despite all the evidence for the picture superiority effect (PSE) there are counter examples. We conducted an experiment whose goal was not to test the PSE, but the results frequently elicited comments from memory experts who were surprised by the superior memory for words than for photographs and pictures. Reder et al. (2006) conducted a recognition memory experiment that tested recognition memory performance under the drug midazolam compared with saline control for three types of stimuli: photographs, abstract pictures and words. Midazolam is a benzodiazepine and an anxiolytic that causes temporary anterograde amnesia. We believe that midazolam causes this memory impairment by blocking the ability to form new bindings in memory (Park, Quinlan, Thornton, & Reder, 2004). We conjectured that it would be easier to bind words to the experimental context than it would be for unfamiliar faces, places or abstract pictures.

After receiving an injection of either midazolam or saline (placebo), a study list was presented one stimulus at a time for the subject to rate for pleasantness. The list consisted of words displayed in a common font, pictures of unfamiliar abstract paintings, and unfamiliar photographs (non-famous faces, cityscapes, and landscapes). After completing the rating task and a filler task, subjects were tested on their ability to recognize the studied/rated items compared with foils that had not been seen before. Subjects served as their own control (receiving midazolam at one session and saline at the other). This meant that subjects completed two separate occasions and rated different exemplars of each stimulus class for each session (stimuli of each type were randomly assigned for each subject to be targets or foils for session 1 or session 2). The order of drug conditions was counter-balanced over subjects and the experimenter and subject were blind to drug condition. (see Reder et al., 2006 for more details).

It would be reasonable to assume that under saline, memory for pictures would be best and under midazolam, memory for pictures would be the most attenuated (since midazolam hurts recollection and pictures would have two representations to strengthen memory). In contrast to the conventional wisdom, Figure 1 shows that under saline, accuracy for abstract pictures was worst, followed by photographs, and words were the best recognized. Furthermore, under midazolam, words showed the greatest decline, followed by photographs. Recognition memory for abstract pictures was unaffected by drug condition. Why should words be better remembered under the placebo and show the greatest decrement under the amnestic drug when pictures have a two-representation advantage? ---Insert Figure 1 about here---

Pictures that do not have a meaningful label are hard to remember.

As we discussed earlier, the abstract pictures used in Reder et al., could not be easily labeled. That is, a good descriptor of the picture would not easily come to mind that would be unique ("abstract picture" would not distinguish among the abstract pictures, targets or foils). The label or descriptor not only needs to be unique but the subject needs to regenerate the same label at test for it to be useful. Even if a subject is successful for a particular picture, the process requires more effort (working memory resources) to create labels that are not automatically elicited (already assigned to the picture).

It is much easier for a subject to generate rather generic labels such as “abstract picture”, or the name of colors used in the painting (red, blue, pink etc…), or a general shape such as “rectangle” or “blob.” Those labels are not very useful since they are unlikely to discriminate one target from another or from a foil. Without a concrete label to aid memory, perceptual features alone are not sufficient to make abstract pictures as memorable as words.

Abstract pictures tend to be processed as the conjunction of their features or a subset of their features that attract attention at any one encoding. With many repetitions, more and more pieces of the design are bound together to form larger and larger chunks. Famous abstract pictures often have labels consisting of either the title of the piece or the artist who created it. Novel abstract pictures are much more difficult to encode holistically and recognize later. They do not benefit from a dual code as there is no pre-existing verbal code for them. To summarize, it is our position (see also Reder et al., 2006; Reder, Park, Kieffaber, 2009) that unfamiliar abstract pictures are ambiguous and do not foster labels. When subjects attempt to generate one they may fail to bind it to context because cognitive resources are exhausted or they may fail to generate the same one when the stimulus is next seen. Converging evidence for this position comes from the developmental literature, described below.

If you cannot name it, you will probably not remember it.

