Cognitive Development in Retarded and Nonretarded Persons: Piagetian Tests of the Similar Sequence Hypothesis

John R. Weisz
University of North Carolina at Chapel Hill

Edward Zigler
Yale University

From the debate over developmental "universals" in Piagetian theory and the controversy between developmental and difference theories of mental retardation, an important hypothesis emerges—one that is testable via cognitive-developmental comparisons between retarded and nonretarded persons. This similar sequence hypothesis holds that retarded and nonretarded persons traverse the same stages of cognitive development in the same order, differing only in the rate at which they progress and in the ultimate developmental ceiling they attain. Current evidence relevant to this hypothesis is drawn from 3 longitudinal and 28 cross-sectional studies of developmental phenomena described by Piaget. The great preponderance of this evidence supports the hypothesis with respect to every subject group, with the possible exception of individuals suffering from pronounced electroencephalogram abnormalities. The quality of current evidence is critically evaluated, and procedures by which more precise tests of the hypothesis might be fashioned are proposed. Overall, the review illustrates that developmental research with atypical populations can be a potent tool in testing general developmental theory. Conversely, it illustrates the power of general developmental theory to enrich our understanding of atypical development.

In recent years two important theoretical issues have stimulated interest in Piagetian research with retarded and nonretarded populations. One is the question of whether developmental "universals" exist. Many psychologists regard the sequence of developmental stages described by Piaget (e.g., 1970) and elaborated by other cognitive-developmental theorists (e.g., Kohlberg, 1969) as one of psychology's few current candidates for universality (see Weisz, 1978). Piaget (1956, 1966) took a psychological universalist position, with qualifications, and Kohlberg (1969, 1971) argued strongly for the invariance of what he regarded as a cognitive-developmental sequence rooted in an inherent logic and in universal characteristics of both the nervous system and the environment.

Of course it is impossible to know that any given developmental phenomenon occurs everywhere without exception, since one can never test all possible exceptions (Popper, 1959). However, if one is not to let the claims of cognitive-developmental theorists go unchallenged, it is important to evaluate the extent to which transcontextual validity (see Weisz, 1978) has been demonstrated for the Piagetian account of development.

One approach to assessing such validity across experimental contexts is to examine developmental sequences across various cultures (Buck-Morss, 1975; Simpson, 1974). Another approach, of particular interest because of the cognitive emphasis of Piagetian theory, is to compare groups of children who differ markedly in measured intelligence, that is, groups of mentally retarded and nonretarded children. If children at very different IQ levels were to show identical Piagetian developmental sequences, then the transcon-
textual validity of the Piagetian account of development would be substantially supported. If retarded and nonretarded children were to differ in their sequence of development, then universality could hardly be claimed for the Piagetian account.

A second theoretical issue that has sparked recent interest in comparative cognitive research is reflected in the ongoing debate between proponents of developmental and difference theories of mental retardation. The developmental position, set forth by Zigler (1969), is intended to apply to retarded individuals not suffering from organic impairment. Zigler has maintained that the retarded child passes through cognitive-developmental stages in the same order as the nonretarded child, with only two differences: The retarded child passes through the stages more slowly and attains a lower upper limit relative to the nonretarded child.\(^1\)

A number of theorists hold what Zigler labeled the general difference position. One aspect of this position is the view that the cognitive development of retarded persons differs from that of nonretarded persons in ways that go beyond mere differences in rate and ceiling of development. Milgram (1973), for example, maintained that the cognitive levels, or stages, of retarded children differ from those of the nonretarded in that the former are more likely to contain traces of developmentally earlier levels and are more likely to show regression to those earlier levels. (For further discussion of the developmental position, the specific difference positions, and the rationale underlying them, see Weisz & Zigler, in press.)

This theoretical conflict has generated a new emphasis on comparative research into the processes (rather than the products) of learning and reasoning (Weisz, 1977; Weisz & Achenbach, 1975) and into processes of reasoning, as described in Piagetian theory (see Wilton & Boersma, 1974).

The growing interest in the pursuit of developmental universals, and the growing intensity of the developmental versus difference debate, have thus combined to lend theoretical force to research comparing the cognitive development of retarded and nonretarded persons along Piagetian lines. This body of research has grown rapidly within the past decade; it now appears to be substantial enough to serve as a resource in our efforts to answer the principal question raised by the universality issue and the developmental versus difference debate. This question can be stated in the form of a testable hypothesis.

**Similar Sequence Hypothesis**

An appropriate label seems to be the similar sequence hypothesis. The hypothesis holds that during development retarded and nonretarded persons traverse the same stages in precisely the same order and differ only in rate of development and in the ultimate ceiling they attain. To be precise, the developmental position (Zigler, 1969, 1971) generates this hypothesis only with respect to nonretarded and cultural-familial retarded persons (thus excluding, for example, brain-damaged and genetically impaired individuals).\(^2\) In

---

\(^1\) An additional postulate of the developmental position is that familial retarded and nonretarded persons who are equivalent in developmental level (often operationally defined as mental age) do not differ in the formed cognitive processes they employ in learning and reasoning. This particular proposition is not germane to the present review and consequently is not discussed here. However, Piagetian evidence bearing on this proposition is being reviewed (Weisz & Zigler, in press).

