An auditory analog of the picture superiority effect

Robert J. Crutcher · Jenay M. Beer

Published online: 2 November 2010
© The Psychonomic Society 2010

Abstract Previous research has found that pictures (e.g., a picture of an elephant) are remembered better than words (e.g., the word "elephant"), an empirical finding called the *picture superiority effect* (Paivio & Csapo, *Cognitive Psychology* 5(2):176–206, 1973). However, very little research has investigated such memory differences for other types of sensory stimuli (e.g. sounds or odors) and their verbal labels. Four experiments compared recall of environmental sounds (e.g., ringing) and spoken verbal labels of those sounds (e.g., "ringing"). In contrast to earlier studies that have shown no difference in recall of sounds and spoken verbal labels (Philipchalk & Rowe, *Journal of Experimental Psychology* 91(2):341–343, 1971; Paivio, Philipchalk, & Rowe, *Memory & Cognition* 3(6):586–590, 1975), the experiments reported here yielded clear evidence for an auditory analog of the picture superiority effect. Experiments 1 and 2 showed that sounds were recalled better than the verbal labels of those sounds. Experiment 2 also showed that verbal labels are recalled as well as sounds when participants imagine the sound that the word labels. Experiments 3 and 4 extended these findings to incidental-processing task paradigms and showed that the advantage of sounds over words is enhanced when participants are induced to label the sounds.

Keywords Dual-coding theory · Picture superiority effect · Auditory imagery · Imagery · Memory · Recall

R. J. Crutcher (✉)
Department of Psychology, University of Dayton,
300 College Park,
Dayton, OH 45469-1430, USA
e-mail: crutcher@udayton.edu

J. M. Beer
Georgia Institute of Technology,
Atlanta, GA, USA

The *picture superiority effect* refers to an empirical finding in which pictures are remembered better than words (Paivio, 1979; Paivio & Csapo, 1969; Paivio & Csapo, 1973; Paivio, Rogers, & Smythe, 1968; Shepard, 1967). For example, a picture of an elephant will be remembered better than the word *elephant* when tested with free recall. What is surprising is that even though learners recall the presented items by writing down the words for them, more of the presented pictures than of the words are recalled. According to dual-coding theory, the source of the memory advantage for pictures versus words is that pictures are coded twice in the cognitive system: first, in a sensory-based, visual code and then, in a symbolic, verbal code; words, on the other hand, are coded once in a verbal code (Paivio, 1975, 1991; Paivio & Csapo, 1973). Given, the dual-coding explanation, it follows that other sensory stimuli such as sounds or odors might also be remembered better than verbal labels for those sensory stimuli. Unfortunately, there has been little research exploring this possibility for sounds. The purpose of the experiments reported here was to investigate the possibility of an auditory analog of the *picture superiority effect*, or what we shall call the *auditory picture superiority effect*. In other words, are environmental sounds (e.g. the sound of thunder) remembered better than the words that refer to these sounds (e.g. the word "thunder")?

The picture superiority effect has been demonstrated for a range of memory tasks, as well as under a variety of conditions (Madigan, 1983; Paivio & Csapo, 1969, 1973; Whitehouse, Maybery, & Durkin, 2006). It has been demonstrated with free recall (e.g., Paivio & Csapo, 1973; Paivio et al., 1968), cued recall (e.g., Nelson & Reed, 1976), serial recall (e.g., Nelson, Reed, & McEvoy, 1977; Paivio & Csapo, 1969), recognition memory (Paivio & Csapo, 1969, 1973), and implicit memory tasks such as

picture fragment completion (Weldon & Roediger, 1987) and implicit category production (Wippich, Melzer, & Mecklenbrauker, 1998). The effect has also been obtained with intentional and incidental instructions (Paivio & Csapo, 1973), and the effect increases with developmental age (Whitehouse et al., 2006). In short, the picture superiority effect is a robust empirical finding. However, under some conditions, the picture superiority effect can be eliminated or reversed. For example, with serial recall tasks, words are sometimes recalled better than pictures, particularly when stimulus presentation rates are fast (Boldini, Russo, & Punia, 2007; Paivio & Csapo, 1969, 1971). More recently, experiments using implicit memory tasks (e.g., implicit category production) have sometimes eliminated the effect (Weldon & Coyote, 1996) or reversed it, producing a word superiority effect (Weldon & Roediger, 1987; Weldon, Roediger, & Challis, 1989).

There are various explanations for the picture superiority effect, including explanations that suggest that pictures trigger deeper semantic levels of processing than do words (Nelson, 1979; Nelson & Brooks, 1973; Nelson & Reed, 1976; Nelson et al., 1977; Nelson, Reed, & Walling, 1976) as well as explanations that emphasize the match between encoding and retrieval processes to explain the effect (McBride & Doshier, 2002; Weldon & Roediger, 1987; Weldon et al., 1989). However, the classic explanation of the effect is based on dual-coding theory (Paivio, 1979, 1983, 1990, 1991, 2007). Dual-coding theory is a general theory with many specific assumptions and hypotheses that taken together, can explain these various results. Most specifically, dual-coding theory accounts for the picture superiority effect because of its assumption of independent and additive representational codes. When words are read or heard, they are generally encoded verbally, whereas when pictures are processed, participants are more likely to encode the stimulus visually and by verbally labeling the picture. With an item represented in two codes, the likelihood of accessing it from memory increases. However, activation of representational codes is probabilistic and optional, not automatic, and therefore various situational, task, and participant variables can influence which codes are activated and when. For example, the reversal of the picture superiority effect with serial recall is accounted for by the theory's assumption that the verbal system is better at sequential processing tasks than is the nonverbal imagery system. Similarly, the lack of a picture superiority effect with faster presentation rates is explained by specific processing assumptions of the theory: A picture must activate its corresponding *imagen* before cross-activation of the corresponding verbal representation (the *logogen*) can occur. Therefore, under faster stimulus presentation rates, labels for pictures are less likely to be activated.

The most important assumption of dual-coding theory for the present research is the multisensory nature of the nonverbal representation systems. In a review of the dual-coding theory literature, Paivio (1991) noted that the "multiple sensorimotor side of the theory has often been overlooked, perhaps because DCT (Dual Coding Theory) research has emphasized visual imagery more than other modalities, just as imagery researchers have done." However, dual-coding theory is not limited to verbal and visual codes; nonverbal codes include additional sensory-based codes, such as auditory and olfactory codes. If words are used to label other sensory stimuli, it follows that the picture superiority effect should extend to other nonverbal, sensory-based stimuli, such as odors or sounds. Does it?