Ducharme and Fraisse (1965) did not find superior memory for pictures for 7 and 8 year old children when using free recall. They suggested that children do not automatically label pictures, much like subjects were not able to automatically label abstract pictures in Reder et al. (2006). Later, Jenkins, Stack, and Deno (1969) reported that second grade students exhibited PSE for recognition, but that PSE was reversed for free recall; words were better recalled than pictures. They speculated that children must be at a disadvantage for labeling and recall of those labels is more difficult than recognizing a reinstated picture.   It is also possible that they found spelling very difficult but when they had viewed the word earlier, the spelling was primed (reinforced) making it easier to output the word than the picture. Whitehouse and colleagues (2006) tested students from 4 different age groups: 2nd & 3rd graders, 4th & 5th graders, 7th & 8th graders, and 10th & 11th graders. The subjects were given a mixed list of concrete words and pictures and then were asked to recall as many of the stimuli as they could. Whereas the number of words recalled remained constant across all 4 groups, the number of pictures recalled increased with age. Younger subjects were clearly able to read, comprehend, and retain verbal information, since recall for words was as good at the youngest age compared to all the older ages. Whitehouse et al. concluded that their results were due to younger children lacking ability to use inner speech. We expand on that conclusion and hypothesize that the diminished PSE for the younger subjects is due to an inability to maximize dual code. Younger subjects must not be able to label pictures as easily or as automatically as older subjects, and if a label is not generated at study then recalling a picture concept at test becomes very difficult.

Further evidence that the ability to label a picture affects its subsequent memory comes from Robertson and Köhler (2007). They tested children aged 4-6 years and discovered that the children’s ability to label a picture affected subsequent recognition. Those pictures that children could successfully name aloud during encoding were more likely to be remembered at test. Empirical work in our lab using famous faces supports this finding. Walsh, Xu, Dutcher, Oates & Reder (in preparation) used pictures of famous and non-famous people and discovered that famous people who were identified as being famous by subjects were better remembered when subjects could generate the name of the person depicted in the photograph. In summary, when people find it easy to generate unique labels at encoding that they can reproduce at the time of memory retrieval, memory is better. We believe that this is because the label can be bound to context, enabling recollection as well as familiarity-based judgments.

If the picture label is not discriminative, the verbal code is useless.

Another result from the Reder et al. 2006 study, shown in Figure 1, is that the hit rate is lower for photographs than for words, while the false alarm rate is higher. The pattern of results is similar to those found for the word-frequency mirror effect: high frequency words have lower hit rates and higher false alarm rates than low frequency words (Glanzer & Adams, 1985). We speculate that photographs are behaving like high frequency (higher familiarity) words. The concepts, or labels, of the photographs in the experiment were few and had been encountered over and over again in the study phase, making them highly familiar in the experimental context. In other words, subjects saw many different photographs, but they came from a small number of concept categories (e.g. pictures of landscapes divided into only a handful of sub categories, such as mountains, deserts, bodies of water, forests). Subjects could only rely on visual information to discriminate old from new items ("I saw many pictures of mountains, but did I see a picture of *this* particular mountain?"). The labels, if used, are not distinctive because they are used over and over again with many similar photographs. In other words, the verbal part of dual code to help discriminate "old" from "new" photographs was useless. Even if one ignores the problem that the verbal labels are useless in this situation, the high perceptual familiarity/similarity among stimuli also interferes with familiarity-based recognition. Familiarity based responding is used in absence of remembering contextual details. In absence of contextual details, or episodic retrieval, a person can use the level of activation of a concept to help indicate if an item was previously experienced. If the level of activation of the concept is high enough, then it can be used as an index to make a memory judgment. Given a high level of concept activation, a subject can determine that an item is sufficiently familiar and therefore must have been experienced recently, despite lacking details about the experience. However, the familiarity-based process is more vulnerable to false alarms, and is less accurate than using contextual cues to discriminate amongst items. The familiarity-based process does not help discriminate between an item that is highly familiar because it has been frequently experienced in the past, and an item that is highly familiar because it has been recently experienced.

In a clever experiment, Chandler (1994) demonstrated that familiarity with similar pictures hurts accuracy while increasing confidence. She showed subjects segmented portions of unique nature pictures (e.g., “lake thawing”, “sand dunes”). The pictures were divided into 3, dimensionally equal portions (A, B, and C) that were all visually similar since they were from the same scenic photograph. One of the three was presented during study (A). For some of the triples, a second portion B was shown between the study and test phase. At test, subjects had to discriminate the originally viewed segment A from the never viewed segment C. When subjects saw the related counterpart B between study and test, they were *less* accurate but *more* confident in their discrimination between A and C.

One interpretation of this result is that more associations to the same stimulus label (e.g., Lake scene) negatively impacted ability to discriminate. Or, the memory trace "I saw a lake scene" was reinforced more when there was a second presentation of a lake scene. It seems likely that the specific perceptual details could be confused from one presentation to the next and generally create more interference.

