Developmental Psychology
1975, Vol. 11, No. 3, 304-310

Effects of IQ and Mental Age on Hypothesis Behavior in Normal and Retarded Children

JOHN R. WEISZ AND THOMAS M. ACHENBACH
Yale University

Mental-age-matched normal and retarded subjects at mental-age (MA) levels 5.5, 7.5, and 9.5 years received blank-trial discrimination problems designed to expose hypothesis behavior. There was evidence that subjects at all MA levels used hypotheses. Use of feedback indicating that a response was wrong increased significantly with MA, while use of feedback indicating that a response was right increased significantly with IQ. On simple problems involving two stimulus dimensions, retarded and normal groups used about equal numbers of hypotheses, but on four-dimensional problems the retarded used fewer hypotheses than normals. This IQ Level × Problem Complexity interaction may explain contradictions among previous findings regarding IQ effects on learning.

A major issue in developmental psychology concerns the respective roles of rate of learning and level of cognitive development in determining intellectual performance. This issue has been especially central in research with familially retarded children. Views of cultural-familial retardation tend to cluster around two contrasting poles which Zigler (1969) has labeled *developmental* and *difference* theories. Difference theories ascribe specific intellectual deficiencies to the familially retarded (e.g., deficient attention and retention processes, cf. Zeaman, 1973; Zeaman & House, 1967), while Zigler's developmental view is that the cognitive competence of familially retarded persons should be essentially similar to that of mental-age-matched normals.

This study was conducted while the first author was a National Institute of Mental Health Predoctoral Research Fellow (1 F01 MH54552-01). It was supported by National Institute of Child Health and Human Development Grant HD03008 to the second author. The authors are grateful to David Audette, John Mazzarella, Ruben Swimmer and the staffs of the Middletown and Milford, Connecticut, school systems for their generous assistance in obtaining subjects, and to David Balla, David Feldman, Ann Garvin, Susan Harter, Fred Rothbaum, and Edward Zigler for their help in other aspects of the study.

Requests for reprints should be addressed to John R. Weisz, Department of Psychology, Yale University, New Haven, Connecticut 06520.

The degree to which the mental age (MA) yielded by traditional IQ tests is an adequate index of cognitive competence is in itself open to question, but in recent studies it has been shown that normal and familially retarded subjects matched for total Binet MA (both Short-Form and Standard) are similar in the patterns of Binet items they pass (Achenbach, 1970, 1971) and in success on minimally verbal Piagetian tasks (Achenbach, 1969, 1973; Brown, 1973). Nevertheless comparisons of MA-matched retarded and normal groups on measures of learning rate, such as trials to criterion in discrimination learning tasks, sum to a highly equivocal body of evidence. In several such studies more rapid learning by normals than by the retarded is reported, but in a number of others equivalent performance by the two groups is reported (cf. Zeaman & House, 1967).

In traditional discrimination learning tasks, the seeming simplicity of trials to criterion as an index of ability masks the fact that the attainment of criterion is merely the end product of a complex succession of behaviors. The tendency to focus on such end products alone is unfortunate for the developmental–difference debate, for at the heart of this controversy is the question of whether retarded–normal differences exist in the underlying cognitive processes. The primary purpose of the present study was to compare normal and familially retarded groups with respect to a particularly important process in virtually all learning, that is, hypothesis behavior. To expose the specific hypotheses used by each subject, a "blank-trials" discrimination learning procedure was employed whereby subjects received no feedback as to the correctness of their responses on selected blocks of trials (cf. Levine, 1966). Hypotheses were inferred from subjects' response patterns across blank-trial blocks, and the percentage of such blocks reflecting hypotheses constituted a subject's score for frequency of hypothesizing. Changes in hypotheses after feedback were also scored in order to determine how effectively hypotheses were used.