Some research on memory for odors has shown that immediate retention of odors is poor and is not enhanced by verbal labels, although odors that are retained tend to be quite durable over long-term retention intervals (Engen & Ross, 1973; Richardson & Zucco, 1989). Other research has shown rather good retention for odors (Lawless & Engen, 1977; Rabin & Cain, 1984) and has shown that retention of odors is enhanced when verbal labels are assigned to odors (Lyman & McDaniel, 1990; Rabin & Cain, 1984). Lyman and McDaniel examined several different forms of elaborative encoding for odors, including verbal and visual elaboration, as well as encoding odors alone. They reported that odors were recognized more readily than verbal labels for those odors. They also found that memory was better when additional elaboration, visual or verbal, was performed at encoding, a finding that the authors acknowledged is consistent with dual-coding theory.

There has been relatively little research comparing memory for auditory and verbal stimuli. Miller and Tanis (1971) investigated recognition memory for sounds, spoken verbal labels of sounds, and printed verbal labels. They tested each type of material in the same mode as that for learning, using a forced choice recognition test. Recognition memory for sounds was poorer than that for printed or spoken verbal labels. Moreover, recognition memory for the sounds was less than 70%, on average, somewhat low, as compared with recognition memory for pictures. However, Lawrence and Banks (1973) found that recognition memory for sounds was quite good. Unfortunately, they did not compare memory for sounds with that for verbal labels. Using a yes/no recognition procedure, Huss and Weaver (1996) compared memory for recorded sounds with that for spoken verbal labels. They manipulated the match between encoding and retrieval presentation mode so that recognition was tested in the same or the opposite mode. Overall, recognition was better for sounds than for words; however, recognition was best of all when sounds were presented but

the testing mode was spoken words. However, the recognition tests were conducted after free recall testing, which may have influenced the recognition results. There are very few studies that have compared *recall* of sounds and verbal stimuli (Huss & Weaver, 1996; Paivio et al., 1975; Philipchalk & Rowe, 1971; Rowe, Philipchalk, & Cake, 1974). Philipchalk and Rowe compared free and serial recall of sounds and verbal labels. For serial recall, recall was greater for verbal items than for sounds, a finding similar to that for serial recall of pictures versus verbal items (Paivio & Csapo, 1969). However, free recall of sounds was not greater than recall of verbal labels, in contrast to the frequently found recall advantage for pictures over words. In subsequent research, Paivio et al. (1975) investigated free and serial recall of nonverbal stimuli (sounds or pictures) versus verbal labels (printed or spoken words). In two experiments, they found that *overall* recall of nonverbal items (sounds and pictures) exceeded recall of verbal items (printed and spoken words) but that, with serial recall, verbal items were recalled better than nonverbal items. However, in a focused comparison of the recorded sound stimuli and their corresponding spoken verbal labels, the pattern of results was less clear. In their first experiment, Paivio et al. (1975) reported that the mean number of sounds recalled ($M = 9.51$) tended to be greater than the number of spoken words ($M = 8.75$) recalled. However, the difference was small and apparently not significant, since no statistics are reported for the comparison. In their second experiment, the authors reported that free recall of sounds and spoken verbal labels did not significantly differ. In summary, Paivio and colleagues found that recall for sounds was not significantly different than recall for corresponding spoken verbal labels of those sounds. In contrast, Huss and Weaver (1996) found that recall was greater for sounds than for verbal labels of those sounds. Thus, whether sounds are recalled better than verbal labels remains an open question.

The purpose of the experiments reported here was to reinvestigate the possibility of an auditory picture superiority effect but to enhance the possibility of obtaining the effect by employing a within-subjects design for one of the studies and increasing the number of stimuli and participants in the experiments employing between-subjects designs. Experiment 1 tested the possibility of an auditory picture superiority effect, using a within-subjects design. The prediction was that sounds would be recalled better than the spoken verbal labels for those sounds, as predicted by dual-coding theory. Experiment 2, employing a between-subjects design, further tested the dual-coding explanation by adding a third stimulus-encoding condition in which participants listened to spoken verbal labels of sounds and imagined hearing

the corresponding sounds for the labels. On the basis of the earlier Paivio and Csapo (1973) study of pictures and words, the prediction was that environmental sounds would be recalled better than spoken verbal labels; however, when asked to *imagine* hearing the corresponding sounds for the words, the words would be recalled as well as the sounds. Experiment 3 tested whether the auditory picture superiority effect obtains under incidental- and intentional-processing paradigms, as has been found for the picture superiority effect. Finally, Experiment 4 tested another implication of dual-coding theory: specifically, whether labeling auditory sounds would enhance the auditory picture superiority effect, relative to repeating the words, and how both of these were compared with a free strategy condition in which sounds and words were coded in whatever way participants chose. The prediction derived from dual coding theory was that labeling the sounds would produce greater recall than would repeating the words or using other strategies.

Experiment 1

The goal of Experiment 1 was to test whether sounds are remembered better than spoken verbal labels of those sounds. Given the mixed findings of earlier studies, all of which employed between-treatment designs, this experiment used a within-subjects design.

Method

Participants Twenty-four undergraduate students at the University of Dayton participated in this experiment as part of a requirement for an introductory psychology course.

Apparatus and stimulus materials The stimuli consisted of 40 recorded sounds and 40 spoken verbal labels corresponding to the sounds (e.g. a recording of the sound of thunder and a recording of someone saying the word "thunder").¹ The sounds were taken from a CD of recorded sound effects for radio station use; the spoken verbal labels were made by recording one of the experimenters reading the words into a high-quality digital recording system. The sound levels of all the stimuli were edited and normalized so that they were equal in loudness and the durations of the

¹ From earlier sound studies, 52 sounds likely to be easily and correctly identified were selected for the present experiments. Eight randomly selected pilot participants then listened to our recordings of those sounds, writing down a label for each sound. The 40 best sound stimuli—those that listeners easily and correctly identified—were selected for these experiments.

recordings were 1–2 s in length. Each recorded stimulus included a brief fade-in and fade-out, to avoid sudden changes in volume. To equalize the presentation time of the sounds and spoken verbal labels, silence was added to the end of each sound or spoken word to fill out the time to 6 s. After the editing process, all the stimuli were converted to MP3 format at a 192-kb/s sampling rate. The items were randomly divided into two groups, and two lists were constructed so that list 1 was composed of 20 sounds and 20 verbal labels and list 2 was composed of the same 40 items, but with the assignment of sounds and verbal labels conditions reversed (see the [Appendix](#)). The equipment used to run the study included Dell Microcomputers running Windows 2000 with software to present the recorded sounds used in the experiment, as well as individual headphones connected to each computer over which the recordings were heard.

