Variability in maze drawings of young children: Effects of stimulus change and chronological age*

LAUREN HARRIS
Michigan State University, East Lansing, Mich. 48823

Thirty children between 2½ and 6 years of age traced a series of mazes under three different conditions of stimulus (crayon color) constancy. On the basis of several different measures, older children proved to be substantially more variable than younger children and, for all children, variability tended to increase with increasing stimulus variability.

In a recent study by Ellis & Arnoult (1965), nursery school children traced a series of T-mazes under varying and nonvarying stimulus conditions. The stimulus conditions were established by either varying or holding constant across trials the color of the outline of the T-mazes, the color of the background of the cards on which the mazes were drawn, or the color of the crayon with which the child traced the maze. Based primarily on Berlyne's (1960) hypothesis on the activating properties of novel stimuli, Ellis and Arnoult predicted that alternation of route traced would exceed chance level when stimulus conditions remained constant on successive trials, would not exceed chance when stimulus conditions varied on successive trials, and, most generally, would be less under conditions of stimulus variability than under conditions of stimulus constancy. Their results supported all three predictions, leading the authors to conclude that spontaneous alternation appeared to be related to the activating properties of novel stimuli.

If we take spontaneous alternation behavior as an instance of variability behavior, we recognize this study as an exemplar of a totally situation-based approach to the investigation of variability. The child is variable only insofar as external cues impel him to be variable. The design, employing a presumably homogeneous sample of children (in Ellis and Arnoult's case, 4- to 5-year-olds), does not permit hypothesis-based testing of individual differences in variability.

A second approach is exemplified in a study by Hodgden (1961). Hodgden also was concerned with investigating the amount of response variability shown by a child given freedom either to make a new response or to repeat a previous response. Hodgden, however, was influenced by the early writings of Lewin (1936) and Kounin (1941), who, on the basis of their studies of behavioral rigidity, hypothesized that response variability increased with increasing intelligence level. Hodgden, then, predicted that certain characteristics of the S would interact with the external stimulus conditions to produce response variability. Her results, from child Ss at three levels of intelligence (retarded, average, bright) bore out Lewin's and Kounin's hypothesis.

For developmental psychologists, these two studies, while complete in their respective ways, do not allow for examination of the effects of chronological age within children of normal intelligence. Hanlon (1960) met this need. Using Hodgden's apparatus and substantially similar procedures, she found that 6- to 7-year-old children were significantly more variable than were 3- to 4-year-old children.

The current research attempted to mesh these two approaches by considering both chronological age and external stimulus conditions as variables contributing to variability. The experimental materials consisted of a series of complex mazes which S traced with crayons. The questions were: whether or not older children are more variable than younger children, whether or not response variability increases as stimulus constancy increases, and whether or not older and younger children are affected differently by various levels of stimulus constancy.

SUBJECTS

The Ss were 30 children divided into two age groups ($\bar{X}$ CA "old" group = 5 years 3 months, 12 days, range = 4:7:9-6:1:14; $\bar{X}$ CA "young" group = 3:9:23, range = 2:6:5-4:7:2). The young group contained seven boys and eight girls, the older group, eight boys and seven girls. All Ss attended a day-care center in Lansing, Michigan.

APPARATUS

Each S drew with crayons on 20 identical sheets of white paper, 8½ x 14 in. in size. Each sheet was overprinted with cross-hatched sections representing a simple maze (see Fig. 1). The crayons were red, green, brown, blue, and yellow.

PROCEDURE

The Ss were tested individually in an extra room in the day-care center. The S sat across from E at a table masked on E's side to hide the crayons and papers not in use. The E presented the first sheet to S and spoke the following instructions from memory: "Look at this [pointing to 'house']. Let's pretend that this is your house [pointing with finger of one hand] and this is the grocery store [pointing with finger of other hand].

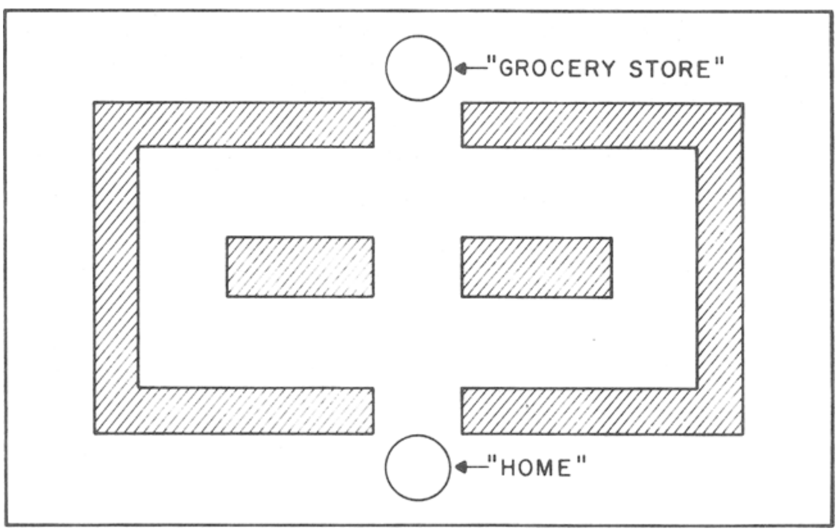

Fig. 1. Maze printed on 8½ x 14 in. white paper.

*This research was supported in part by a faculty grant from Michigan State University to the author. Thanks are due Alice Granskog for her enthusiastic and skillful help in testing the children and tabulating the scores and also the staff of the Carol Lee Nursery, Lansing, Michigan, for their cooperation.

"You want to go to the grocery store to buy some food. You pretend that this crayon [indicating crayon] is your car, and you drive your car from your house [pointing] to the grocery store [pointing] by drawing with your crayon.