A second purpose of the study was to explore one possible explanation for the inconsistent findings on IQ effects in learning. Harter (1965) and Weir (1967) have argued that since an IQ reflects a subject's lifetime learning rate, it should also predict his rate of learning in an experimental setting, provided that the experimental task is relatively complex. Accordingly, Harter and Weir have proposed that in studies in which MA-matched retarded and nonretarded subjects are compared, complex tasks are more likely than simple tasks to reflect the inferior learning ability ascribed to the retarded. In order to test this proposal with respect to hypothesis behavior, half of the subjects in the present study received relatively simple discrimination problems, while the other half received more complex problems. Since the "simplicity" of a task depends partly upon the MA of the person attempting it, it was deemed important to compare normal and retarded subjects across a broad range of MA levels. The lowest MA level at which pilot subjects clearly comprehended the task was approximately 4.5 years. The MA ceiling was determined by the upper MA limit of retarded pupils in public schools. An intermediate MA level was included to provide for a "mapping" of developmental functions.

Method

Experimental Design

The 2 × 3 × 2 factorial design employed IQ levels of approximately 68 and 100, crossed with three levels of MA. The third factor, problem type, included discrimination problems involving two stimulus dimensions (2D problems) and four stimulus dimensions (4D problems). Stanford-Binet Short-Form MAs and IQs were obtained using an "optimizing" testing procedure whereby easy items were alternated with difficult items (cf. Zigler & Butterfield, 1968) 3 to 5 weeks prior to the experiment. Approximately half of the subjects in each cell were administered the Binet and experimental task by a male experimenter and half by a female experimenter.

Subjects

The 156 subjects attended regular and special classes in urban public schools. IQs ranged from 90 to 125 for normals and 48 to 81 for the familially retarded. Mean Hollingshead (Note 1) social class scores were 6.0 for retarded and 5.4 for normal subjects (7 = the lowest level; retarded–normal difference nonsignificant, $p > .15$).

Thirty-nine quartets were formed, with 2 normal and 2 retarded subjects in each, and with members matched for sex, race, MA, and experimenter. Within each quartet, 1 retarded and 1 normal subject received 2D problems, the others 4D problems. There were 13 quartets each at mean MA levels of approximately 5.5 years (range: 53–78 months MA), 7.5 years (range: 80–99 months), and 9.5 years (range: 102–126 months). No differences in mean MA within an MA level approached significance among the four Type of Problem × IQ Level groups, nor did any IQ differences within normal or retarded groups. Over the entire sample, MA and IQ were unconfounded ($r = .06$). All subjects were white except for one black quartet at the 5.5-year level, two at the 7.5-year level, and one at the 9.5-year level. Each cell at the low-MA level contained 7 boys and 6 girls, while each cell at the upper MA levels contained 8 boys and 5 girls.

Blank-Trials Procedure

Sample problems. After asking the subject to help "try out some new things for kids to do," the experimenter placed a deck of nine white 7.5 × 15 cm Plexiglas cards face up on a table between himself and the subject. Two shapes appeared side by side on each card. Each shape was cut from Pantone, 5 × 5cm in size and colored black for 2D problems. The 2D stimuli differed in shape and position on the card (left or right), while 4D stimuli differed in shape, position, color, and size (6.5 × 6.5 cm versus 2.5 × 2.5 cm). The same two shapes appeared on all cards of a problem; 12 shapes (e.g., circle, square, star) and 12 colors were employed, each being used for two problems.

For Sample Problem 1, the subject was asked to name the two shapes, their respective "sides of the card" ("this side," subject's left; "that side," subject's right), and, for 4D subjects only, their colors and relative sizes (bigger and smaller). Then the experimenter said (portions adapted from Eimas, 1969):

> What we are going to do is look through these cards, one at a time. For each card, you will point to one of the two things, and I will tell you whether you are right or wrong. Then [turning to Card 2] we will try the next card; you will point to one of the two things, and I will tell you "right" or "wrong." In this way, you can begin to figure out whether it's one of the shapes, or one of the colors, or one of the sizes, or one of the sides of the card, that's right for the whole bunch of cards. [The experimenter then replaced Card 1 on top of the deck.] Now let's try this bunch of cards just for practice. Remember your job is to figure out whether it's one of the shapes, or colors, or sizes, or sides of the card that is always the right answer. When you figure that out, you'll be able to be right every time. Now let's begin. Point to one of these things.