Procedure and design The experiment was a single factor within-subjects design with stimulus type (sound versus spoken verbal label) as the independent variable and proportion of items recalled as the dependent measure. Each participant listened to 20 sounds and 20 spoken verbal labels. Assignment of items to the sounds and verbal label conditions was counterbalanced across the two lists so that an item presented as a sound for list 1 was presented as a verbal label for list 2 and vice versa. Presentation order of the sounds and words was blocked so that each participant heard 20 sounds and then 20 spoken words or the reverse (20 spoken words and then 20 sounds). Stimulus assignment and presentation order were completely counterbalanced across participants, so that each stimulus was presented equally often as a recorded sound or a spoken verbal label and in the first or the second presentation block.

The experiment was conducted in a single session that lasted approximately 45 min, with participants run in groups of 5. The experimenter began by instructing participants that they would hear a series of “sound effects” and “spoken words” played through a set of headphones connected to a computer. The stimuli would be presented one at a time, and participants were to listen to the recordings while keeping their eyes focused on a fixation point on the computer screen during the entire presentation. Afterward, participants were to recall the stimuli by writing a single word for each presented sound or spoken word. The experimenter presented several example sounds and spoken words, one at a time, on an overhead projection display and asked participants to recall the items. The experimenter answered questions about the procedure and then instructed the participants to start up the presentation program on their individual computers. Participants listened to 20 spoken words and 20 sounds presented 1 at a time for

6 s each, with the presentation advancing immediately and automatically from one trial to the next. The 6 s included the 1- to 2-s duration of the sound itself, along with the silence that filled out the trial to exactly 6 s. After all 40 stimuli had been presented, a tone sounded, and participants recalled the items by writing their responses on the provided response sheet. They were again instructed to write a single word for each stimulus they recalled. After 3 min, a second tone sounded, cuing participants to stop writing and to give their response sheets to the experimenter.

Results and discussion

The recall responses were scored using a procedure similar to that used in previous studies of memory for sounds (e.g., Philipchalk & Rowe, 1971). Responses that were the same as the predetermined verbal labels for the sounds but in another form (e.g. “laugh” or “laughed” instead of “laughing”), as well as synonyms or words that clearly referred to the sound (e.g., “applause” for “clapping” or “belch” for “burp”) were acceptable. We also analyzed the responses using a stricter coding scheme that did not allow synonyms. The pattern of the means and the results of all the statistical analyses were similar for both sets of scores.

As was predicted, the average proportion of sounds recalled ($M = .48$, $SD = .11$) was greater than that for recall of the spoken verbal labels ($M = .29$, $SD = .13$), $t(23) = 5.07$, $p < .001$, $d = 1.534$. In contrast to earlier studies that showed no difference between sounds and spoken verbal labels (Paivio et al., 1975; Philipchalk & Rowe, 1971), there was a clear recall advantage for recorded sounds over the corresponding verbal labels for those sounds. In short, our results for auditory stimuli mirror the findings for visual stimuli in the picture superiority effect.

Experiment 2

Experiment 2 had two goals. The first was to directly test an implication of dual-coding theory. According to dual-coding theory, when someone sees a picture of a chair, they code a visual representation of the stimulus; at the same time, they are likely to label or name the object storing a verbal representation of the stimulus as well. However, in the case of words, we are much less likely to image or picture the referent of the word. Thus, pictures are often remembered better than words, because two codes are better than one. This suggests an interesting test of the dual-coding account: If participants read the word “chair” and visualized a chair while reading, they would, in effect, code the stimulus visually and verbally, just as in the picture condition. The effect should be to equalize memory for

pictures and words. In fact, previous research has demonstrated this result (Paivio & Csapo, 1973). Analogously, if the auditory picture superiority effect reported here is explained by dual coding, imagining the sound of thunder while studying the word “thunder” should enhance recall of the words, relative to a verbal-label-only condition.

The second goal of Experiment 2 was to replicate the effect obtained in Experiment 1, which used a within-subjects design, with a between-subjects design. Previous studies of sounds versus verbal labels have employed between-subjects designs but have failed to find significance, except for the Huss and Weaver (1996) study which used 46 participants and 15 stimuli in each condition. Here, we used a sample size of 24, but with 40 stimuli per condition, hoping to increase the sensitivity of the experiment in detecting an auditory picture superiority effect.

Our predictions were the following: Recall of actual sounds would exceed recall of verbal labels, replicating the result in Experiment 1; however, listening to spoken words and imagining the sounds for those words would enhance recall of the verbal items and produce recall more similar to that in the sound-alone condition.

Method

Participants Seventy-two University of Dayton undergraduate students participated in this experiment as part of a requirement for an introductory psychology course.

Materials and apparatus The stimulus materials and apparatus were the same as those in Experiment 1, except that each list consisted entirely of 40 sounds or 40 verbal labels for those sounds.

Procedure and Design The design was a single factor between-subjects design with type of stimulus as the independent variable. The three stimulus conditions were *sound*, *verbal label*, and *verbal label plus imagined sound*. Each participant was randomly assigned to one of the conditions and listened to 40 sounds, 40 spoken verbal labels, or 40 spoken verbal labels while *imagining* hearing the sound to which the word referred (e.g., for the word “thunder,” they were to imagine hearing the sound of thunder).

The procedure was the same as that in Experiment 1. Participants were run in groups of 5–10. The instructions for the sound and verbal label conditions were the same as before, but each group received only the instructions relevant to that condition. For the verbal-label-plus-imagined-sound condition, participants were instructed to listen to the word and “try to imagine hearing the sound that the word refers to.” This single phrase was the only difference between the instructions for the spoken word

and the spoken-word-plus-imagined-sound conditions. As in Experiment 1, immediately after participants had listened to all 40 items, a tone sounded, and they had 3 min to recall them.