Further evidence for high concept familiarity having a negative effect on picture memory comes from Goldstein and Chance (1970). They tested recognition memory for snowflakes, inkblots, and pictures of faces. They reported that memory was best for faces, followed by inkblots, with accuracy worst for snowflakes. Goldstein and Chance predicted faces to be best remembered since subjects have expertise in recognizing human faces. They expected memory for snowflakes to be worst because of configuration complexity (defined as the number of “corners” or “turns” contained in an item).

Complexity may not be the only variable that affects retention. It is worth noting that the stimulus sets differed in their degree of homogeneity. That is, all pictures of snowflakes would be labeled “snowflake” such that there would be no added value of the verbal code. According to Goldstein and Chance, inkblots should be more memorable than snowflakes that have more detail that must be encoded. To expand on that hypothesis, we propose that whereas the verbal label for snowflakes was useless, inkblots could be coded by similarity to real world objects much like the practice of labeling/interpreting inkblots for a Rorschach Test. Likewise, even non-famous faces could be labeled by gender, race, age, similarity to friends, etc.

Comparing picture vs. word recognition when the picture foils share the verbal code.

We reviewed evidence earlier (Reder et al., 2006) that when the picture label is *generic*, meaning that it is shared with many other targets and foils, performance is worse than for words. Now we consider the case in which the picture label is distinctive (i.e., not shared with other studied items) but the foils are other pictures that represent the same meaning or label. Will the perceptual richness of the picture when combined with meaning be better than the abstracted meaning of a word when the visual information is all that allows discrimination of a target from a foil? We hypothesize that the dual code is of little value when the only discriminative facet is the picture representation even though there are no other stimuli encoded that share that meaning.

Work in our lab, previously unpublished, has examined memory for words vs. pictures of comparable words (concepts are randomly assigned to be shown as a picture vs word for each subject). We wanted to give words a “fighting chance” so each word was presented in a unique font and that font was used to present the word at test as well. Our goal was to make words more visually distinctive and to reduce the benefit of pictures by diminishing the benefit of the conceptual aspect of the dual code for pictures.

Some of our prior work has demonstrated that visual distinctiveness modulates memory for words. Arndt and Reder (2003) and Park, Arndt and Reder (2006) showed that the number of words that share a font inversely affects memorability. Therefore, as the number of words that share a font increases, the distinctiveness of the font decreases, and so does the memorability of those particular words. Words that are presented in reinstated, relatively unique fonts are recognized more often than words that are reinstated in unusual fonts shown with other words in the experiment. Therefore we expect that presenting words in unique fonts and reinstating those fonts at test will boost recognition memory for words while using pictures of words whose foils share the same meaning will reduce memory for those pictures.

Experiment 1. Subjects were presented with a list of 60 pictures and 60 words. Each item was displayed for 2 seconds and the orienting task involved subjects indicating "Typically, how often would you encounter the concept depicted by this item?" The response choices were either "Very Rarely", "Rarely", "Frequently" or "Everyday". After the encoding task and a short break, subjects were given a surprise recognition memory test. They were told that some of the items would be "old" and that some would be "new", and were explicitly instructed to respond "old" to a picture only if it was *the exact picture* that they saw previously.

Test stimuli consisted of 30 "old" pictures, 30 "old" words, 30 "new" pictures, and 30 "new" words. The lure words shared the same unique font shown with the words, but the meaning was changed. Studied words were shown in the same unique font seen earlier. No font was shown twice. In other words, foil pictures had the identical concept but a different image while foil words had the identical font but a different concept.

As predicted, d′ prime was not statistically different and words were as memorable as pictures, $F(1, 15) = .54$, $p > .05$. Subjects had reliably higher hit rates for pictures than words, $F(1,15) = 14.4$, $p < .05$, but also reliably higher false alarm rates $F(1,15) = 33.2$, $p < .01$, which is not surprising because subjects were most likely responding "old" to the concept, not the visual information despite explicit instruction to only judge based on visual information. The higher false alarm rate indicates that subjects were not always able to remember exact visual information that corresponded with a label. By relying on the familiarity of the concept ("I know I saw a picture of a 'house'") subjects were depending on the label to make a judgment when they should have been judging the visual content only.