\(^2\) The reasoning underlying this qualification bears brief explanation. The developmental position holds that mental retardation can be viewed as a developmental phenomenon most appropriately among persons whose retardation does not result from specific physiological defects. Such investigators as Benton (e.g., 1962), Cruikshank (e.g., 1967), and Reitan (e.g., 1973) have devoted many years to demonstrating idiosyncratic performance characteristics that distinguish brain-injured individuals from those with intact nervous systems. Furthermore, a number of studies employing the specific kinds of problem-solving tasks most often used in research on the developmental-difference controversy have revealed effects of organ-  

icity on retarded children’s performance (Balla, Styfco, & Zigler, 1971; Balla & Zigler, 1964; Elkind, Koegler, Go, & Van Doorninck, 1965; Harter, Brown, & Zigler, 1971). In harmony with such findings, proponents of the developmental position have adhered to the two-group approach (see Zigler, 1969), whereby familial retarded individuals are distinguished from those suffering from organic impairment (including genetic anomalies such as Down’s syndrome). There is some disagreement among investigators over the need for the two-group approach (Ellis, 1969; Milgram,
cognitive-developmental theory, however, the claims for the universality of the developmental sequence appear to be broader. Piaget (1956) held that “the minimum program for establishment of stages is the recognition of a distinct chronology, in the sense of a constant order of succession” (p. 13). According to Kohlberg (1969), the claim that there is an invariant order of cognitive stages rests upon an assumed invariance in certain features of the environment and of the nervous system and upon “a logical analysis of orderings inherent in given concepts” (p. 355). These inherent orderings are seen as logically essential and as independent of individual differences among people. Kohlberg continued, “The invariance of sequence in the development of a concept or category is not dependent upon a prepatterned unfolding of neural patterns; it must depend upon a logical analysis of the concept itself” (p. 355). Thus, the similar sequence hypothesis as advanced by cognitive-developmental theorists seems to predict a truly universal ordering of stages—an ordering that is the same for retarded children of all etiologies (including genetic impairment, brain injury, and other neurological anomaly) as it is for all nonretarded children. There is a conservative version of the similar sequence hypothesis that applies only to familial retarded and nonretarded persons and a liberal version that applies to all persons. In the present article we present evidence bearing on both versions.

In contrast with both these versions, Milgram (1973) has argued that the retarded child’s cognitive stages differ from those of the nonretarded child. In contrast with the liberal version of the hypothesis, Rogers (1977) has described a rationale for (though she has not necessarily endorsed) the hypothesis that profoundly (and thus nonfamilial) retarded children have abnormal developmental patterns.

Material Excluded From the Present Review

The present article is an attempt to synthesize studies relating to the similar sequence hypothesis. In selecting studies to be reviewed, we excluded studies of reading per se and of language development per se. Although both areas can be viewed from the perspective of Piagetian theory, neither is central to the theory; furthermore, the research in both areas is now so voluminous as to warrant separate review. We also excluded studies designed to accelerate cognitive development, since it is not the purpose of this article to determine whether retarded or nonretarded children can be trained more readily.

The studies we do include in this review vary widely in their sampling procedures, their experimental methodology, and their approaches to data analysis and reporting. Consequently, the studies differ in their level of importance vis-à-vis the hypothesis of particular interest here. For this reason, we reserve the right to vary the level of detail in which we describe the studies and give relatively greater space to those that seem to afford the clearest tests of the hypothesis.

Tests of the Similar Sequence Hypothesis

Cross-Sectional and Order-of-Difficulty Evidence

One approach to testing the similar sequence hypothesis is to assess groups of mentally retarded children at more than one developmental level with respect to their performance on various Piagetian tasks. If the direction of the difference in performance from one developmental level to another is the same for the retarded as for the nonretarded, or if the direction is consistent with the developmental sequence posited by cognitive-developmental theory, then the similar sequence hypothesis is supported. A second general approach to testing the similar sequence hypothesis is to rely less upon the developmental levels of the groups sampled than upon the relative-difficulty levels of the various tasks or behavioral items being employed. Perhaps the simplest, but also the least informative, of the variants of this approach

1973). Moreover, there is a strong Piagetian rationale for applying the similar sequence hypothesis to all persons, retarded or nonretarded, organically impaired or intact (see the remainder of the paragraph for details). In what follows we describe this rationale, and we go on to review evidence in a manner that bears directly on the conservative, two-group-oriented version of the hypothesis and on the more liberal version in which the similar sequence hypothesis is applied to all retarded groups regardless of etiology.
is to rank order the items with respect to the number of subjects who pass each one; if this rank ordering of a retarded sample matches either the rank ordering of a nonretarded sample or the developmental sequence posited by cognitive-developmental theory, then the similar sequence hypothesis is supported, albeit modestly. A more informative type of order-of-difficulty evidence is the type that employs scaling procedures, allowing one to determine, for example, how many of the children who grasp Concept A also grasp Concept B and vice versa. Such evidence, when combined with Guttman-type (e.g., Guttman, 1950) scalogram analyses, can provide a relatively strong test of the similar sequence hypothesis. The studies reviewed in this section all employed some type of cross-sectional evidence, order-of-difficulty evidence, or a combination of the two.

Development in the sensorimotor period. Early evidence bearing on the similar sequence hypothesis was provided by the research of Woodward (1959, 1961, 1962, 1963). The first of her studies (1959) focused primarily on a group of 65 institutionalized children and adolescents with a chronological age (CA) range of 7–16 years who were so profoundly retarded that they failed to attain a basal age of 2 years on the Terman-Merrill scale. Although the author maintained that this sample excluded cases involving motor or sensory disability, the cases involved a diversity of medical problems (e.g., 19 subjects were epileptic), and 38 of the children were “emotionally unstable.” Woodward used three means of assessing the sensorimotor stages of this heterogeneous group. First, she observed their spontaneous mannerisms and their manipulation of toys presented individually and in a standardized order. Second, each subject was presented with three pairs of tasks, each pair tapping one of Piaget’s (1953, 1955) last three sensorimotor stages (there are six stages in all). Third, Woodward presented each child with a series of object concept tasks in which a piece of candy or a toy was first used to attract the subject’s attention and was then withdrawn and concealed to varying degrees.

