over trial blocks but increased sharply in Group V-V.

Percent Overlap of Routes Within Trials

Each S's two trips, constituting each trial, were scored for the extent (expressed as a percentage) to which the second trip overlapped the first, regardless of the complexity of route taken. Thus, S's percentage score on any particular trial was 1.00 if the two trips comprising that trial completely overlapped, whether or not they were direct routes. To measure overlap, the total length in inches of overlapping section was divided by the total length of the two trips. The mean percent overlap of routes per trial was .67 for younger Ss, .35 for older Ss ($F = 5.61$, $df = 1, 24$, $p < .05$).

Correlation Between CA and Variability

Correlations between each of the four measures of variability and CA (expressed in days of age) were: (1) number of different routes, $r = .34$, $.10 > p > .05$; (2) frequency of choice of direct routes, $r = -.55$, $p < .05$; (3) number of turns, $r = .69$, $p < .01$; (4) percent overlap of routes within trials, $r = -.48$, $p < .01$.

These correlations therefore lend additional support to the age findings in the various analyses of variance.

DISCUSSION

The results indicate high relation among the various measures of performance on the maze-tracing task. Together or separately, they indicate first that the 4½- to 6-year-olds were substantially more variable than were the 2½- to 4½-year-olds (as measured by total number of different routes, choice of routes without turns, number of turns, and percent overlap of routes within sets of "trips"); second, variability tended to increase, not decrease, with increasing stimulus variability (number of different routes, number of turns).

The suggestion of a direct relation between stimulus variability and response variability disagrees with results of Ellis & Arnoult's (1965) study. The disagreement may stem from the slightly different methods employed: Ellis and Arnoult's Ss traced T-mazes, while Ss in the current study traced much more open-ended mazes. But it would seem that the stimulus-variability/response-variability phenomenon, if a phenomenon of any importance, would endure across such differences in the testing situation. As yet, there is no apparent explanation of this discrepancy.

Because of this discrepancy, it now becomes important to repeat these tests in order to assess both the reliability of these findings and the

extent to which a child's variability on maze-tracing tasks reflects general variability. For this latter reason, there is a need to obtain and to intercorrelate variability scores on a number of very different game-like tasks. Eventually, performance on such a test battery could be compared with performance in problem-solving tasks where variability plays an important role. Such data might prove useful in the following way: For example, the different response styles discovered in numerous studies of "probability learning" in children (e.g., Weir, 1964) have been presumed to be peculiar to such tasks, that is, to be "strategies" brought into play specifically for the purpose of solving some problem. But what seem like problem-solving strategies in probability-learning (or, more generally, in all discrimination-learning) situations might really be more general, characteristic, and enduring "response tendencies"—perseveration in the case of young children, variability in the case of older children. The discovery of

reliable differences in response variability across chronological age, as well as across other distinct S dimensions, could be helpful, then, in shaping new interpretations of such developmental learning phenomena.

REFERENCES

- BERLYNE, D. E. *Conflict, arousal, and curiosity*. New York: McGraw-Hill, 1960.
- ELLIS, N. C., & ARNOULT, M. D. Novelty as a determinant of spontaneous alternation in children. *Psychonomic Science*, 1965, 2, 163-164.
- HANLON, C. Response variability in children as a function of age, pretraining, and incentive conditions. *Journal of Comparative & Physiological Psychology*, 1960, 53, 267-269.
- HODGDEN, L. H. Variability of behavior as a function of intelligence and incentive. *Journal of Personality*, 1961, 29, 183-194.
- KOUNIN, J. S. Experimental studies of rigidity. I. The measurement of rigidity in normal and feeble-minded persons. *Character & Personality*, 1941, 9, 251-273.
- LEWIN, K. *A dynamic theory of personality*. New York: McGraw-Hill, 1936.
- WEIR, M. W. Developmental changes in problem-solving strategies. *Psychological Review*, 1964, 71, 473-490.

Recognition memory for common sounds*

JAMES D. MILLER and DAVID C. TANIS†
Central Institute for the Deaf, St. Louis, Mo. 63110

Recognition memory was measured separately for printed words, spoken words, and common sounds. Ss in each group inspected 350 stimuli and then were tested with 100 pairs of "old" and "new" stimuli in a two-alternative forced-choice task. Mean scores were 84%, 75%, and 69% correct for printed words, spoken words, and common sounds, respectively. It appears that recognition memory for common sounds is inferior to that for familiar pictures and no better than that for printed or spoken words.