Sample Problems 1 and 2 were then administered, with feedback provided on every trial. Before Sample Problem 3 the experimenter said, "So far I've said 'right' or 'wrong' after every card, haven't I? Well, from now on I won't always tell you whether you are right or wrong. After some cards I won't say anything. Don't let this bother you. You just keep trying to be right all the time." For Sample Problem 3, feedback occurred on every second trial, and for Sample Problem 4, on every third trial.

Sample problems were designed *only* to familiarize subjects with the general problem format and the partial feedback arrangement and to satisfy the assumption that experimenter and subject shared a common set of "legitimate" solution hypotheses. Thus, practice in actual problem solving was limited. On every sample problem, the subject was informed after three feedback trials: "The right answer for this bunch of cards is one of the [the relevant dimension, e.g., "sides of the card"]. Which [e.g., "side"] do you think it is that's right every time?" If the subject made the wrong choice, he was informed of the correct answer. The experimenter then said, "Now let's see if you can be right every time," and proceeded through enough cards to provide six feedback trials. The procedure insured that subjects were nearly always correct, but corrective hints were given when errors were made. (Practice problems, unlike the experimental problems, *did* have correct answers rather than prearranged feedback). For 2D subjects, the correct answers were one shape, position (right), a second shape, and position (left), respectively. For 4D subjects the correct answers were a particular size, color, shape, and position, respectively., At the end of each sample problem the experimenter reviewed with the subject the correct answer and the fact that no other stimulus dimension was relevant. After Sample Problem 4, the experimenter said:

> Okay, do you see how these things will be? They aren't just guessing games—they are things you have to figure out. You have to figure out how to be right every time. Just like with the things we've done, the right answer for each bunch of cards will always be one of the shapes or (for 4D subjects) one of the colors, or (for 4D subjects) one of the sizes, or one of the sides of the cards. The right answer will never be anything else.

Experimental problems. In order to maximize motivation, two clear plastic bowls were introduced, one empty, one containing 15 marbles. The subject was told:

> Those things we just did were for practice. Now we're going to have a contest. We're going to do the same thing we've been doing, but whenever I say "right" I'll give you a marble, and whenever I say "wrong" you give me back a marble. When we're all finished, if you have seven marbles or more you can pick out a prize from that big box [pointing]. I'll give you two to start with—now let's begin.

After each experimental problem the subject checked his progress toward the goal of 7 marbles, using a score sheet. This procedure was included because recent evidence (e.g., Harter & Zigler, 1972) indicates that reward-plus-penalty conditions enhance discrimination learning in the familially retarded, presumably by heightening their motivation.

At the beginning of every problem all stimulus dimensions were reviewed and all instances named; then the experimenter said: "Remember, your job is to figure out what the right answer is for this bunch of cards. It's the same answer that's right for every one of the cards in this bunch, but you have to figure out what it is—whether it's one of the two shapes that's right for every card, or one of the two. . . ."

Each of the eight 2D experimental problems consisted of 12 cards. The two possible stimulus arrangements for each problem appeared equally often in a randomly determined order, with the following restrictions: (a) that no one stimulus arrangement appear more than twice in succession; (b) that no one arrangement appear in successive feedback trials; (c) that no one arrangement appear on a feedback trial and on the immediately subsequent blank trial.

Each of the eight 4D experimental problems consisted of 20 cards on which the pairs of shapes followed the same order as in the 2D problem series. In each problem there were eight possible stimulus pairings in which the paired stimuli differed along all four dimensions. Four of these pairings were used only on feedback trials; the remaining four appeared in four different randomly determined series, forming the blank-trial blocks. Each set of four stimulus cards was internally orthogonal in the sense that both levels of every dimension appeared twice with both levels of every other dimension. The internal orthogonality of the blank-trial stimuli provided for a simple index of which hypothesis, if any, was being employed on a given blank-trial block. Each of the eight possible response patterns involving two or all four choices to one side was consistent with one of the legitimate solution hypotheses. None of the eight patterns involving one choice on one side and three on the other was consistent with a legitimate hypothesis.