All but the object concept tasks were analyzed in a way that sheds light on the similar sequence hypothesis. Each task was classified with respect to the Piagetian sensorimotor stage it was designed to represent; then the tasks were ranked ordered with respect to the percentage of subjects passing each. The difficulty level rankings of these 11 items were identical to the Piagetian stage level order, with one exception: A task involving coordination of vision and hearing (Sensorimotor Stage 2) proved to be slightly more difficult than a task involving manipulation of objects (Stage 3); 53 subjects passed the manipulation task, and only 49 passed the coordination task. Furthermore, when the possibly insensitive coordination task was removed from the analyses, 59 of the 65 children passed all of the items at stages below their highest stage level response. Given the extreme diversity of this sample, the high incidence of emotional instability, and the apparent tendency of many not to show responses of which they were actually capable (e.g., some delayed for a half hour before grasping an object placed before them), these data lend surprisingly strong support to the similar sequence hypothesis.

Recently, Rogers (1977) undertook an investigation similar to that of Woodward in several respects. The subjects, 40 profoundly retarded children ranging in age from 8–14 years, with IQs below 20, were given a series of Piagetian tasks. By means of these tasks, each child’s performance was classified into Sensorimotor Stage 3, 4, 5, or 6 in each of four conceptual domains: object permanence (tasks involving searches for a hidden object), spatiality (tasks involving visual anticipation and rotation of objects), causality (tasks involving the use of physical prompts and tools, the removal of obstacles, and inference as to the cause of a jingling sound inside a box), and imitation (tasks involving the reproduction of both self-initiated and experimenter-initiated movements and sounds). Performance within each of the four domains was analyzed using scaling techniques, and Guttman’s (1950) coefficient of reproducibility and Green’s (1956) index of scalability were calculated for each scale. The object permanence and imitation tasks formed highly reproducible scales in the orders hypothesized by Piaget (1955, 1962, 1972). Causality tasks also
formed a highly reproducible scale, although the item order differed from the predicted sequence in one respect: One Stage 6 item preceded one Stage 5 item. The author attributed this irregularity to a poor choice of Stage 6 task (i.e., opening a box to obtain a bell when box opening had just been demonstrated to the subject), “since the task used might have been accomplished using imitation rather than problem-solving skills” (Rogers, 1977; pp. 841–842). Finally, the individual spatiality tasks did not all form highly reproducible scales, but when the tasks within each stage were combined (and subjects were credited with a stage level for passing one or more of the tasks from that level), the stages did form a highly reproducible scale. Rogers concluded convincingly that her findings support “the invariant sequentiality of sensorimotor stages” (p. 841).

The preoperational–concrete operational transition—the Inhelder study of conservation. One of the earliest studies bearing on the similar sequence hypothesis was carried out by Piaget’s associate, Barbel Inhelder, in the early 1940s. This study, now published in English (Inhelder, 1943/1968), involved the assessment of conservation of substance, weight, and volume in 159 persons who had been labeled mentally retarded by Swiss education officials. The sample was extremely heterogeneous (see Jordan, 1976). Ages ranged from 7½–52 years, IQs ranged from 35–104, institutionalized and noninstitutionalized persons were included, and the range of etiologies and physical maladies included such diverse states as “defective environment,” rickets, hearing defect, “abandoned,” and schizophrenia. The procedure involved semistructured clinical interviews with each subject. Since the procedure was not perfectly standardized and little in the way of formal data analysis was presented, it is difficult to evaluate Inhelder’s conclusions, including her references to “oscillations” in the reasoning of retarded subjects, discussed later in this article. However, in Piaget’s (1968) description of the Inhelder study, he explained that in the entire sample, not one [individual] understood the conservation of weight without having the conservation of substance, nor the conservation of volume without both weight and substance, while the conservation of substance was found without the other two, and the conservation of weight was found without the conservation of volume. (p. 11)

Given the marked heterogeneity of the sample, such uniform support for the similar sequence hypothesis is noteworthy.

Other studies of conservation and related concepts using retarded samples only. Studies of conservation and related concepts done since the Inhelder investigation have a bearing on the similar sequence hypothesis, despite the fact that they only sampled retarded subjects. Klaus and Green (1972) assessed conservation of number and volume in 27 trainable mentally retarded subjects ranging in age from 13–19 years and in IQ from 29–57. These investigators found that volume conservation presented greater difficulty than did number conservation, a finding consistent with the pattern apparent in the nonretarded. Marchi (1971) tested conservation of mass, weight, and volume in 106 educable mentally retarded children. Difficulty level evidence suggested that contrary to Marchi’s prediction, the retarded “follow a similar sequence in the acquisition of mass, weight, and volume as postulated for normals” (p. 6442).

Roodin, Sullivan, and Rybash (1976) assessed qualitative identity, quantitative identity, and equivalence conservation (see Elkind, 1967) in 60 institutionalized retarded children averaging 13 years of age and about 47 in IQ. Dyed water was poured from a standard 100 ml beaker; to test qualitative identity, subjects were asked, “Is the water in this glass (comparison) the same water that was in that glass (empty standard)?” To assess quantitative identity, subjects were asked, “Is there as much water in this glass (comparison) as there was in that glass (empty standard)?” To assess equivalence conservation, two standard beakers were filled with equal levels of water, and the contents of one were then poured into a comparison beaker; the experimenter then asked, “Is there as much water in this glass (standard) as there is in this glass (comparison)?” Previous research (e.g., Papalia & Hooper, 1971) with nonretarded children had suggested that the developmental order for the attainment of these three concepts would be qualitative identity, quantitative identity,
and equivalence conservation. In the Roodin et al. study, analyses of the number of conservers on each task indicated a parallel order of difficulty.