Experiment 2 attempted to address the problem of task switching between words and pictures with different rules by creating separate blocks for words and pictures memory tests. We used an ABBA design in which A was a test for words and B a test for pictures for half of the subjects, and the reverse for the other half. Other than that, the experiment was the same as Experiment 1 (see Table 1 for a summary of the data from both experiments). In the first experiment, picture memory performance might have been attenuated from switching tasks between words and pictures. It might have been too difficult for subjects to decide if they saw an exact picture (foils shared the same concept as targets), since previously viewed words were always presented in a reinstated font and could be discriminated by concept alone. Using the blocked design, the hit rate for pictures increased slightly while the false alarm rates decreased: 95.2% hit rate, up 2.1%, 7.7% false alarm rate, down 15.2%. However, the same pattern occurred for words: 90.4% hit rate up 14.4%, 3.9% false alarm rate down .7%. Again, d' was not reliably different between the two modalities, $F(1,19) < 1.0$.

---Insert Table 1 about here---

It is noteworthy that Reder and Thornton's findings in Exp. 2 are similar to those reported by Brady et al (2008). Recall that Brady et al. showed subjects 2,500 pictures and then administered a 2 alternative forced choice test. In their task, like ours, memory depended on the visual code alone for pictures because lures shared concepts with targets. The recognition test compared three different types of foils that could be paired with the target. Directly related to our experimental manipulation is the “Exemplar” condition that required a discrimination between an old item and a foil that shared the same concept (e.g. an “old” picture of *bread* paired with a “new” picture of *bread*). In Brady et al., subjects’ accuracy was 87.6% in the “Exemplar” condition. When we apply the same correction, subtracting false alarms from hit rates, our subjects also performed at the same corrected accuracy (87.5%).

The results of Reder and Thornton’s two experiments underscore our claim that when the picture label is not discriminative, the verbal code is useless. Similar picture foils (same label) at test forced subjects to rely on visual information to make old/new judgments. This made pictures more comparable to words in that only one code, in this case, the image, could be used to discriminate targets from foils. In most recognition experiments involving words only the verbal code is available for discriminating targets from foils. In summary, these two experiments demonstrate a constraint to Dual Code Theory: if the label of the picture is not discriminative, pictures will be no better remembered than words.

Conclusions

In this chapter we have illustrated that the Picture Superiority Effect in recognition memory, as explained by Paivio’s Dual Code Theory, does not always hold. The failure to observe the PSE can also be understood within Paivio's Dual Code Theory by postulating the variables that affect (a) when people are able to generate a second code, and (b) when that code will or will not be helpful. Other researchers have also demonstrated poor memory performance for pictures (Amrhein, McDaniel, & Waddill, 2002; Weldon & Roediger, 1987), although their interpretations are different from ours. Our elaborations to Paivio's Dual Code are thus: the picture must afford a meaningful label that discriminates it from other test probes. „When those labels are shared with other test items, a dual code does not offer a memory advantage.

To summarize: PSE occurs when items are presented as distinctive, easy to label pictures and the foils do not share the same labels (Brady et al. 2008; Standing, 1973). However, when a picture's verbal code is shared with other pictures, the conceptual fan effect makes retrieval of the encoding episode difficult and makes spurious recognition based on the concept's familiarity likely. Remembering that you saw a picture of a "cat" is not going to help if you must discriminate between a picture of the previously shown cat and a different "cat" at test. Subjects must rely only on the image code to determine if an item is old or new. When the visual stimulus is difficult to identify, that is, generation of a consistent label is not easy or possible, ("abstract picture" will not suffice if there are many such stimuli) the picture is not even as memorable as a single word.

References

Ally, B.A. & Budson, A.E. (2007). The worth of pictures: Using high density event-related potentials to understand the memorial power of pictures and the dynamics of recognition memory. *NeuroImage, 35*, 378-395.

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

Anderson, J.R. (2009). *Cognitive Psychology and its Implications: Seventh Edition*. Worth Publishers, New York.

Arndt, J. & Reder, L.M. (2003). The effect of distinctive visual information on false recognition. *Journal of Memory and Language, 48*, 1-15.

Brady, T.F., Konkle, T., Alvarez, G.A. and Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences, 105*(38), 14325-14329.

Chandler, C.C. (1994). Studying related pictures can reduce accuracy, but increase confidence, in a modified recognition test. *Memory & Cognition, 22*, 273–280.

Ducharme, R., & Fraisse, P. (1965). Etude genetique de la memorisation de mots et d'images. *Canadian Journal of Psychology, 19,* 253-61.

Fraisse, P. (1968). Motor and verbal reaction times to words and drawings. *Psychonomic Science, 12*, 235-236.

Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory & Cognition, 13,* 8–20.