Results and discussion

The average proportion of items recalled for the sound condition ($M = .45$, $SD = .11$) exceeded that for the verbal label condition ($M = .32$, $SD = .11$), $t(46) = 4.026$, $p < .001$, $d = 1.187$. In addition, recall for the verbal-label-plus-imagined-sound condition ($M = .45$, $SD = .09$) was greater than recall for the verbal label condition ($M = .32$, $SD = .11$), $t(46) = 4.500$, $p < .001$, $d = 1.326$. However, recall for the verbal label plus imagined sound condition ($M = .45$, $SD = .09$) was not significantly different from that for the sound condition ($M = .45$, $SD = .11$), $t(46) = 0.146$, $p = .884$.

First, as was predicted, average recall for the verbal-plus-imagined-sound condition exceeded recall for the verbal-labels-alone condition. Second, imagining the sounds for the words increased recall for the actual sounds. Finally, recall was greater for sounds than for the verbal labels, extending the auditory picture superiority effect to a between-subjects design.

Experiment 3

Experiment 2 showed that when asked to imagine the sound referenced by a verbal label, participants recall verbal labels as well as sounds, extending Paivio and Csapo’s (1973) finding for visual stimuli to auditory stimuli and strengthening a dual-coding account of the auditory picture superiority effect. In addition, Experiments 1 and 2 showed that, in general, sounds are recalled better than the verbal labels for those sounds. However, in both of these experiments, participants were informed prior to encoding the stimuli that they would be tested on their memory for the items. Research on the picture superiority effect has demonstrated the effect under incidental, as well as intentional, free recall task instructions (Paivio, 1979, 2007; Paivio & Csapo, 1973). The effect has been found even with an incidental-processing task that does not require the participants to code the items in any specific way but simply to predict, on each trial, whether the presented stimulus will be a picture or a word (Paivio & Csapo, 1973).

The goal of Experiment 3 was to compare recall of sounds versus words under incidental and intentional memory instructions with an incidental-processing task that engaged participants’ attention, requiring them to register the sounds and words, but not to further elaborate on the

stimuli or expect a memory test. Thus, on each trial, participants heard a sound or a word and predicted whether a sound or a spoken word would occur on the next trial. In addition, the intentional group was informed that they would be tested on their memory for the items. The main prediction was that, overall, sounds would be recalled better than words but that recall would be greater when participants anticipated a memory test than when they did not. Consistent with Paivio and Csapo's (1973) earlier findings, it was expected that, overall, recall for both sounds and words would be lower given the engagement of attention in an orienting task that discouraged higher level encoding of items.

Method

Participants Seventy-two University of Dayton undergraduate students participated in this experiment as part of a requirement for an introductory psychology course or as an extra-credit assignment for another psychology course.

Materials and apparatus The stimulus materials were the same as those used in Experiments 1 and 2. Apple Macintosh computers running the 10.5.8 version of the OS X operating system presented the recorded stimuli.

Procedure and Design The experiment was a two-way mixed factorial design with stimulus type (sound vs. spoken verbal label) as the within-subjects factor and encoding processing task (incidental vs. intentional) as the between-treatment factor. The proportion of items recalled was the dependent measure. Participants were randomly assigned to either an incidental or an intentional condition. Each participant listened to 20 sounds and 20 spoken verbal labels. However, unlike in the previous two experiments, in which presentation of sounds and words was blocked, in the present experiment, the presentation of the stimulus items was mixed and randomized for each participant.

The experiment was conducted in a single session lasting approximately 25–45 min. The procedure was the same as before, with a few changes. First, no mention of a memory test was made until just before the computer presentation of the actual stimuli. Second, participants were run individually, with the participant sitting at the computer and the experimenter to the side and behind the participant. After explaining, as before, how the stimuli would be presented one at a time, the experimenter stated, "While listening to the stimuli, your task is to predict which type of stimulus will appear on each subsequent trial. If you think the stimulus will be a *sound* then you should say 'sound' if you think the next stimulus will be a *word* then say 'word'." Participants went through several

practice trials on the computer, listening to a stimulus then predicting whether the next stimulus would be a sound or a word. Questions about the procedure were then answered. At this point, the intentional group received one additional instruction before the actual presentation began: "One more thing, before we start. After you finish listening to all of the items, you will be tested on your memory for the items." This was the only difference in the instructions for the two groups.

At this point, the computer stimulus presentation program started up, and participants predicted whether the first item would be a sound or a spoken word by saying "sound" or "word." They then listened to the recorded stimulus, which was presented for 6 s, including the interval of silence. During this interval participants noted whether their response was correct or not and predicted what the next item would be. Immediately after all the items had been presented, a tone sounded and participants had 3 min to recall the items on a response sheet.

Results and discussion

The results of a mixed analysis of variance, revealed that the average proportion of items recalled for the sound condition ($M = .23$, $SD = .11$) exceeded the proportion of items recalled for the verbal label condition ($M = .16$, $SD = .10$), $F(1, 70) = 21.24$, $MSE = .009$, $p < .001$, $\eta_p^2 = .233$. Recall for the intentional group ($M = .22$, $SD = .08$) was greater than that for the incidental-processing group ($M = .17$, $SD = .08$), $F(1, 70) = 5.14$, $MSE = .013$, $p = .027$, $\eta_p^2 = .060$. However, the interaction of encoding condition and stimulus type was not significant, $F(1, 70) = 1.019$, $MSE = .009$, $p = .316$.

The results were generally in accord with the predictions from dual coding theory. Sounds were recalled better than words overall and recall was better under intentional than incidental instructions, although this difference was small. Additional comparisons confirmed that for the incidental group recall of sounds ($M = .22$) was greater than recall of verbal labels ($M = .13$), $F(1, 35) = 17.6$, $MSE = .008$, $p < .001$, $\eta_p^2 = .335$; likewise, for the intentional group, recall of sounds ($M = .24$) exceeded that of the verbal labels ($M = .19$), $F(1, 35) = 5.87$, $MSE = .010$, $p = .021$, $\eta_p^2 = .144$.