In a similar study also employing 60 institutionalized retarded children (age range of approximately 10–16 years and average IQ of approximately 57), McManis (1969c) investigated identity and equivalence conservation with three types of material (Styrofoam balls, clay, and water). Like Roodin et al. (1976), McManis found evidence that the developmental sequence of his retarded subjects replicated that of nonretarded children. The notion that identity conservation must precede equivalence conservation was supported by the finding that no subject who failed to achieve identity conservation showed equivalence conservation, whereas 13%–18% of the subjects (precise percentage depending on the particular task used) displayed identity conservation without equivalence conservation.

Three studies that examined conservation of number, and number concepts generally, in mentally retarded groups yielded similar findings, despite some differences in methodology. Woodward (1961) investigated numerical concepts of 94 institutionalized individuals (50 adults with average CA of 19 years and 44 children and adolescents with average CA of 12 years) ranging in IQ from 25–73. Tests given to the subjects included assessments of their understanding of (a) one-to-one correspondence and equivalence of corresponding sets, (b) ways of equalizing unequal groups, (c) seriation, and (d) conservation of continuous quantity (water and sand). Performance was scored as indicative of one of two preoperational stages or of concrete operational thinking. When the stage level assignments were plotted as a function of the IQs (and thus roughly of the mental ages or MAs) of the adult subjects sampled, the table reflected precisely what would be expected from the application of Piaget's stage scheme to nonretarded individuals.

In the second of the three studies, Mannix (1960) administered eight of Piaget's (1952) number concept tasks to 48 "educationally subnormal" individuals ranging in MA from 5–9 years. The tasks included two tests of additive composition, one test of coordination of equivalence relations, two tests of judgment of correspondence between sets of items, and two conservation tasks (continuous and discontinuous quantities). Responses to these tests were classified into Piagetian stage levels, a scalogram was constructed, and the coefficient of reproducibility was .94. Mannix's brief report gave little information as to the precise nature of the scale types; but apparently the scalogram was consistent with Piaget's stage theory, because the author concluded that educationally subnormal children "pass through the three stages of development described by Piaget" (Mannix, 1960, p. 181).

The third of these studies on number concepts was conducted with 20 institutionalized mentally retarded persons in New Zealand (CA range of 8–17 years; IQs of 29–65). Singh and Stott (1975) presented these subjects with a series of number conservation tasks designed to classify them with respect to three Piagetian number stages: Stage 1—child fails to attend to relevant cues and fails to conserve; Stage 2—child selectively attends to only certain relevant cues, can match perceptually, but cannot conserve; Stage 3—child conserves, showing understanding of invariance of properties despite transformation in appearance. Data bearing on the similar sequence hypothesis are not reported in detail, but the authors' conclusion is quite clear: "Retarded children apparently develop sequentially in the same order as normals but at a slower rate and at a later CA" (Singh & Stott, 1975, p. 220).

One other study that used only a retarded sample deserves mention both because of its scope and because of an important issue it raises. In this study, Lister (1972) assessed six types of conservation among 115 educationally subnormal pupils in Great Britain. The subjects were aged 8–16 years, and their IQs ranged from 47–81. Both difficulty level rankings and a scaling procedure strongly suggested the following developmental sequence in the emergence of these types of conservation: number, substance, length, weight, volume, and area. Although no scalogram statistics were calculated, only 6 of the 115 subjects showed a scalogram response pattern inconsistent with the preceding order.
Lister noted that the order with respect to substance, weight, and volume was consistent with previous Piagetian research, whereas the suggested order of the remaining attributes differed from at least some previous findings with nonretarded subjects. Her own interpretation of the discrepancies was that they resulted from experiment-to-experiment variations in the specifics of the problems used to assess the various types of conservation. This is a very real possibility, and it is one reason why tests of the similar sequence hypothesis that expose retarded and nonretarded subjects to the same experimental procedures must be regarded as stronger evidence than experiments that test only retarded children and compare the findings with those of different experiments. We now turn to six studies of the former type.

Studies of conservation and related concepts employing both retarded and nonretarded subjects. Four of the experiments in which the performance of retarded and nonretarded subjects was directly compared were conducted by McManis (1969b, 1969d, 1969e, 1970). In one of these, 90 institutionalized retarded subjects (IQs of 47–73) and 90 nonretarded elementary school children (IQs of 85–115) were tested for conservation of mass, weight, and volume of clay, using Piaget and Inhelder's (1941) "sausage" technique. About half the retarded subjects were organically impaired. Analyses of the mean scores for the conservation tasks indicated that conservation of mass was easiest and conservation of volume most difficult for both the retarded and the nonretarded group, providing some support for the notion that the order of emergence of these types of conservation in groups of both average and below-average IQ is as follows: mass, weight, then volume.