In a recognition memory experiment, the S inspects for a few seconds each member of a series of stimuli that have been sampled from a particular population. After completion of the inspection series, the S is tested. During the test, he must recognize a stimulus as "old" (included in the inspection series) or as "new" (not previously presented). When there are many stimuli in the inspection and test series, delays between the inspection and the test series of a few minutes to several hours can be treated alike. Our discussion is limited to these conditions. Several

different test procedures have been used. These range from the recognition of singly presented stimuli to a two-alternative forced-choice task (2AFC) in which the test pairs are presented either simultaneously or successively. Since, in many of the previous experiments, the last procedure was used, we have converted, where necessary, the data reviewed below to the equivalent percent correct for the 2AFC procedure (Swets, 1964, Tables I and II).

It has been found that recognition memory for pictures of objects, scenes, and faces is excellent (Dallett, Wilcox, & D'Andrea, 1968; Goldstein & Chance, 1970; Hochberg & Galper, 1967; Nickerson, 1965, 1968; Shepard, 1967; and Standing, Conezio, & Haber, 1970). In each of these experiments, the 2AFC scores, either reported or computed, have mean

*This research was supported by Public Health Service Grant NS 03856 from the National Institute of Neurological Diseases and Stroke to Central Institute for the Deaf.
†Predoctoral fellow (PHS Grant GM 1900) in the Department of Psychology, Washington University, St. Louis, Mo.

values in excess of 90% correct; the average over all experiments is 94%. This is true for black-and-white or color pictures, for inspection series that range in length from 14 to 2,500 items, for similarly posed female faces, and even for old and new items that differ by only a single element in the scene. [Lower scores are obtained when the Ss are not instructed to remember the stimuli in the inspection series (82%), when the pictures are blurred or inverted during both the inspection and test series (70%-86%), when changes in focus or vertical orientation are interposed between the inspection and test series (65%-81%), or when the pictures are of unfamiliar patterns (inkblots, 83%; snowflakes, 71%).]

The overall excellence of recognition memory that has been shown for pictures does not appear to be a general characteristic of human recognition memory, since scores of about 87% correct are found when the stimuli are printed words or sentences (Egan, 1958; Shepard, 1967).

The remarkable recognition memory for pictures of familiar material led us to wonder whether similar results would be obtained if the stimuli were common sounds. For comparison, spoken and printed words were also used as stimuli.

SUBJECTS

The Ss were 27 students (24 female, 3 male) enrolled in high school or college.

MATERIALS

Three groups of stimulus materials, each consisting of an inspection series of 350 stimuli and a test series of 100 pairs of stimuli, were prepared. Two groups of stimulus materials (printed and spoken words) were prepared by random selection from published phonetically balanced lists of monosyllabic words (Egan, 1948). The *printed words* were typed on 3 x 5 in. index cards to form the inspection series, but for the test series a pair of words, one "old" and one "new," was typed on each card. The same vocabulary was used to prepare the *spoken words*. In this case, each word was spoken by one of the authors (DCT) and tape-recorded on separate Language Master Cards (Bell & Howell, 9-in. card).

The *sounds* were obtained primarily from sound-effects records and were chosen to be familiar (although not necessarily identifiable out of context) and discriminable. Each sound was recorded on a standard Language Master Card. The durations of the sounds ranged from about 0.5 to 4.8 sec. The categorization of these sounds and an example from each category are: animal (20%)—rooster crowing; industrial (18%)—pile driver;

nonspeech human (12%)—baby crying; musical (10%)—banjo melody; travel (10%)—bus horn; domestic (10%)—squeaky door; war (7%)—helicopter; miscellaneous (7%)—shuffling cards; artificial (5%)—300-Hz tone; and weather (1%)—thunder. The 350 sounds used for the inspection series were randomly selected from a set of 450 sounds. The test series was formed by pairing the 100 remaining cards with 100 cards chosen randomly from the inspection series.