"The streets that you can drive on are the white spaces. You can drive on any of the streets that you like. You can go to the grocery store any way you want.

"[Trial 1] Here's your car [giving crayon to S]. You start from home [E puts S's hand and crayon on 'home'] and drive to the store [pointing with other hand] to buy some food. [After S reaches 'store,' E takes the crayon from him and removes it from sight.] Now, here's your car [giving appropriate crayon to S], drive back home with the food [pointing to 'home']. [If S hesitates at least 5 sec, E points again to 'home' and says, 'Drive your car home.']

"[After S reaches 'home,' E takes the crayon away and, out of sight, removes the used page and turns it over onto the used-page pile.] Now you've eaten all the food, so you have to go to the store to buy some more. Here's your car [E gives S the appropriate crayon]. Drive to the store."

The first sheet, comprising two trips (to the store and from the store), was Trial 1. Each S made 40 trips (20 trials).

STIMULUS CONDITIONS

There were three conditions designed to represent three levels of stimulus variability. In constant-constant (C-C), S received the same color crayon both "going to the store" and "coming back" and the same color for each trial; in constant-variable (C-V), the color was the same for going to and from the store but varied from one set of trips to another (i.e., from trial to trial); and in variable-variable (V-V), S received one color for "going to the store," another color for "coming back," and still a different color for the first trip of the next trial. For Ss in Conditions C-V and V-V, sequences of 20 trials comprising five colors, each color appearing four times, were generated as follows: (1) Four different series of the five colors were generated such that, among the series, no color followed another color or itself twice; (2) the four series then were treated in the same way such that no series followed another series twice, making four different orders of the four series; and (3) the four orders were assigned equally among Ss according to age (young and old) and sex.

The 30 Ss were divided approximately equally by age and sex among the three stimulus conditions. Each S was tested in only one condition. In each condition, S returned the crayon to E after completing each trip.

RESULTS

Each S's drawings were scored for: (1) number of different routes; (2) frequency of choice of direct routes (routes without turns); (3) number of turns; (4) percent overlap of routes within trials. Each set of scores was analyzed in an Age by Sex by Stimulus Condition by Trials analysis of variance. For brevity's sake, only significant ($p < .05$) main effects and interactions will be discussed, though marginally significant effects will be reported where of special interest.

Number of Different Routes

The Ss received 1 point for each route taken that differed from the first route. Ss who "drove" to or from the store in five different ways received 4 points; Ss who took the same route throughout the 20 trials received zero points.

For the analysis of variance, scores were grouped into five blocks of four trials each. Table 1 presents the mean number of different routes across trial blocks traced by Ss in the two age groups. The age effect was significant: older Ss took more different routes than did younger Ss ($F = 4.77$, $df = 1,24$, $p < .05$). Additional significant effects were trial blocks ($F = 9.96$, $df = 4,96$, $p < .001$) and the interaction of Trial Blocks by Age ($F = 4.72$, $df = 4,96$, $p < .005$). As Table 1 shows, the average number of different routes for both young and old Ss was greater in Trials 1-4 than in later trials, and this difference was greater among the older Ss. These latter effects are difficult to interpret. Each S received 1 point for each route that differed from the first route. Assuming that each S had available to him a conceptually finite population of different routes, sampling without replacement from this population would make it be more and more difficult with each completed trial to trace still another different route. The trial blocks effect, therefore, is probably an artifact of the measure of variability and as such would not warrant the conclusion that Ss' response variability (meaning the psychological characteristic underlying the score) decreased over trials.

Table 1
Mean Number of Different Routes Per Trial (Two Trips) by Young and Old Ss

	Trials				
	1-4	5-8	9-12	13-16	17-20
Young Ss N = 15	.14	.10	.50	.50	.02
Old Ss N = 15	.40	.02	.09	.10	.09

The same point applies to the interaction. The largest difference appears in Trials 1-4 where older Ss traced three times as many different routes as did younger Ss. Thereafter, the two groups were substantially more similar, the older Ss losing their lead in Trial Block 2 and regaining it by a small margin in Trial Blocks 3, 4, and 5. Though the threefold difference in Block 1 clearly contributed heavily to the interaction, again we could not conclude that the older and younger Ss were substantially different in response variability *only* at an early point in the task. The older Ss' higher scores in Block 1 made it relatively more difficult for them to trace additional different routes thereafter. It therefore appears that, in a maze-drawing test, *consistency* of variability as a psychological attribute is better estimated by other response measures.

As for stimulus condition, only borderline significance was reached ($F = 2.91$, $df = 2,24$, $.05 < p < .10$). The direction of difference, however, was opposite to that predicted: the mean number of different routes taken per trial increased rather than decreased with increasing stimulus variability (for Group C-C, $\overline{X} = .12$; Group C-V, $\overline{X} = .20$, Group V-V, $\overline{X} = .30$).

Frequency of Choice of Direct Routes

The mean number of direct routes (trips without turns) per trial for older Ss was 0.78, for younger Ss, 1.24. The difference was of borderline significance ($F = 3.32$, $df = 1,24$, $.05 < p < .10$).

Number of Turns

A turn was any 90-deg change in direction, whether or not it brought S onto a new "street." For the analysis of variance, the scores were grouped into four blocks of five trials each. The analysis revealed that over all trial blocks, older Ss made more turns per block ($\overline{X} = 1.91$, $F = 10.48$, $df = 1,24$, $p < .01$). The Trial Blocks by Condition effect was also significant ($F = 2.37$, $df = 6,72$, $p < .05$). The mean numbers of turns in Groups C-C and C-V were approximately constant 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 feebleminded 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.