For the eight experimental problems the experimenter said "right" or "wrong" (and dispensed or retrieved a marble) on Trials 4, 8, and 12 for 2D problems and on Trials 5, 10, 15, and 20 for 4D problems, without regard to the subject's response. No feedback was provided on any other trials. Each of the eight possible three-trial right-wrong sequences was used once in prearranging feedback for Trials 5, 10, and 15 of the 4D problems. Each of the four possible two-trial sequences using right and wrong was employed twice in determining feedback for Trials 4 and 8 of all 2D problems. In both cases right-wrong sequences were assigned to particular problems on a random basis, with one limitation: that subjects be told "right" on the majority of feedback trials for the final problem. Final trial feedback was "right" for six problems and blank (i.e., no feedback) for two, placement being determined with two goals in view: (a) to provide roughly parallel progress (including occasional regressions) toward the goal of 7 marbles for 2D and 4D subjects and (b) to assure that no more than three "wrongs" occurred in succession within a problem.

Results

Reliance on Hypotheses

For both 2D and 4D problems, response patterns consistent with one of the simple solutions outlined for subjects were designated as hypotheses. Because 2D subjects had 24 blank-trial blocks on which to manifest hypotheses, while 4D subjects had 32, the extent to which subjects manifested hypotheses was expressed as a percentage of the number of opportunities. In order to assess experimenter effects, all possible subject pairs were formed with members matched for race, sex, IQ level, problem type, and MA, but differing as to the experimenter they saw. To assess sex effects, subjects were matched on all relevant variables except sex. Neither sex nor experimenter effects approached significance with respect to any of the dependent variables.

Normal and retarded groups at all three MA levels used hypotheses at well above the 50% chance baseline, the lowest being 74%; in comparisons of each cell mean with chance expectation, $t(12) > 5.0$, $p < .001$. A 2 × 3 × 2 (IQ × MA × Problem Type) analysis of variance using arcsine-transformed percentage scores revealed that 2D problems evoked a significantly greater percentage of hypotheses than 4D problems, $F(1, 144) = 7.42$, $p < .01$. Underlying the main effect for problem type was a significant IQ × Problem Type interaction such that on 2D problems, retarded subjects used more hypotheses than normals (89% versus 85%), while on 4D problems normals used more than the retarded (86% versus 78%), $F(1, 144) = 4.03$, $p < .05$ for interaction of IQ × Problem Type. The difference between normal and retarded groups on 2D problems was not significant, while the difference on 4D problems was significant, $t(76) = 2.21$, $p < .05$. Viewed another way, normals were about equally reliant upon hypotheses for both the 2D and 4D problems, while retarded subjects relied less on hypotheses for 4D problems than for 2D problems, $t(76) = 2.69$, $p < .01$ for 2D versus 4D retarded. Neither MA nor IQ had significant main effects on hypothesis scores (see Figure 1).

Responsiveness to "Right" and "Wrong"

In order to determine how feedback information was used, 2 × 3 × 2 (IQ × MA × Problem Type) analyses of variance were performed on the arcsine-transformed percentage of (a) hypothesis patterns following the feedback "wrong" that were logically consistent with information provided by that "wrong," and (b) hypothesis patterns following the feedback "right" that were logically consistent with that "right." In these analyses, the only significant effects were that (a) use of the feedback "wrong" generally improved with MA in both IQ groups, $F(2, 144) = 6.94$, $p < .005$, and (b) normals were generally superior to the retarded in using the feedback "right," $F(1, 144) = 4.91$, $p < .05$.