As was expected, overall, recall of sounds and words was lower here than in the other experiments, confirming that the incidental-processing task successfully engaged participants' attention. We specifically chose this orienting task because Paivio and Csapo's (1973) research on the picture superiority effect argued that the task induced participants "to register the pictures cognitively as unelaborated images and the words as verbal traces," but did not

“require active encoding of items to higher levels.” Our pilot testing confirmed that this task fully engaged participants: Not only did participants report that predicting the next stimulus left little time to reflect on the stimulus, but also some participants initially fell behind in their predictions of the subsequent stimuli and realized that they had to more quickly register the current stimulus and predict the next one. These observations were further confirmed in postexperimental interviews that asked participants what they had done to remember the items. Most reported difficulty doing much of anything because they were so engaged in predicting what the next stimulus would be. In short, our results perfectly parallel those of Paivio and Csapo (1973) for the picture superiority effect. In summary, despite a decrease in overall recall performance, sounds were recalled better than words under incidental- and intentional-processing instructions that discouraged extensive elaboration.

Experiment 4

It is important to remember that according to dual-coding theory, the activation of representational codes is probabilistic and optional, not automatic: Situational, task, and participant variables influence which codes are activated and when. Although dual-coding theory assumes that we are much more *likely* to verbally label a picture (or a sound) than we are to form a visual or auditory image to a word, these probabilities can be influenced by task instructions, as was demonstrated in Experiment 2. When instructed to imagine the sounds for the spoken words, participants remembered the spoken words just as well as the sounds. This result strongly supports a dual-coding account of the auditory picture superiority effect and suggests a further test of the account: If labeling sounds is responsible for their memory advantage, then *instructing* participants to label the sounds should enhance recall of sounds even more, whereas similar instructions applied to the spoken words should not benefit word recall much at all; in fact, they might reduce memory for the words by decreasing the likelihood of forming auditory images.

Furthermore, if *labeling* the sounds is responsible for the enhanced recall of sounds but participants are already likely to label the sounds, being informed of a memory test when asked to label the sounds would not enhance recall of the sounds much, as compared with labeling the sounds without the expectation of a memory test. Extending this logic, a free strategy condition that encouraged participants to encode the sounds in any way they wished should not produce higher recall than would simply labeling the sounds; on the other hand, it may produce higher recall of

spoken words, because not having to write down the words would allow participants to employ auditory imaging and other strategies that could improve recall of the words.

Method

Participants One hundred eight University of Dayton undergraduate students participated in this experiment as part of a requirement for an introductory psychology course or an extra-credit assignment for another psychology course.

Materials and apparatus The stimulus materials and apparatus were the same as those used in Experiment 3.

Procedure and Design The experiment was a two-way mixed factorial design with stimulus type (sound vs. spoken verbal label) as the within-subjects factor and encoding processing task (incidental, intentional, or free strategy) as the between-treatment factor. As in the previous experiments, the dependent variable was proportion of items recalled. The experimental procedure was nearly identical to that used in Experiment 3, except for the changes to the processing instructions. The experiment was conducted in a single session that lasted approximately 25 to 45 min. Participants were randomly assigned to the incidental, intentional, or free strategy conditions. The only difference among the three groups was the instructions they received. Participants in all three conditions were told, “In this experiment, you are going to be listening to a series of recorded ‘sound effects’ and ‘spoken words’ through a set of headphones connected to the computer. These stimuli will be presented one at a time at a constant rate of 6 seconds each.” At this point, the incidental- and intentional-processing groups were instructed, “While listening to the stimuli, your task is to write down the name or label for that sound. When you hear a sound effect, write down the name or label for that sound. When you hear a spoken word, write down the word you heard spoken. Please write only a single word. Do you understand what I’d like for you to do?” The free strategy group was told, instead, “While listening to the stimuli, your task is to do whatever you can to try to remember each item. After you have finished listening to all of the items, you will be tested on your memory for the items. Do you understand what I’d like for you to do?” All three groups then practiced the tasks they were to perform with example stimuli on the computer. After answering any questions about the participant’s assigned task, the experimenter started the computer program that ran the experiment. At this point, as in Experiment 3, participants in the intentional group were

told that after listening to all of the items, they would be tested on their memory for the items.

Each participant listened to 20 sounds and 20 spoken words, with the presentation of the items mixed and the order randomized individually, as in Experiment 3. Immediately after all 40 stimuli had been presented, a tone sounded, and participants recalled as many of the items as possible. Responses were typed on the computer, instead of being written on a response sheet, as in the previous experiments.

Results and discussion

Average proportion of recall of the stimuli as a function of the type of stimulus (sounds vs. words) and processing condition (incidental, intentional, and free strategy) are displayed in Fig. 1. The results of a mixed analysis of variance showed that the average proportion of items recalled for the sound condition ($M = .44$) exceeded the proportion of items recalled for the verbal label condition ($M = .25$), $F(1, 105) = 140.03$, $MSE = .014$, $p < .001$, $\eta_p^2 = .571$. In addition, recall for the incidental ($M = .37$), intentional ($M = .36$) and free strategy ($M = .32$) groups differed, as indicated by a significant overall effect of processing task on recall, $F(2, 105) = 5.23$, $MSE = .012$, $p = .007$, $\eta_p^2 = .091$. However, there was also a significant interaction of the stimulus type and processing type factors, $F(2, 105) = 9.793$, $MSE = .014$, $p = .001$, $\eta_p^2 = .157$.