In another article McManis (1969e) reported his assessment of conservation and transitivity of weight (clay) and length (sticks) in what appears to be the same sample used in his 1969d experiment. The study was designed to test the hypothesis, derived from Kooistra (1964), that for any given property (e.g., weight) conservation will appear developmentally earlier than transitivity. The results supported this hypothesis for both weight and length in retarded and nonretarded subjects. McManis (1970) then explored the relations among conservation, seriation, and transitivity (of length) within groups of 80 institutionalized mentally retarded persons (IQs of 46–72) and 80 nonretarded elementary school children (IQs of 85–116). Among both retarded and nonretarded children who showed discrepant performance on the conservation and seriation tasks, nearly all showed conservation without seriation. Among both retarded and nonretarded children who showed discrepant performance on the seriation and transitivity tasks, nearly all showed seriation without transitivity. These findings indicate that seriation falls developmentally between conservation and transitivity (at least with respect to the property of length, as measured in this experiment) for both retarded and nonretarded persons.

In a fourth article based on the same sample used in two of the preceding studies (McManis, 1969d, 1969e), McManis (1969b) tested Piaget's (1952) view that there are three hierarchically ordered stages in the development of quantitative comparison processes. In the first stage, children are said to consider only uncoordinated perceptual relations of gross qualitative equality or difference; in the second stage, intensive quantity, children are said to compare quantities by seriating them along more than one dimension (e.g., width and height) simultaneously; in the third stage, extensive quantities, children are said to be capable of overruling apparent differences between two equal quantities by imposing equal units of measurement upon them. McManis tested his young subjects' performance of these three types of comparison, using sticks, colored water, and beads. The analysis of scores on these tasks indicated that for both the retarded and the nonretarded group, gross comparisons were the simplest (they were passed by nearly all subjects in both groups) and extensive comparisons were the most difficult. These findings are consistent with the view that for children at both IQ levels the developmental order is as follows: gross, intensive, and extensive quantity (the order posited by Piaget). One other study (McManis, 1969a) should be mentioned in this connection. McManis's tests of quantity com-
parison were given to 140 institutionalized mentally retarded persons, who were divided into equal groups representing different IQ levels (IQs of 30-49 and of 50-69). The procedure and analyses were similar to those employed in the preceding study (McManis, 1969b). The results indicated, as in the preceding experiment, that comparison of gross quantities was easiest and comparison of extensive quantities was most difficult, regardless of the IQ level of the subjects.

Three other comparative studies were designed to address the problem of order of events across the types of conservation. Gruen and Vore (1972) assessed conservation of number (poker ships), continuous quantity (water), and weight (clay rolled into various shapes) in familial retarded (IQs of 55-80) and nonretarded (IQs of 90-120) public school pupils. Both retarded and nonretarded groups were divided into three subgroups of MA: 5, 7, and 9 years. Evidence on developmental ordering was in the form of mean scores for the three types of conservation, analyzed within each MA level. In one set of analyses, conservation judgments alone (i.e., disregarding the subjects' verbal explanations) constituted the dependent variable. With this criterion, performance of nearly all subjects at MA 9 was correct; for the other two MA levels, both retarded and nonretarded subjects tended to score significantly better on the number task than on the quantity or weight tasks. For nonretarded children at MA 7, however, the differences were not significant. The quantity and weight tasks did not differ significantly in difficulty for retarded and nonretarded children.

In a second set of analyses by Gruen and Vore, the dependent measure was conservation judgment in combination with the subject's explanation for that judgment. Using this criterion, there was no significant task effect at the MA 5 level. At the MA 7 level both retarded and nonretarded subjects tended to score significantly better on the number task than on the quantity or weight tasks. For nonretarded children at MA 7, however, the differences were not significant. The quantity and weight tasks did not differ significantly for the retarded and nonretarded children.

We conclude this section of the review by discussing two conservation studies by Achenbach. Building on the work of Charlesworth (1969) and Mermelstein and Shulman (1967), Achenbach (1973) inferred children's identity concepts with respect to color, number, length, and continuous quantity from their surprise reactions to contrived changes in those properties. For example, to test number identity concepts two toy Indians were placed in a box, and when the bottom was opened three Indians dropped out. Among nonretarded subjects (M IQ = 116), there were significantly more frequent surprise reactions to a change in color than to changes in the three quantitative properties, a finding consistent with Piaget's view (see Piaget & Voyat, 1968) that children develop identity concepts for qualitative properties such as color prior to the emergence of identity concepts for quantitative properties such as number, length, and continuous quantity. Surprise reactions to change in color and number were virtually identical in 45 familial and 16 Down syndrome retarded subjects (M IQ = 47). The frequencies of surprise reactions to changes of the three different quantitative properties for both retarded and nonretarded subjects are consistent with the findings of others (including Gruen & Vore, 1972) that successful performance on conventional length and number conservation tasks is simpler than, and thus presumably developmentally prior to,
success on conventional continuous quantity tasks. Thus, once again we see fairly strong support for the view that the sequence of developmental events for the retarded child is similar to that for the nonretarded child.

A different type of similarity is illustrated by Achenbach’s (1969) study of nonretarded public school children (IQs of 94–168) and nonorganically impaired retarded children (IQs of 31–78) from public schools and institutions. He assessed children’s conservation concepts with respect to length, area, and volume\(^3\) (employing 4 tasks for each of the three properties) by using optical illusions to create discrepancies between the actual and apparent sizes of various stimuli. To test conservation of length, for example, the experimenter presented each child with a barbell illusion in which a small metal rod that fit into a groove that just touched the inner edges of two circles was placed into another groove that passed through two circles and touched their outer edges. The effect was to make the rod appear longer in the second position than in the first. Subjects were then asked whether the rod would fit into the original groove. An important feature of the study, for our purposes, is that the 12 tasks were designed to be free of intellectual demands in the areas of additivity, numeration, conservation of equivalence, or complex verbal expression—dimensions along which more traditional conservation tasks often vary. This made it possible to test the contention of Braine and Shanks (1965a, 1965b) that the attainment of conservation with respect to the various properties would be parallel if the performance criteria used were standard across the types of conservation. Consistent with the Braine–Shanks view, Achenbach (1969) found a “total absence of evidence for a horizontal décalage” (p. 677) in the three types of conservation, for both the retarded and the nonretarded group. In both groups, there were neither consistent nor significant differences in success rates for length, area, and volume tasks. This finding, together with the reasoning of Braine and Shanks, suggests that some of the order-of-difficulty and scalogram-type evidence reviewed in the preceding paragraphs may be more indicative of differences in the specific requirements of the contrived tasks employed than of actual differences in the order of emergence of the various types of conservation.

