Bulletin of the Psychonomic Society
1982, Vol. 19(2), 101-104

Birdsong learning and intersensory processing

RICHARD D. WALK
George Washington University, Washington, D.C. 20052

and

MICHAEL L. SCHWARTZ
Yale University School of Medicine, New Haven, Connecticut 06510

Two experiments were performed in which subjects learned to attach names to birdsongs. In the first experiment, subjects who were instructed to generate their own visual codes were much superior to those not given any instructions except those of learning the birdsongs. In the second experiment, both those given a model code for half of the birdsongs and those who made their own visual codes were superior to controls without visual codes. The experiments show the way in which learning in the auditory modality can benefit from visual symbols, and it has implications for the study of "higher order invariances" or relations between the modalities of vision and audition.

Eleanor Gibson, in her influential book, *Perceptual Learning and Development*, reports that W. H. Thorpe used visual spectrograms to help his students learn to identify birdsongs (Gibson, 1969, p. 211). This is an example of the use of a visual model or crutch to help to learn a complex auditory pattern. It has implications for intersensory processing, the way in which one modality can influence learning in another one. Birdsongs are readily available in our environment and on commercial phonograph records, yet few of us can identify with certainty any bird sounds except the most obvious, such as the caw of the crow or the cock-a-doodle-doo of the rooster. But these common bird sounds are not songs. Byron once characterized, with poetic insight, a birdsong as "the sweetest song ear ever heard."

The present series of experiments concerned the learning of birdsongs. The first experiment investigated the effect on such learning of having the subjects construct their own visual code of the birdsong as compared with those given instructions just to learn the songs. The second experiment had an additional group that had model visual patterns of half of the birdsongs.

Do the visual codes help learning? What are the theoretical implications of such auditory-visual studies?

GENERAL METHOD

Stimulus Materials

Birdsongs. A sample of 10 common birdsongs was selected from a record that contained a large number of birdsongs. The songs chosen were: Baltimore oriole, cardinal, common goldfinch, mockingbird, purple finch, robin, song sparrow, towhee, tufted titmouse, and winter wren.

Reprints may be obtained from Richard D. Walk, Department of Psychology, George Washington University, Washington, D.C. 20052.

The songs were placed in random order on a cassette tape. A total of six random orders, one for each phase, was constructed as follows: Phase 1 (pretest)–birdsongs followed by 8-sec pause; Phases 2-5 (training)–birdsongs identified at end of 8-sec pause; Phase 6 (posttest)–birdsongs followed by 8-sec pause. Each phase played all 10 birdsongs, with the trial identified for a phase so it could be matched with the answer sheet, for example: "*Trial one.* Birdsong [e.g., cardinal]. *Trial two.* Birdsong [e.g., winter wren]. . . ," and so on. With six phases and 10 trials for each phase, the experiment had a total of 60 trials. In the training section, as described, the name of the song came at the end of the 8 sec, whereas no feedback was given for pretest and posttest. The tape lasted about 25 min because of the taped instructions, the birdsongs, pauses for writing the answers, and 30-sec pauses between phases.[1]

Response sheets. The response sheets had a place for the subject to respond on each trial for each phase. The experiment required three sheets of paper to accommodate all 60 trials. Each sheet had the 10 birds listed in alphabetical order in the upper right corner to help identification. Subjects wrote their names on the response sheets.

Procedure

All subjects were told that they would hear a tape with birdsongs on it and that they would be asked to match the songs they heard with the names of the birds listed on the response sheet. The taped instructions reiterated this. On the pretest, the taped instructions said that no identification would be given of a song, but it asked that the subjects guess on every trial. The training phases (Phases 2-5) told the subjects to guess before the song was identified at the end of 8 sec. The posttest instructions told the subjects the birdsongs would not be identified and asked them to answer every trial.

EXPERIMENT 1

In this experiment, some of the subjects were asked to construct their own visual code to help identify the birdsongs.

Method

Subjects. The subjects were 87 students in a laboratory

0090-5054/82/020101-04$00.65/0

course in psychology at George Washington University. They were from five different laboratory sections.

Procedure. Two laboratory sections were given no additional instructions. The other three sections were told that a way to assist learning would be to take notes on the characteristics of each birdsong. For example, high and low marks might illustrate high and low notes in the songs.

Analysis. Not all subjects requested to use notes actually did so. Those requested to use notes, but who used none at all, compose a second "no-code" group, referred to in the results as "No-Code 2," meaning a second group of subjects without a visual code.

Results

The results in terms of accuracy by group are shown in Table 1. An analysis of variance showed that the groups did not differ on the pretest; all were at the chance level of about 10% correct. On the posttest, however, all groups were more accurate. Pretest-posttest difference scores showed that both the first no-code group ($p < .01$) and the second no-code group ($p < .05$) had significantly improved scores and the visual code group was markedly improved ($p < .001$). An analysis of variance of posttest scores of the three groups found a significant group effect [$F(2,84) = 17.13$, $p < .01$], and the studentized range statistic demonstrated that the visual code group was significantly superior ($p < .01$) to the other two groups.