Analyses of variance (2 × 3 × 2) were also performed on the arcsine-transformed percentage of times that (a) a hypothesis pattern followed by the feedback "wrong"

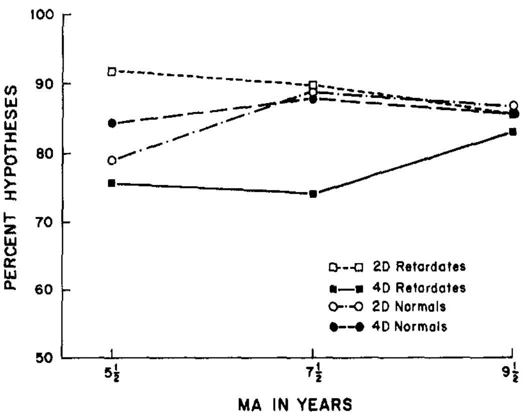

Figure 1. Percentage of blank trial blocks on which hypotheses were manifest. (Each point represents 13 subjects. 2D = 2 stimulus dimensions; 4D = 4 stimulus dimensions.)

led to a change to any other hypothesis pattern, and (b) a hypothesis pattern followed by the feedback "right" led to repetition of the same hypothesis pattern. In effect, these two variables reflect only the subject's tendency to repeat or change a particular hypothesis pattern following positive or negative feedback, while the variables described above reflect responding that was logically consistent with either positive or negative feedback. As is clearly shown in Table 1, changes in hypotheses following "wrong" increased significantly with MA, $F(2, 144) = 6.31, p < .005$. However, for repetition of a hypothesis following "right," an IQ × MA interaction occurred because normals at MAs 5.5 and 7.5 years repeated more often than the retarded, while the retarded repeated more at MA 9.5 years, $F(2, 144) = 3.50, p < .05$. Furthermore, there was an IQ × Problem Type interaction similar to the interaction effect on total number of hypotheses used: On 2D problems, the retarded repeated more hypotheses than normals (79.9% versus 75.1%), but on 4D problems the retarded repeated fewer hypotheses (61.7% versus 80.8%), $F(1, 142) = 4.90, p < .05$. Thus, differential responsiveness to the feedback "right" may be related to the different frequencies of using hypotheses by the normals and retarded on simple and complex problems.

Table 1: Mean Percentage of Hypothesis Patterns Repeated Following "Right" and Replaced Following "Wrong"

Mental age level (yrs.)	2D subjects Normal	2D subjects Retarded	4D subjects Normal	4D subjects Retarded
5.5				
Right	74	84	84	70
Wrong	27	32	32	28
7.5				
Right	85	68	78	45
Wrong	51	57	34	36
9.5				
Right	66	85	80	70
Wrong	58	53	48	53

Note. The figures after Wrong represent the mean percentage of times that a hypothesis pattern and the subsequent feedback "wrong" were followed by a different hypothesis pattern. 2D = 2 stimulus dimensions; 4D = 4 stimulus dimensions.

Response Patterns

The finding that subjects in the MA-5.5-years group used hypotheses differs somewhat from a finding by Gholson, Levine, and Philips (1972) that normal 5-year-olds failed to do so when responses based on position were disallowed as hypotheses. Several procedural differences between the present study and that of Gholson et al. make direct comparisons with the present data impossible, but data are now available from another study (Weisz, Note 2) in which six groups of 10 children each, all at mean MA 5.5 years, received discrimination problems similar to those of Gholson, et al., with position-oriented response patterns disallowed. In the Weisz study, with instructions and motivating conditions similar to those of the present investigation, each of the six groups responded with hypothesis patterns significantly more often than would be expected by chance, $p < .05$, in every instance.

In the Gholson et al. analysis of "nonhypothesis" patterns, the only position-oriented response pattern appearing more often than expected by chance was position alternation (e.g., left, right, left, right). On the 4D problems of the present study, every position alternation pattern corresponded to some legitimate hypothesis, while in 2D problems this was true only in 6 of 24 blank-trial blocks. This raises the question of whether position alternation responding may help to account for the IQ × Problem Type interaction reported above for number of hypotheses—a question which is answered by simply repeating the analysis of variance with all position-alternation patterns subtracted from the data. Hence, a 2 × 3 × 2 (IQ Level × MA Level × Problem Type) analysis of variance was calculated, using as a dependent variable the arcsine-transformed percentage of hypothesis patterns relative to the number of opportunities for a hypothesis, with all position alternation patterns subtracted. Just as with the original analysis, an IQ × Problem Type interaction revealed fewer hypothesis patterns among the retarded than among normals, but only on the complex 4D problems, $F(1, 144) = 4.12, p < .05$.