Goldstein, A.G. & Chance, J.E. (1970). Visual recognition memory for complex configurations. *Perception and Psychophysics, 9*, 237-241.

Jenkins, J.R., Deno, S.L., & Stack, W.B. (1969). Children's recognition and recall of picture and word stimuli. *AV Communications Review, 17*, 265-271.

Kirkpatrick, E.A. (1894). An experimental study of memory. *Psychological Review, 1*, 602-609.

Kroll, J.F., & Potter, M.C. (1984). Recognizing words, pictures, and concepts: A comparison of lexical, object, and reality decisions. *Journal of Verbal Learning and Verbal Behavior, 23*, 39-66.

Madigan, S. (1974). Representational storage in memory. *Bulletin of the Psychonomic Society, 4*, 567-568.

Mayer, R.E. (1984). Aids to text comprehension. *Educational Psychology, 19*, 30-42.

Mayer, R.E. (1989). Models for understanding. *Review of Educational Research, 59* 43-64.

Mayer, R.E. (1993). Illustrations that instruct. In R. Glaser (Ed.), *Advances in Instructional Psychology,* Volume 5. Hillsdale, JH: Erlbaum.

Mayer, R.E., Steinhoff, K., Bower, G., & Mars, R. (1995). A generative theory of textbook design: Using annotated illustrations to foster meaningful learning of science text. *Educational Technology Research and Development, 43*(1), 31-41.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review, 76,* 165-178.

Nelson, D. L., Reed, U. S., & Walling, J. R. (1976). Pictorial superiority effect. *Journal of Experimental Psychology: Human Learning & Memory, 2*, 523-528.

Nickerson, R.S. (1965). Short-term memory for complex meaningful visual configurations, a demonstration of capacity. *Canadian Journal of Psychology, 19,* 155-60.

Nickerson, R.S. (1968). A note on long-term recognition memory for pictorial material. *Psychonomic Science, 11*, 58.

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

Paivio, A (1986). *Mental representations: a dual coding approach*. Oxford. England: Oxford University Press.

Paivio, A. & Csapo, K. (1973). Picture superiority in free recall: Imagery or dual coding? *Cognitive Psychology, 5*(2) 176-206.

Paivio, A., Rogers, T.B., & Smythe, P.C. (1968). Why are pictures easier to recall than words? *Psychonomic Science, 11*, 137-138.

Paivio, A.D. (1991). *Images in the Mind, the Evolution of a Theory*. Harvester Wheatsheaf, New York.

Park, H., Arndt, J.D., & Reder, L.M. (2006). A contextual interference account of distinctiveness effects in recognition. *Memory & Cognition, 34*(4), 743-751.

Park H., Quinlan, J.J., Thornton, E.R., & Reder, L.M. (2004). The effect of midazolam on visual search: Implications for understanding amnesia. *Proceedings of the National Academy of Sciences, 101*(51), 17879-17883.

Reder, L.M., Nhouyvansivong, A., Schunn, C.D., Ayers, M.S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember/know judgments in a continuous recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 26*(2), 294-320.

Reder, L.M., Oates, J.M., Thornton, E.R., Quinlan, J.J., Kaufer, A., & Sauer, J. (2006). Drug induced amnesia hurts recognition, but only for memories that can be unitized. *Psychological Science, 17*(7), 562-567.

Reder, L.M., Park, H., & Kieffaber, P.D. (2009). Memory systems do not divide on consciousness: Reinterpreting memory in terms of activation and binding. *Psychological Bulletin, 135*(1), 23-49.

Robertson, E.K. & Köhler, S. (2007). Insights from child development on the relationship between episodic and semantic memory. *Neuropsychologia, 45*(14), 3178-3189.

Schacter, D.L, Israel, L., & Racine, C. (1999). Suppressing false recognition in younger and older adults: The distinctiveness heuristic. *Journal of Memory and Language, 40*(1), 1-24.

Shepard, R. (1967). Recognition Memory for Words, Sentences and Pictures. *Journal of Verbal Learning and Verbal Behavior*, *6,* (156-163).

Standing, L. (1973). Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology, 25*, 207-222.

Walsh, M.M., Xu, J., Dutcher, J.M., Oates, J.M., & Reder, L.M. (in preparation). ERP correlates of familiarity and recollection in face recognition.

Weldon, M.S. & Roediger, H.L. (1987). Altering retrieval demands reverses the picture superiority effect. *Memory & Cognition, 15*(4), 269-280.