The performance of the three groups on each individual birdsong is shown in Table 2. The use of a visual code helped performance on all songs, except that the superiority of the visual code group was slight for the common goldfinch. The use of a visual code particularly helped in discriminating the songs of the mockingbird, robin, towhee, and tufted titmouse. The high performance of the visual code group on the Baltimore oriole, cardinal, towhee, and winter wren should also be noted. Of course, contextual factors, such as the place of a song on the list and the confusability of competing songs, is a factor in the task.

To conclude Experiment 1, the use of a visual code markedly improved performance on this auditory task. But the visual code must be used. Those given instructions to use the code who ignored the instructions performed no better than the control group that did not receive such instructions. Further discussion of these results is deferred until after Experiment 2 is presented.

Table 1
Average Percent Correct Identifications for the Three Groups in Experiment 1

	Group		
Time	No-Code 1	No-Code 2	Visual Code
Pretest	9	11	10
Posttest	24	25	53
N	38	17	32

Table 2
Average Percent Correct Identifications on the Posttest for Each Birdsong, by Group, in Experiment 1

	Group		
Birdsong	No-Code 1	No-Code 2	Visual Code
Baltimore Oriole	37	29	63
Cardinal	47	47	81
Common Goldfinch	18	6	25
Mockingbird	11	12	47
Purple Finch	13	12	34
Robin	24	0	47
Song Sparrow	24	47	56
Towhee	21	35	72
Tufted Titmouse	11	24	44
Winter Wren	32	35	66

EXPERIMENT 2

Method

Subjects. The main subjects were 37 students from three laboratory sections. Detailed records were lost for an additional 34 subjects from two laboratory sections.

Procedure. Two of the laboratory sections had the same procedure as Experiment 1. The third laboratory section had a model code pattern made for five of the birdsongs (see Figure 1). These birdsongs were chosen to have equal performance on Experiment 1 by those who used a visual code. The songs chosen, mockingbird, robin, song sparrow, towhee, and tufted titmouse, were correctly identified 52.8% of the time on the posttest by the visual code group of Experiment 1. The other birdsongs had 53.4% correct identifications on the posttest by the visual code group.

The modeled birdsong group was given a dittoed sheet with models of the songs. These subjects were not allowed to turn the sheet over until after the pretest phase of the experiment.

Analysis. Because the number of subjects in each of the experimental groups was small, a capsule summary of the results from the two laboratory sections for whom statistical analysis could not be performed follows the analyzed results. One laboratory section ($n = 18$) was a modeled code group; the other ($n = 16$) was a group that constructed its own visual codes. All subjects requested to use a visual code actually did so.

Results

The main results are the three sections for whom complete data were available. The other two sections will serve to fill in and help give perspective for the experiment.

The results are shown in Table 3. They are very comparable to those of the first experiment. Pretest scores were around the chance level of 10%; the no-code group had 24% on the posttest and the visual code group had over 50%. The subjects furnished model patterns of the birdsongs were similar in their performance to those who made their own visual code.

Statistical analysis confirmed this. The groups did not differ on the pretest. A posttest showed a significant difference [$F(2,34) = 8.61$, $p < .01$] among groups, and the two code groups were significantly superior on the studentized range statistic ($p < .01$) to the group not instructed to use a visual code.

Two further questions are relevant to this analysis.

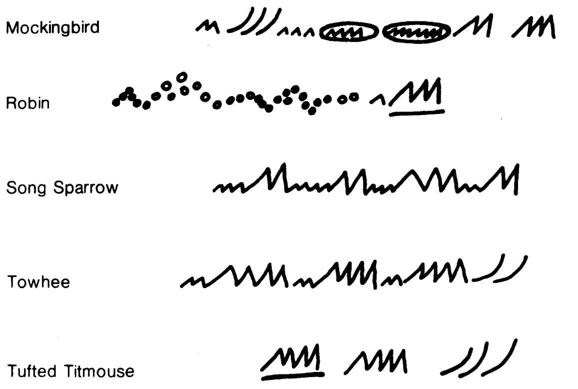

Figure 1. The model or pattern furnished subjects in the model code group of Experiment 2.

Table 3
Average Percent Correct Identifications for Each Group in Experiment 2, Showing Posttest Performance on Modeled and Nonmodeled Songs

Group	N	Pretest	Posttest Over-all	Posttest Model Songs*	Posttest No Model
No Code	14	10	24	21	26
Own Code	11	7	59	55	64
Model Code	12	11	50	57	43

**Only the "model code" group actually had a model or pattern of the five birdsongs.*

Are those with a pattern model superior on the modeled birdsongs to those who construct their own visual code? Does the use of a patterned model for half of the songs generalize to improve performance on the other songs?