The preceding analysis indicates that position alternation responding did not produce the IQ × Problem Type interaction. Additional analyses were undertaken in an effort to determine which of the differences between 2D and 4D problems might explain the interaction. These analyses included comparisons between normal and retarded subjects in frequency of choice of each of the four stimulus dimensions; frequency of position alternation, double alternation, and perseveration; frequency of hypotheses in the first three trials of each 4D blank-trial block (i.e., equated in length to 2D blank-trial blocks); frequency of hypotheses before and after feedback. None of these comparisons revealed differences that could account for the IQ × Problem Type interaction. A description of these analyses may be obtained from the senior author.

Discussion

The present findings suggest that MA and IQ exercise relatively independent effects on hypothesis behavior. Increasing mental age appears to have had its principal impact on the ability to respond correctly following the information that a response is wrong. IQ, on the other hand, had its most direct effect on correct responding following the information "right"; in addition, however, a significant IQ × Problem Type interaction indicated that normal and retarded subjects used hypotheses with about equal consistency when presented with relatively simple problems, but that the retarded used fewer hypotheses when confronted with more complex problems. The same pattern occurred with respect to repetitions of hypothesis patterns following the information "right."

The main effect of IQ on correct responding following "right" is consistent with the general "difference" position that retarded children's learning ability is inferior to that of their nonretarded MA peers, and both the IQ × Problem Type interactions are consistent with the more refined difference hypothesis of Harter (1965) and Weir (1967) that IQ effects are a function of task complexity. Such interactions, reflecting the extent to which hypotheses are sustained across trials, may account in part for the pattern of previous findings by which complex tasks resulted in faster learning by normals than by MA-matched retarded subjects, while relatively simple tasks yielded no reliable group differences (cf. Harter, 1965; Weir, 1967; Zeaman & House, 1967; Zigler, 1966). However, the precise explanation for the role of IQ in these findings still remains a problem for further study.

In an effort to minimize potentially detrimental motivational effects, the present procedure included only noninstitutionalized retarded children, attractive stimuli, an engaging and minimally fatiguing task, reward-plus-penalty conditions, and prior contact with the experimenter (during Binet testing) in ways the literature suggests will enhance motivation to do well. These factors may have contributed to the finding of hypothesis generation among untrained children at a kindergarten MA level, a finding not previously reported. There have been two previous contrary findings. One of these (Rieber, 1969) may have resulted from failure to apprise subjects of the possible forms that a correct solution might take; the investigator's implicit definition of an "hypothesis" as one of those patterns that the investigator presumed might reflect a hypothesis, makes for an inadequate analysis. The failure of kindergartners to manifest hypothesis behavior in Gholson et al.'s (1972) study may have resulted from fatigue caused by the lengthy, impersonally presented (machine projected) problem set, the absence of incentive (e.g., prizes) for good performance, or the potentially stress-inducing feedback arrangements employed: "No matter which response the subject made it was characterized as wrong during the first few feedback trials of most problems" (Gholson et al., 1972, p. 440). Recent evidence from Gholson and McConville (1974) reveals that kindergartners given training designed to enhance their attention to, and differentiation of, nonposition dimensional cues, subsequently generated nonposition hypothesis patterns at above-chance levels (73.2%). This, like the present finding, indicates that once the relevant problem dimensions are recognized and differentiated, hypothesis generation, as defined by the blank-trials format, is not beyond the ability of kindergarten age children.

In interpreting the present findings, a word of caution is in order. The blank-trials procedure developed by Levine (1966) requires further examination before we can state with confidence precisely what is being measured. To be credited with using hypotheses, a subject must clearly go beyond simply formulating a reasonable notion as to the correct answer; he must sustain this notion across a series of unreinforced trials. There are a number of possible reasons why retarded subjects confronted with the more complex problems in this study were relatively less effective at sustaining hypotheses and at repeating hypotheses following the information "right." One possible explanation is that the retarded suffer from a retention deficit (cf. Zeaman, 1973), but an alternative possibility is suggested by evidence that the retarded child lacks confidence in, and readily abandons, his own cognitive strategies when problem solving proves difficult (cf. Gruen & Zigler, 1968; Turnure & Zigler, 1964). An investigation is now in progress to explore the influence of such factors as self-confidence and feelings of personal efficacy on hypothesis behavior as presently defined.