Given the significant interaction, follow-up analyses were conducted. Simple effects testing for the stimulus variable showed that sounds were recalled better than words for

the incidental group ($M = .51$ vs. $M = .24$), $F(1, 35) = 73.34$, $MSE = .017$, $p < .0001$, $\eta_p^2 = .677$; for the intentional group ($M = .46$ versus $M = .25$), $F(1, 35) = 76.25$, $MSE = .011$, $p < .0001$, $\eta_p^2 = .685$; and for the free strategy group ($M = .36$ vs. $.27$), $F(1, 35) = 11.14$, $MSE = .014$, $p = .002$, $\eta_p^2 = .241$. Simple effects testing for the processing task factor showed a simple main effect of processing task for the sounds, $F(2, 105) = 13.76$, $MSE = .014$, $p < .001$, $\eta_p^2 = .208$, but not for the words, $F(2, 105) = 0.166$, $MSE = .012$, $p = .847$. Additional post hoc comparisons for sounds showed that sound recall for the incidental condition ($M = .51$) was greater than that for the free strategy condition ($M = .36$), $t(105) = 5.109$, $p < .001$, and sound recall for the intentional condition ($M = .46$) exceeded that for the free strategy condition ($M = .36$), $t(105) = 3.586$, $p = .001$. However there was no difference in recall of sounds for the incidental- and intentional-processing conditions, $t(105) = 1.523$, $p = .131$. Given the absence of a simple effect of processing task for the verbal items, no additional comparisons were performed. In short, recall of words did not differ for the incidental ($M = .24$), intentional ($M = .25$), and free strategy ($M = .27$) conditions.

The results of Experiment 4 extend the auditory picture superiority effect and further strengthen the dual-coding account of the effect. First, across three different encoding process instructions (incidental, intentional, and free strategy), sounds were recalled better than spoken words. Second, as was predicted, labeling sounds and words enhanced memory for the sounds, relative to the words. Nearly twice as many sounds were recalled as words for both the incidental- and intentional-processing conditions. In addition, labeling the sounds and knowing that there would be a memory test (intentional group) did not improve recall over simply labeling the sounds (incidental group), whereas not labeling the sounds and using whatever strategy they wished (free strategy group) significantly lowered recall of sounds, as compared with spoken words. These results are all consistent with a dual coding account of the auditory picture superiority effect. The one expected result that did not materialize was that words were not recalled better in the free strategy condition than in the incidental- and intentional-processing conditions, as dual-coding theory might predict.

One further result merits comment. Recall for the sounds in the free strategy condition ($M = .36$) not only was lower than that for the sounds in the intentional ($M = .51$) and incidental ($M = .46$) conditions in the present experiment (as was predicted) but also was lower than recall of sounds in Experiment 1 ($M = .48$) and Experiment 2 ($M = .45$). Why was the recall of sounds in the free strategy condition ($M = .36$) not as high as that for the sounds in Experiments

Fig. 1 Mean proportion of items recalled as a function of stimulus type and encoding condition in Experiment 4

1 and 2? We do not know. One possibility is that, overall, recall was lower in Experiment 4 than in Experiments 1 and 2. This is supported by the lower free strategy recall not only for the sounds, but also for the words. Thus, the recall for the free strategy group represents the recall level for an “intentional” group that was not instructed to label the sounds and words. This possibility, although admittedly speculative, could be tested by running a new experiment that included the three groups here (i.e., incidental and intentional groups that labeled the stimuli and a free strategy group that did not label the stimuli) as well as an additional intentional group that did not perform the labeling task. We predict that recall of sounds would be quite similar for the intentional (no-labeling) and free strategy groups but that recall for both of these groups would be lower than that for the incidental-and intentional-labeling groups.

General discussion

The main goal of the experiments reported here was to test several of the assumptions of dual-coding theory that explain the picture superiority effect and predict that the effect should occur for auditory stimuli as well. In contrast to earlier studies that have investigated memory for environmental sounds versus words and have failed to show a difference in recall, the experiments reported here provide clear evidence for superior recall of sounds over words. Experiment 1, using a within-subjects design, showed a reliable auditory picture superiority effect in which recorded sounds were recalled better than the spoken verbal labels of those sounds. The second experiment replicated this finding with a between-subjects design. Both findings are consistent with a dual-coding explanation in which sounds are coded auditorily and verbally, whereas words are coded only verbally. Strengthening this interpretation, Experiment 2 also showed that when spoken verbal labels were presented and participants were asked to imagine hearing the corresponding sounds, recall performance for the words equaled that for the sounds. This result parallels Paivio and Csapo’s (1973) earlier finding for the picture superiority effect, in which the advantage of pictures over verbal items was eliminated by asking participants to visually imagine the object corresponding to the presented verbal label (Paivio, 1983).

The results in Experiment 3 further strengthen the dual-coding interpretation by demonstrating that sounds are recalled better than words with an incidental-processing task that did not require elaborative processing of the stimuli. Participants simply predicted whether they would

hear a sound or a spoken word on each trial. That this task fully engaged participants’ cognitive resources is suggested by the decrease in overall recall, as compared with Experiments 1 and 2. Nonetheless, recall for sounds still exceeded that for the spoken words. Finally, Experiment 4 demonstrated that an incidental-processing task that induces participants to label sounds *strengthens* the effect. Twice as many sounds as words were recalled. Furthermore, being informed of a memory test added nothing over simply labeling the sounds. There was no difference in the recall of sounds for the incidental and intentional groups. Allowing participants to encode the sounds and words in any way they wished, as they did in the free strategy condition, also did not improve performance, as compared with simply labeling the sounds. In fact, the free strategy condition showed *lower* recall of the sounds than did the groups that labeled the sounds. In short, the results from Experiment 4, as well as those from Experiment 2, strongly suggest that multiple, independent memory codes seems the best explanation of why sounds were recalled better than words in these experiments.

Although the present experiments were specifically designed to test dual-coding theory, naturally, there are other possible explanation for an auditory picture superiority effect. A sensory-semantic account (Nelson, 1979; Nelson et al., 1976, 1977) might try to explain the memory advantage for sounds over words by arguing that sounds trigger deeper semantic levels of processing than do words or that the auditory codes for sounds are more distinct from one another than are the verbal codes for different words are. However, neither of these explanations seems sufficient to explain the present results. The dramatic improvement of recall of sounds by labeling them, as compared with encoding them, however they wished in the free strategy condition (Experiment 4) seems difficult to explain with semantic elaboration. Nor does deeper processing seem a compelling explanation of why sounds were recalled better than words when participants were actively engaged in a task predicting whether the next stimulus would be a sound or a word (Experiment 3). A distinctiveness account might argue that words share a fixed set of phonemes, whereas environmental sounds vary along several different acoustic dimensions, such as pitch and timbre. However, the experiments here did not specifically test this possibility and, thus, cannot really speak to the distinctiveness account. Moreover, we would suggest that distinctiveness may be more relevant to tasks in which item differentiation is more important, such as recognition memory, which we also did not test here.