PROCEDURE

Nine of the Ss were tested with printed words, nine others with spoken words, and nine others with common sounds. When the stimuli were printed words, the Ss were tested individually; otherwise they were tested in groups. All Ss were instructed in detail with regard to the procedures, were urged to "try to learn and remember the inspection stimuli," and were informed of the nature of the test to follow. Then the 350 inspection stimuli were presented and, after a 5-min rest, the 100 test pairs were given. The members of a test pair were presented either simultaneously (printed words) or in rapid succession (spoken words or common sounds). The spoken words and common sounds were presented at a comfortable listening level via a Language Master that had been modified for use with an external loudspeaker. During the inspection series, the common sounds were presented at a rate of about one every 8 sec, while the spoken words were presented at a rate of one every 6 sec. For printed words, each S looked at each card for about 3 sec, as instructed.

RESULTS AND DISCUSSION

The Ss tested with printed words made correct decisions on 69%-100% of the test trials, with a mean of 84.4%. Those tested with spoken words had scores between 58%-93%, with a mean of 74.8%. Those tested with common sounds had scores between 63%-75%, with a mean of 69.2%. A one-way analysis of variance was significant ($p < .01$), but pairwise comparisons only yielded a significant difference ($p < .01$) between the scores of those tested with printed words and common sounds.

The scores for printed words of about 84% correct agree closely with those of Egan (1958), who used words from the same population. The fact that the scores for spoken words were lower than those for printed words cannot be interpreted because of differences in procedure, the fidelity of the recordings, and the variance of the scores.

It is clear that the recognition memory scores for common sounds (69%) are not nearly as high as those for familiar pictures (94%). One factor limiting the scores for common sounds was probably the low fidelity of the recording and playback system. However, it is unlikely that the quality of this system accounts for all of the difference between scores for sounds and those for pictures, since scores of 86% correct are found for highly similar and badly blurred pictures (Dallett et al, 1968). Rather, it is likely that auditory stimuli must be of rather long duration or extremely familiar before they can be differentiated and stored in memory as well as pictures of familiar objects and scenes. This speculation is given some support by the data for talker identification (Bolt et al, 1970) which show that the number of errors made in identifying voices of a set of 10-15 talkers decreases rapidly as the duration of the speech sample is increased to 2 sec and that if leisurely listening is allowed Ss can identify without error as many as 65 talkers (Pollack, Pickett, & Sumby, 1954).

REFERENCES

- DALLETT, K., WILCOX, S. G., & D'ANDREA, L. Picture memory experiments. *Journal of Experimental Psychology*, 1968, 76, 312-320.
- BOLT, R. H., COOPER, F. S., DAVID, E. E., JR., DENES, P. B., PICKETT, J. M., & STEVENS, K. N. Speaker identification by speech spectrograms: A scientist's view of its reliability for legal purposes. *Journal of the Acoustical Society of America*, 1970, 47, 597-617.
- EGAN, J. P. Articulation testing methods. *Laryngoscope*, 1948, 58, 955-991.
- EGAN, J. P. Recognition memory and the operating characteristic. Indiana University, Hearing and Communication Laboratory, Technical Note AFCRC-TN-58-51, 1958.
- GOLDSTEIN, A. G., & CHANCE, J. E. Visual recognition memory for complex configurations. *Perception & Psychophysics*, 1970, 9, 237-241.
- HOCHBERG, J., & GALPER, R. E. Recognition of faces. I. An exploratory study. *Psychonomic Science*, 1967, 9, 619-620.
- NICKERSON, R. S. A note on long-term recognition memory for pictorial material. *Psychonomic Science*, 1968, 11, 58.
- NICKERSON, R. S. Short-term memory for complex meaningful visual configurations: A demonstration of capacity. *Canadian Journal of Psychology*, 1965, 19, 155-160.
- POLLACK, I., PICKETT, J. M., & SUMBY, W. H. On the identification of speakers by voice. *Journal of the Acoustical Society of America*, 1954, 26, 403-406.
- SHEPARD, R. Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning & Verbal Behavior*, 1967, 6, 156-163.
- STANDING, L., CONEZIO, J., & HABER, R. N. Perception and memory for pictures: Single-trial learning of 2500 visual stimuli. *Psychonomic Science*, 1970, 19, 73-74.
- SWETS, J. A. (Ed.) *Signal detection and recognition by human observers*. New York: Wiley, 1964.