**Concepts of time.** A number of studies have addressed the similar sequence hypothesis in content areas other than conservation. In one study on the concept of time, Lovell and Slater (1960) interviewed 50 educationally subnormal children (IQs not reported) aged 8, 9, 10, 11, and 15 years and 50 “average to above average” children aged 5–9 years. The interview included tasks (some were Piaget’s) designed to measure concepts of simultaneity and equality of synchronous intervals (e.g., asking the child to judge whether two dolls traveling at different rates but starting and stopping at the same time actually traveled for the same amount of time). There were also tasks involving chronological ordering of events and children’s concepts of age and interior time. Little in the way of statistical analysis was reported, but nonetheless, Lovell and Slater concluded that the understanding of these five concepts of time follows roughly the same sequence in retarded as in normal children, although the stages in understanding are reached some years later by retarded children.

**Concepts of space.** Two studies that included a retarded sample were specifically concerned with spatial concepts. In one of these, Houssiaudas and Brown (1967) sampled 40 institutionalized, mentally retarded Australians (\(M\) IQ = 55) who showed no evidence of mongolism or other specific defects and who ranged in age from 8–15 years. These subjects were presented with two perspective-taking tasks (one pictorial and one using manipulation of actual objects) in which they were asked to identify their own perspective on a perceptual array as well as the perspective of another person seated at a different position. Although no statistical analyses were reported, the pattern of passes and failures on the different items was consistent with the view that retarded individuals pass through the stages identified by Piaget; that is, first there is difficulty in identifying one’s own perspective

---

\(^3\) The volume conservation task used by Achenbach (1969) actually tapped what Piaget (has) called conservation of continuous quantity.
and that of another, second there is only
difficulty in identifying how a perceptual
array might look from another position, and
third the individual is able to “coordinate
perspectives,” identifying not only his own
perspective but that of another person as well.
Summing up, Houssiadas and Brown (1967)
concluded, “It is clear that the pattern of
predominant responses follows the same
sequence suggested by Piaget, whose data were
derived from normal children” (p. 213).

In a more fine-grained analysis of spatial
concepts, Woodward (1962) tested the same
institutionalized retarded group of 50 adults
and 44 children used in the study of number
concepts described earlier (Woodward, 1961).
In this sample, 50% of the adults and 61%
of the children showed some type of organic
impairment. The spatial tasks included mea-
sures of the ability to reproduce a spatial
order under varying degrees of transformation
(e.g., reproducing a circular array of beads on
a horizontal rod). Using a similar procedure
with nonretarded children, Piaget and Inhelder
(1956) identified seven stages through which
their children passed as they improved on the
tasks. Woodward (1962) constructed a table
of scale types to assess the comparability of
her results with those of Piaget and Inhelder.
Although no scalogram statistics were cal-
culated, the great majority of Woodward’s
subjects fit scale types consistent with the
developmental sequence posited by the
Genevans.

A second task employed by Woodward
involved drawing copies of 21 geometric
figures used by Piaget and Inhelder. The
compatibility of subjects’ scores on these
tasks with a four-stage sequence advanced
by Piaget and Inhelder (1956) was demon-
strated by the fact that “subjects classified
by the features of a given stage showed the
features of the lower stages in most cases”
(Woodward, 1962, p. 31). However, once
again no scalogram statistics were reported,
and 5 of the 14 performance criteria by
which stage assignments were made were
outside the appropriate difficulty level for
at least some of the subjects. The third task
employed by Woodward was a “reference
points” problem in which adult subjects only
were presented with drawings of a bottle
	tilted at various angles and said to be about
one-fourth full of water. The subjects’ task
was to pencil in the portion of the bottle
occupied by the water. The performance data
presented for this task were extremely sketchy,
but Woodward indicated that the order of
difficulty of the tasks was the same as that
found by Piaget and Inhelder. In her overview
of her findings bearing on what we have called
the similar sequence hypothesis, Woodward
(1962) concluded that for her retarded subjects,
“The sequence suggested by Piaget and
Inhelder [for nonretarded children] was
confirmed for all three spatial concepts that
were investigated” (p. 35).

Relative thinking. In investigating the “logic
of relations” in children, Piaget (1928) used a
“brothers and sisters” problem and a “right
and left” problem. In the former problem,
children’s understanding of the relation be-
tween being and having a sibling was explored
by asking such questions as “George has
three brothers, Paul, Henry, and Charles.
How many brothers has Paul? How many
brothers are there in this family?” In the
right and left problem, children were in-
structed, “Show me your right hand, your left.
Show me my right hand, my left” Lane and
Kinder (1939) used these two Piagetian
problems with 50 institutionalized retarded
individuals of unspecified etiology who were
grouped, for purposes of data analysis, into
different IQ levels: 38, 51, 64, and 77.
Instead of scalogram statistics, relative levels
of the questions were reported for each IQ
group. These data indicated that the rank
ordering of difficulty for the 11 questions was
similar across the different IQ levels—a
parallelism consistent with the similar sequence
hypothesis.