Another class of explanations that has been proposed as explaining the picture superiority effect that might be

offered to explain the auditory picture superiority effect emphasizes the match between encoding and retrieval processes (McBride & Doshier, 2002; Weldon & Roediger, 1987; Weldon et al., 1989). However, these accounts have been invoked to explain differences in implicit and explicit memory task versions of the picture superiority effect. Our tasks here, employing free recall, were all explicit memory tasks. Obviously, an important goal for future research will be to determine whether the auditory picture superiority effect occurs with other memory tasks, including recognition and cued recall, as well as various implicit memory tasks. However, the Huss and Weaver (1996) investigation of sounds and words that we mentioned previously may be relevant to accounts that emphasize the match between encoding and retrieval processes. Huss and Weaver examined recognition memory for sounds and words and found that, overall, sounds were recognized better than words. However, they completely crossed the encoding and testing presentation formats for their stimuli, so that each type of stimulus (sound or spoken word) was tested in the same and an opposite format. Sounds, for example, were encoded as sounds and were tested as sounds for half the subjects, but as spoken words for the other half; similarly, spoken words were encoded as spoken words and were tested as spoken words for half the subjects, but as sounds for the other half. Sounds were recalled better than words overall, but there was a significant interaction of encoding and testing format, with the highest recognition occurring for stimuli presented as sounds but tested as spoken words. This result seems more consistent with a dual-coding interpretation than with an explanation in terms of the match between encoding and retrieval processes.

There are a number of directions that future research on the auditory picture superiority effect might take. First, very few reported studies have investigated the possibility of an auditory picture superiority effect. Of the few studies reported in the literature, some have reported no difference in remembering sounds and words; others have reported an advantage for sounds over words. Here, we found clear evidence for an advantage of sounds over words. We obtained an auditory picture superiority effect using within- and between-subjects designs, whereas previous studies have used only between-subjects designs. However, as was noted earlier, the Huss and Weaver (1996) study used a fairly large sample size, with 46 participants per stimulus encoding condition. Although we used fewer (24) participants per condition in some of our experiments (Experiments 1 and 2), we used twice as many stimuli (40) than did previous studies. Taken together, Huss and Weaver's and our results suggest that the auditory picture superiority effect may not be as robust an effect as the visual picture

superiority effect and that it will have to be investigated further to understand why this might be. One possibility is that auditory stimuli are not as quickly or as easily labeled as visual stimuli. In fact, Clark, Stamm, Sussman and Weitz (1974) found a significant difference in recognition memory for relatively labelable and nonlabelable sounds and suggested that this difference indicated that "a dual coding process operates for sounds as well as for visual pictures and objects." They also suggested that difficulty in attaching labels to sounds may explain the failure of sounds to show an advantage over verbal labels in the Philipchalk and Rowe (1971) study.

Other possibilities for future research include testing the boundary conditions of the auditory superiority effect and extending the effect to other memory tasks besides free recall. For example, research on the picture superiority effect usually shows superior memory for pictures over words. However, with faster stimulus presentation rates or speeded response tasks, the effect can sometimes be eliminated or reversed (Boldini et al., 2007; Paivio & Csapo, 1969). The reason, according to dual-coding theory, is insufficient time to label the pictures. If sounds are more difficult or take longer to label, presentation rates might prove an important factor in whether the auditory picture superiority effect is observed or not. Studies that manipulated the rate of presentation of the sounds and labels seem warranted. Unlike pictures, though, which can be presented for extremely brief time intervals (e.g., a fraction of a second), sounds, which are recognized temporally, may provide interesting challenges in manipulating presentation times. Reversals of the picture superiority effect have also been obtained using implicit memory tasks. Investigations of the auditory picture superiority effect, using implicit memory tasks as well as traditional memory tasks such as cued recall and recognition, will be necessary to extend the findings here and establish the boundary conditions of the phenomenon.

In conclusion, the results reported here, all of which were consistent with a dual-coding explanation, extend the traditional picture superiority effect to a new sensory modality: the auditory domain. Future work must address to what extent these findings generalize to other memory tasks and whether or not dual-coding theory is the best account of these findings.

Acknowledgement The studies reported here were conducted at the University of Dayton. Parts of this research were reported at the annual meeting of the Psychonomic Society, Houston, November 2006, and at the Midwest Psychological Association Meeting, Chicago, May 2006.

We would like to thank Courtney Castle, Alison Danforth, Fitore Musmurati, Timothy Hawk, Maureen O'Marro, Jennifer Shea, and Allison Walden, for their help in conducting these studies and Bradlee Beer for his expert technical assistance in creating the auditory stimuli for the experiments. We would also like to thank Greg Elvers for his insightful comments on several drafts of the manuscript.

Appendix A

List 1	List 2
Sounds	Sounds
Airplane	Siren
Whistle	Footsteps
Clapping	Kiss
Gong	Quacking
Piano	Snoring
Mooring	Whip
Laughing	Hiccup
Gallop	Thunder
Cough	Hooting
Typing	Slurp
Scream	Gunshots
Gargle	Chirping
Motorcycle	Belch
Telephone	Shatter
Harp	Drum
Meow	Roar
Doorbell	Train
Wind	Sneeze
Barking	Heartbeat
Bells	Guitar
Words	Words
Siren	Airplane
Footsteps	Whistle
Kiss	Clapping
Quacking	Gong
Snoring	Piano
Whip	Mooring
Hiccup	Laughing
Thunder	Gallop
Hooting	Cough
Slurp	Typing
Gunshots	Scream
Chirping	Gargle
Belch	Motorcycle
Shatter	Telephone
Drum	Harp
Roar	Meow
Train	Doorbell
Sneeze	Wind
Heartbeat	Barking
Guitar	Bells