This is also shown in Table 3. Both the modeled group and the own-code group were superior to the no-code group on the five modeled songs [$F(2,34) = 7.13$, $p < .01$]. The two code groups were significantly different from the no-code group but not from each other. The answer to the first question is negative. The second question is more complex. On songs without a model, the analysis of variance among groups was again significant [$F(2,34) = 6.05$, $p < .01$], but the significant difference in the means was for the own-code group compared with the no-code group, and the difference among means almost showed the own-code group (64%) significantly superior to the patterned model group (43%). The modeled song group seemed somewhat superior on nonmodeled songs to the no-code group (43% vs. 26% accuracy), but the difference was not significant and the trend may have been due to an artifact. The artifact is that the song models raise the chance level on the nonmodeled songs. Thus, to answer the second question, we have no evidence that the use of models generalizes (perhaps by increasing attention to identifying features) to improve performance on songs without a model.

Results from those for whom statistical analysis could not be performed were similar. The own-code subjects had 9% correct on the pretest and 47% on the posttest, with 40% correct on the other group's modeled songs and 54% on the nonmodeled songs. The modeled group had 11% correct on the pretest and 46% on the posttest, with 54% accuracy on the songs with a model and 37% on the songs without one. Compared with the controls, with 26% accuracy, the 37% correct on the songs without a model for the modeled group would seem to show that the modeled experience does not generalize to help performance on the nonmodeled songs. The slight superiority could be an artifact.

Thus, those for whom no analysis could be performed reinforced the conclusions from the other subjects. The modeled songs are a definite aid to learning, but the use of the models does not seem to generalize beyond them.

DISCUSSION

These experiments show that visual codes, whether generated by the subjects themselves or from models provided by the experimenters, help subjects to learn auditory songs. The results seem to be fairly specific: Those subjects instructed to use a code who used none did not benefit, and learning of the nonmodeled songs in Experiment 2 was not helped by the models provided for some of the songs. Explanation of the effect in terms of the rehearsal of, participation in, or attention to the task provided by the suggestion of or presence of the codes does not seem to be appropriate because of the specificity. We seem to have specific intersensory facilitation based on vision.

The auditory modality seems to have a certain plasticity that is influenced by a visual anchor. The common "ventriloquism effect" is one in which we localize a sound at the place where we see lips move on the ventriloquist's dummy. Being able to see the environment, even though no sound source is visible, helps auditory localization (Warren, 1970). An experiment by McGurk and MacDonald (1976) showed that auditory comprehension is influenced by the sighted lips of a model. If a tape is sounding "ba-ba" while the model's lips are forming "ga-ga," comprehension is affected, and the subjects hear some compromise, such as "da-da." This effect is based on experience, since children are less influenced by it.

Birdsongs are hard to describe in words, but complex visual patterns can help in their identification. Neither the subjects' own codes nor the models shown in Figure 1 are obvious sound spectrograms, but they do contain intermodal elements that aid learning. Some of the self-generated visual codes from Experiment 1 are reproduced in Walk (1978). They have few obvious intersubject similarities, and yet they did aid learning.

The broader theoretical question concerns high-level similarities and differences in the coding of different modalities. One thinks of Gibson's (1969) concept of "higher order invariances" or similarities, such as sharpness or smoothness, that interconnect modalities. Even 6- to 8-month-old infants can identify as similar or different patterns in the visual and auditory modalities (Allen, Walker, Symonds, & Marcell, 1977).

Unfortunately, visual symbols for coding language and music often seem more arbitrary than helpful, the result of cultural heritage rather than higher order invariances. One can see more audiovisual relationships in the vowels and in pitch on the musical scale, but these are a minor part of linguistic and musical symbols. In English, we have "cough" and "though" and "knife," "nice," and "kit" as a sample of words for which even the arbitrary letter codes are not related to pronunciation in a constant manner.

Birdsongs have no obvious matching codes like those the culture has provided for reading and music. One could teach oneself musical notation, of course. Sound spectrograms may help, but our subjects did not need them.

We need more research on how the modalities are related to each other. Where should the present research lead? Is it best to ask what "distinctive features" best interconnect the two modalities and how we can characterize the intersensory principles of perceptual organization? Or will any symbol help? We doubt that. If we named the birdsongs (the "squeaky one," "the busy one," etc.), we might get just as much facilitation as we do with visual models, but we could not randomly connect the names (or the models) to the songs.

REFERENCES

Allen, T. W., Walker, K., Symonds, L., & Marcell, M. Intrasensory and intersensory perception of temporal sequences during infancy. *Developmental Psychology,* 1977, 13, 225-229.

Gibson, E. J. *Principles of perceptual learning and development.* New York: Appleton-Century-Crofts, 1969.

McGurk, H., & MacDonald, J. Hearing lips and seeing voices. *Nature,* 1976, 264, 746-748.

Walk, R. D. Perceptual learning. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 9). *Perceptual processing.* New York: Academic Press, 1978.

Warren, D. H. Intermodality interaction in spatial localization. *Cognitive Psychology,* 1970, 1, 114-133.

NOTE

1. Birdsongs were taken from the phonograph record "A field guide to bird songs of Eastern and Central North America" (revised) (HM-4671). Boston: Houghton Mifflin, 1971. A copy of the tape used in this experiment can be obtained from the first author at cost.

(Received for publication December 2, 1981.)