REFERENCE NOTES

1. Hollingshead, A. B. *Two-factor index of social position.* Unpublished manuscript, 1957. (Available from A. B. Hollingshead, Department of Sociology, Yale University, New Haven, Connecticut 06510.)
2. Weisz, J. R. *Hypothesis behavior and learned helplessness.* Manuscript in preparation.

REFERENCES

Achenbach, T. M. Conservation of illusion-distorted identity: Its relation to MA and CA in normals and retardates. *Child Development*, 1969, *40*, 663–679.

Achenbach, T. M. Comparison of Stanford-Binet performance of nonretarded and retarded persons matched for MA and sex. *American Journal of Mental Deficiency*, 1970, *74*, 488–494.

Achenbach, T. M. Stanford-Binet Short-Form performance of retarded and nonretarded persons matched for MA. *American Journal of Mental Deficiency*, 1971, *76*, 30–32.

Achenbach, T. M. Surprise and memory as indices of concrete operational development. *Psychological Reports*, 1973, *33*, 47–57.

Brown, A. L. Conservation of number and continuous-quantity in normal, bright, and retarded children. *Child Development*, 1973, *44*, 376–379.

Eimas, P. D. A developmental study of hypothesis behavior and focusing. *Journal of Experimental Child Psychology*, 1969, *8*, 160–172.

Gholson, B., Levine, M., & Phillips, S. Hypotheses, strategies, and stereotypes in discrimination learning. *Journal of Experimental Child Psychology*, 1972, *13*, 423–446.

Gholson, B., & McConville, K. Effects of stimulus differentiation training upon hypotheses, strategies, and stereotypes in discrimination learning among kindergarten children. *Journal of Experimental Child Psychology*, 1974, *18*, 81–97.

Gruen, G., & Zigler, E. Expectancy of success and the probability learning of middle-class, lower-class, and retarded children. *Journal of Abnormal Psychology*, 1968, *73*, 343–352.

Harter, S. Discrimination learning set in children as a function of IQ and MA. *Journal of Experimental Child Psychology*, 1965, *2*, 31–43.

Harter, S., & Zigler, E. Effects of rate of stimulus presentation and penalty conditions on the discrimination learning of normal and retarded children. *Developmental Psychology*, 1972, *6*, 85–91.

Levine, M. Hypothesis behavior by humans during discrimination learning. *Journal of Experimental Psychology*, 1966, *71*, 331–338.

Rieber, M. Hypothesis testing in children as a function of age. *Developmental Psychology*, 1969, *1*, 389–395.

Turnure, J., & Zigler, E. Outer-directedness in the problem solving of normal and retarded children. *Journal of Abnormal and Social Psychology*, 1964, *69*, 427–436.

Weir, M. Mental retardation. *Science*, 1967, *157*, 576–577.

Zeaman, D. One programmatic approach to retardation. In D. K. Routh (Ed.), *The experimental psychology of mental retardation*. Chicago: Aldine, 1973.

Zeaman, D., & House, B. The relation of IQ and learning. In R. Gagne (Ed.), *Learning and individual differences*. Columbus, Ohio: Charles E. Merrill, 1967.

Zigler, E. Mental retardation: Current issues and approaches. In M. L. Hoffman & L. W. Hoffman (Ed.), *Review of child development research* (Vol. 2). New York: Russell Sage Foundation, 1966.

Zigler, E. Developmental versus difference theories of mental retardation and the problem of motivation. *American Journal of Mental Deficiency*, 1969, *73*, 536–556.

Zigler, E., & Butterfield, E. C. Motivational aspects of changes in IQ test performance of culturally deprived nursery school children. *Child Development*, 1968, *39*, 1–14.

(Received July 22, 1974)