References

- Boldini, A., Russo, R., & Punia, S. (2007). Reversing the picture superiority effect: a speed-accuracy trade-off study of recognition memory. *Memory & Cognition*, 35(1), 113–123.
- Clark, M., Stamm, S., Sussman, R., & Weitz, S. (1974). Encoding of auditory stimuli in recognition memory tasks. *Bulletin of the Psychonomic Society*, 3(3-A), 177–178.
- Engen, T., & Ross, B. M. (1973). Long-term memory of odors with and without verbal descriptions. *Journal of Experimental Psychology*, 100(2), 221–227. doi:10.1037/h0035492.
- Huss, M. T., & Weaver, K. A. (1996). Effect of modality in earwitness identification: Memory for verbal and nonverbal auditory stimuli presented in two contexts. *The Journal of General Psychology*, 123(4), 277–287.
- Lawless, H., & Engen, T. (1977). Associations to odors: Interference, mnemonics, and verbal labeling. *Journal of Experimental Psychology: Human Learning & Memory*, 3(1), 52–59. doi:10.1037/0278-7393.3.1.52.
- Lawrence, D. M., & Banks, W. P. (1973). Accuracy of recognition memory for common sounds. *Bulletin of the Psychonomic Society*, 1(5-A), 298–300.
- Lyman, B. J., & McDaniel, M. A. (1990). Memory for odors and odor names: Modalities of elaboration and imagery. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(4), 656–664. doi:10.1037/0278-7393.16.4.656.
- Madigan, S. (1983). Picture Memory. In J. C. Yuille (Ed.), *Imagery, memory, and cognition: Essays in honor of Allan Paivio* (pp. 65–89). Hillsdale: L. Erlbaum Associates.
- McBride, D. M., & Doshier, B. A. (2002). A comparison of conscious and automatic memory processes for picture and word stimuli: A process dissociation analysis. *Consciousness and Cognition: An International Journal*, 11(3), 423–460. doi:10.1016/S1053-8100(02)00007-7.
- Miller, J. D., & Tanis, D. C. (1971). Recognition memory for common sounds. *Psychonomic Science*, 23(4), 307–308.
- 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). Hillsdale: Lawrence Erlbaum Associates.
- Nelson, D. L., & Brooks, D. H. (1973). Functional independence of pictures and their verbal memory codes. *Journal of Experimental Psychology*, 98(1), 44–48. doi:10.1037/h0034299.
- Nelson, D. L., & Reed, V. S. (1976). On the nature of pictorial encoding: A levels-of-processing analysis. *Journal of Experimental Psychology: Human Learning & Memory*, 2(1), 49–57. doi:10.1037/0278-7393.2.1.49.
- Nelson, D. L., Reed, V. S., & Walling, J. R. (1976). Pictorial superiority effect. *Journal of Experimental Psychology: Human Learning & Memory*, 2(5), 523–528. doi:10.1037/0278-7393.2.5.523.
- 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. doi:10.1037/0278-7393.3.5.485.
- Paivio, A. (1975). Coding distinctions and repetition effects in memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 9) (pp. 179–214). New York: Academic Press.
- Paivio, A. (1979). *Imagery and verbal processes*. Hillsdale: Lawrence Erlbaum Associates, Inc.

- Paivio, A. (1983). The empirical case for dual coding. In J. C. Yuille (Ed.), *Imagery, memory, and cognition: Essays in honor of Allan Paivio* (pp. 307–332). Hillsdale: L. Erlbaum Associates.
- Paivio, A. (1990). *Mental representations*. New York: Oxford University Press. doi:10.1093/acprof:oso/9780195066661.001.0001.
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, 45(3), 255–287. doi:10.1037/h0084295.
- Paivio, A. (2007). *Mind and its evolution: A dual coding approach*. Mahwah: Lawrence Erlbaum Associates, Inc.
- Paivio, A., & Csapo, K. (1969). Concrete image and verbal memory codes. *Journal of Experimental Psychology*, 80(2, Pt. 1), 27–285.
- Paivio, A., & Csapo, K. (1971). Short-term sequential memory for pictures and words. *Psychonomic Science*, 24(2), 50–51.
- Paivio, A., & Csapo, K. (1973). Picture superiority in free recall: Imagery or dual coding? *Cognitive Psychology*, 5(2), 176–206. doi:10.1016/0010-0285(73)90032-7.
- Paivio, A., Philipchalk, R., & Rowe, E. J. (1975). Free and serial recall of pictures, sounds, and words. *Memory & Cognition*, 3(6), 586–590.
- Paivio, A., Rogers, T. B., & Smythe, P. C. (1968). Why are pictures easier to recall than words? *Psychonomic Science*, 11(4), 137–138.
- Philipchalk, R. P., & Rowe, E. J. (1971). Sequential and nonsequential memory for verbal and nonverbal auditory stimuli. *Journal of Experimental Psychology*, 91(2), 341–343. doi:10.1037/h0031845.
- Rabin, M. D., & Cain, W. S. (1984). Odor recognition: Familiarity, identifiability, and encoding consistency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(2), 316–325. doi:10.1037/0278-7393.10.2.316.
- Richardson, J. T., & Zucco, G. M. (1989). Cognition and olfaction: a review. *Psychological Bulletin*, 105(3), 352–360. doi:10.1037/0033-2909.105.3.352.
- Rowe, E. J., Philipchalk, R. P., & Cake, L. J. (1974). Short-term memory for sounds and words. *Journal of Experimental Psychology*, 102(6), 1140–1142. doi:10.1037/h0036379.
- Shepard, R. N. (1967). Recognition Memory for Words, Sentences, and Pictures. *Journal of Verbal Learning & Verbal Behavior*, 6(1), 156–163. doi:10.1016/S0022-5371(67)80067-7.
- Weldon, M. S., & Coyote, K. C. (1996). Failure to find the picture superiority effect in implicit conceptual memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(3), 670–686. doi:10.1037/0278-7393.22.3.670.
- Weldon, M. S., & Roediger, H. L. (1987). Altering retrieval demands reverses the picture superiority effect. *Memory & Cognition*, 15(4), 269–280.
- Weldon, M. S., Roediger, H. L., & Challis, B. H. (1989). The properties of retrieval cues constrain the picture superiority effect. *Memory & Cognition*, 17(1), 95–105.
- Whitehouse, A. J. O., Maybery, M. T., & Durkin, K. (2006). The development of the picture-superiority effect. *British Journal of Developmental Psychology*, 24(4), 767–773. doi:10.1348/026151005X74153.
- Wippich, W., Melzer, A., & Mecklenbrauker, S. (1998). Picture or word superiority effects in implicit memory: Levels of processing, attention and retrieval constraints. *Swiss Journal of Psychology - Zeitschrift für Psychologie - Revue Suisse de Psychologie*, 57(1), 33–46.