Whitehouse, A.J.O., Maybery, M.T., & Durkin, K. (2006). The development of the picture-superiority effect. *British Journal of Developmental Psychology, 24*, 767–773.

Author Note

Preparation of this chapter was supported by grant R01-MH52808 from the National Institute of Mental Health. The authors would like to thank Edward Thornton for help on Experiments 1 & 2, and Christopher Paynter and Anna Manelis for comments on the manuscript.

Footnote

1. In casual conversation we found that a number of internationally recognized memory experts were vulnerable to the erroneous expectation that any picture should be better recognized than any word. That is, they were surprised when presented with data that were inconsistent with the picture superiority effect.

Table 1. Results from Experiments 1 & 2. Mean Proportion of Hits, False Alarms, and *d′* as a Function of Item.

	Item Type	Mean Proportion Response		*d′*
		Hits	False Alarms	
Experiment 1	Pictures	0.93	0.23	2.46
	Words	0.76	0.05	2.67
Experiment 2	Pictures	0.95	0.08	3.30
	Words	0.90	0.04	3.44

Figure Caption

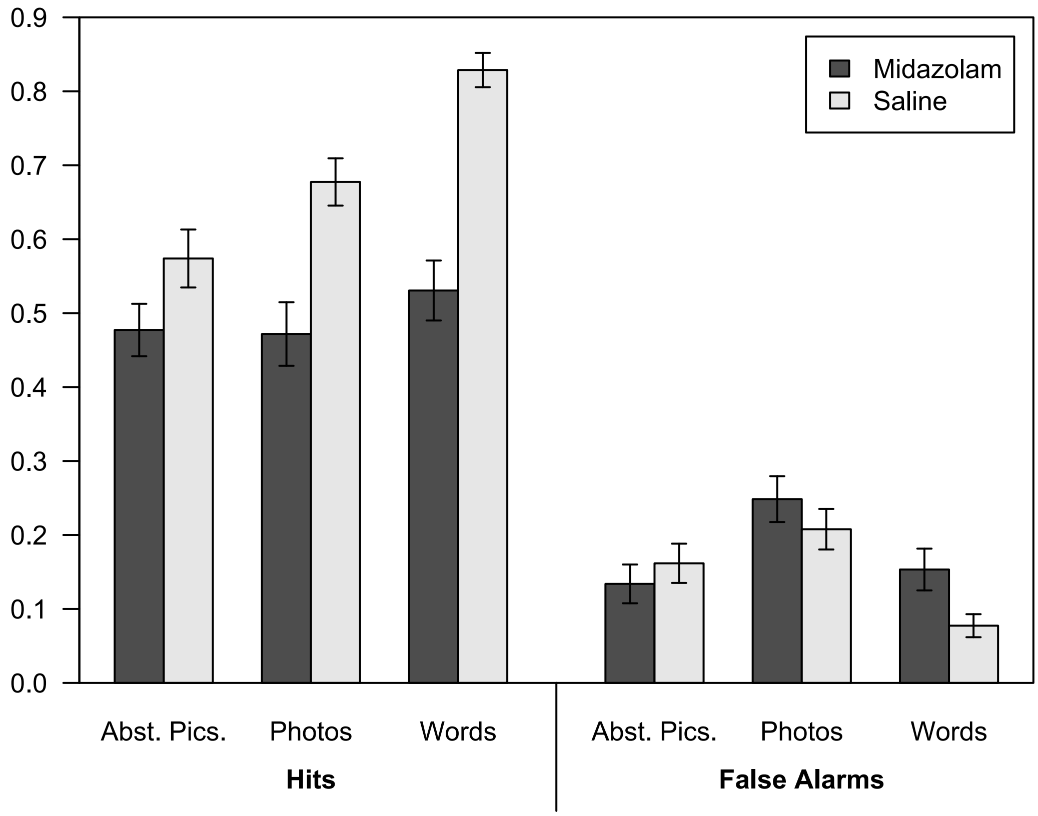

Figure 1. Proportion of hits and false alarms for abstract pictures, photographs, and words as a function of drug condition. From “Drug induced amnesia hurts recognition, but only for memories that can be unitized,” by L.M. Reder, J.M. Oates, E.R. Thornton, J.J Quinlan, A. Kaufer, & J. Sauer, 2006, Psychological Science, 17, p. 565. Copyright John Wiley & Sons, Inc. Reprinted with permission.