Moral judgment. Abel (1941) investigated
moral judgment in 74 institutionalized “sub-
normal adolescent white girls” (aged 15–21
years; IQs unspecified). Subjects were ques-
tioned about seven brief stories concerning
immanent punishment (the inevitability of
punishment following a misdeed), retributive
justice (punishment orientation, particularly
of the “eye for an eye” variety), and judgments
of the gravity of a misdeed (using information
on consequences of the deed and intent of the
transgressor). Mirroring previous findings with
nonretarded individuals (Lerner, 1937, 1938), Abel's findings were that with increasing maturity (defined in terms of MA) subjects gave nonsignificantly greater weight to intent and less weight to consequences in judging the gravity of a misdeed and were significantly less likely to consistently advocate retributive punishment. Unlike the nonretarded persons in at least some research, Abel's more mature subjects (MAs of 9-11 years) did not show any less pronounced a belief in immanent punishment than did her less mature subjects (MAs of 6-8 years). In fact, about 82% of both groups showed such a belief, which Abel (1941) attributed to the "constraining" institutional environment "that controls the girls with threats of immanent punishment" (p. 386). Except for this one anomaly, the Abel data are consistent with the similar sequence hypothesis.

Studies of multiple concepts. We conclude this section on cross-sectional research with a discussion of studies that have assessed concepts in more than one conceptual domain. DeVries (1970, 1973a, 1973b, 1974) assessed a variety of Piagetian concepts in bright (M IQ ≈ 130), average (M IQ ≈ 105), and retarded (M IQ ≈ 72); etiologies not reported) children, all enrolled in public schools. The tasks included the brothers and sisters and right and left problems described earlier; tests of generic and sex identity and of conservation of mass, number, length, and liquid; interviews on magic and dream concepts; object sorting and class inclusion problems; and a guessing game ("Which hand has the penny?") designed to reveal the level of children's role-taking skills. Of all the tasks used, data from the guessing game task were presented in the most complete manner (see DeVries, 1970). Using an independent sample of 64 high-IQ children, DeVries (1970) classified behavior on the guessing game with respect to 10 characteristics (e.g., does not always hide penny in the same hand). These characteristics formed a highly reproducible Guttman-type scale with a reproducibility of .95 and an index of consistency of .66. The scale was then used with the bright, average, and retarded samples and checked against Kohlberg's (1969) criteria for developmental sequentiality, namely, (a) mean scale scores should increase with age, (b) success on each individual scale item should increase with age, and (c) the sequence of items should be justifiable with a logical rationale based on Piagetian theory. DeVries (1973b) maintained that her scale met the third criterion, and her data (DeVries, 1970) indicate that the first two criteria were met within the bright, average, and retarded groups separately. Similar analyses were carried out with respect to the other 14 Piagetian tasks, with a Guttman scale constructed for each. Within the average and retarded groups each scale met Green's (1956) criterion of an index of consistency greater than .50, and the lowest coefficient of reproducibility was .94 (DeVries, Note 1). DeVries (1973b) indicated that all of the Kohlberg criteria for sequentiality "were applied to each ability group (i.e., bright, average, and retarded subjects) separately, and the order of scale items was the same for each ability group on all tasks" (p. 3).

Stearns and Borkowski (1969) investigated conservation of continuous quantity (water) and discontinuous quantity (blocks and marbles) as well as horizontal-vertical space perception in institutionalized retarded individuals (IQs unspecified) ranging in age from 7½-27 years. Consistent with Piaget's (e.g., 1964) view (and supporting findings; see Elkind, 1961) that conservation of continuous quantity is more difficult and emerges developmentally later than conservation of discontinuous quantity, Stearns and Borkowski found performance on their test of the former concept to be significantly poorer than performance on their two tests of the latter concept. Scores were also highly similar for the tests of horizontal and vertical frames of reference; this finding is consistent with Piaget's (see Piaget & Inhelder, 1956) view that concepts of the vertical and of the horizontal are acquired at the same time.

Finally, we turn to two studies by Lovell and his colleagues (Lovell, Healey, & Rowland, 1962; Lovell, Mitchell, & Everett, 1962). The studies reported few relevant statistical analyses, but the diversity of concepts examined makes them worthy of brief attention. In the Lovell, Mitchell, and Everett study, groups of nonretarded and educationally subnormal individuals (no IQs reported)
were divided into separate age groups. The skills investigated included additive and multiplicative classification (of objects and pictures differing in multiple dimensions), seriation, multiplication of asymmetrical transitive relations, hierarchical classification, class inclusion, and visual and tactile classification. For all tasks the tabled data indicated a general improvement in performance with increasing age level for both normal and subnormal groups (no significance tests reported).

In the study by Lovell, Healey, and Rowland, the subjects were again groups of nonretarded and educationally subnormal persons from special schools (IQs unreported) who were divided into separate age groups. The groups were presented with 12 of the tasks used by Piaget, Inhelder, and Szeminska (1960) to study the child’s geometric concepts. Within normal and subnormal groups separately, correlation coefficients were calculated that related Piagetian stage levels on the 12 tasks to subjects' age levels. Of the 24 coefficients, 23 were significant at the .01 level. In both Lovell et al. studies, the details of subject selection, experimental procedure, and statistical analyses are so skimpy that the findings must be regarded as only suggestive. Nonetheless, although they are not by any means definitive, the data are in harmony with the similar sequence hypothesis.

**Summary of the Cross-Sectional and Order-of-Difficulty Evidence on the Similar Sequence Hypothesis**

Thus far we have reviewed 28 studies in which cross-sectional and order-of-difficulty evidence is reported in ways that have some bearing on the similar sequence hypothesis. The degree of retardation involved in the samples ranged from profound to mild; the retarded persons sampled ranged in age from childhood to adulthood, were both institutionalized and noninstitutionalized, and included both cultural–familial cases and individuals with diverse organic and emotional disorders. The nonretarded contrast groups, when employed, ranged from slightly below average to extremely high in IQ. The studies reported also varied widely in their experimental methodology and in their methods of data analysis. Despite this great diversity in methodology and in sample characteristics, the data reviewed show rather consistent support for the similar sequence hypothesis, both in its conservative form, which applies only to nonretarded and familial retarded persons (Zigler, 1969), and in its broader form, in which universality of developmental sequence is held to be independent of individual subject characteristics such as organic impairment (see Kohlberg, 1969).

There were very few exceptions to this generalization: (a) Woodward (1959) found that among her profoundly retarded subjects 1 sensorimotor task out of 11 proved to be slightly more difficult than another task that Piaget designated as being one sensorimotor stage higher. (b) Among Rogers’s (1977) profoundly retarded subjects, a causality task designed to be at Stage 6 proved to be easier than one of the tasks at Stage 5, whereas individual spatiality tasks within each stage had to be combined to yield a highly reproducible scale. (c) In Achenbach’s (1973) familial and Down’s syndrome retarded sample, surprise reactions indicative of color and number identity occurred with equal frequency, whereas in his nonretarded sample, color surprise was significantly more frequent than number surprise. (d) Abel (1941) found no significant decline with increasing MA in belief in immanent punishment in her institutionalized female retarded sample, a finding that differs from some earlier research with nonretarded children.

These four instances of no support for the similar sequence hypothesis are exceedingly minor. They may have resulted, as the authors of these studies generally suggested from idiosyncratic (or misinterpreted) properties of the tasks selected or from other measurement errors, and in some cases (e.g., Abel, 1941) they may reflect the suppressive influence of an idiosyncratic environment that merely delays the shift from one level of reasoning to another. Furthermore, in each of the four studies, the findings of no support were outnumbered by findings supporting the similar sequence hypothesis.

While noting the strong level of support that the reviewed evidence has yielded for the similar sequence hypothesis, we must also note that such cross-sectional and level-of-
difficulty evidence, even at its best, can support only indirect inference regarding the actual process of development. Considerably more direct and potent inference is possible when an investigator observes the same individuals at more than one point during the course of development, that is, in longitudinal fashion. Such research is often expensive and complex, and consequently it is relatively rare, particularly with mentally retarded persons. However, three longitudinal studies have some bearing on the similar sequence hypothesis. We now turn to these.

**Longitudinal Evidence**

*Development in the sensorimotor period.* One longitudinal investigation was designed by Wohlhueter and Sindberg (1975) as an extension of Woodward’s (1959) cross-sectional study of sensorimotor development in profoundly retarded persons (described earlier in this article). These investigators conducted monthly assessments of institutionalized 1-6-year-old profoundly, severely, and moderately retarded children (no IQs reported). The Piagetian object concept tasks used by Decarie (1965) were employed for 1 to 1½ years or until a child performed at the highest of the 10 substage levels for 2 consecutive monthly sessions. Of the principal sample of 49 children, 20 had progressed to the highest substage level by the end of the study; of the remaining 29, 10 showed a generally monotonic increase, and 9 seemed to be at a plateau, with object concept levels the same for most of the 12 or more sessions. Thus, 39 of the 49 subjects showed patterns harmonious with the pattern of object concept stages posited by Piaget (1955) and found by subsequent investigators using nonretarded samples. For the remaining 10 subjects, however, there was a variable developmental pattern in which substage levels appeared to rise and fall from session to session, ranging over as many as 3 or 4 substages during the 12 or more sessions.

In an effort to determine what characteristics might distinguish this group of atypical subjects, Wohlhueter and Sindberg (1975) examined medical histories and clinical findings for their sample; the distinguishing feature of the variable group was that the majority of subjects “were found to have EEG abnormalities, especially dysrhythmias or a history of seizures” (p. 516). This finding raises at least two possible interpretations with respect to the variable developmental pattern: (a) that individuals with brain anomalies associated with electroencephalogram (EEG) abnormalities may show atypical sequences of development with respect to the object concept and (b) that behavioral and attentional abnormalities in individuals with anomalous EEG patterns make accurate assessment of object concept substages difficult.

One other unusual pattern was noted by the investigators, namely, some children seemed to bypass or skip over some of the substages. This apparent skipping phenomenon has been noted in research with nonretarded children as well (see Uzgiris, Note 2). Although this is an interesting phenomenon, it is difficult to know how often skipped substages may actually have been traversed by the children in the intervals between experimental sessions. Furthermore, both the skipping phenomenon and the variability phenomenon might be better understood if we were able to rule out specific method effects, as could have been done if a nonretarded sample had been included in this study for comparative purposes. Nonetheless, the Wohlhueter-Sindberg investigation at least raises significant questions about the validity of the similar sequence hypothesis with respect to certain substages in the development of the object concept.

A longitudinal study reported by Cicchetti and Sroufe (1976), however, yields strong support for the similar sequence hypothesis within the sensorimotor period. The study focused on the relation between cognitive and affective development in home-reared Down’s syndrome infants during the period from 4-18 months of age. Sroufe and his colleagues (e.g., Sroufe & Wunsch, 1972) earlier demonstrated that among normal infants there is a developmental progression from mirth response to auditory and tactile stimulation that is physically intense or vigorous (e.g., tickling the baby’s chin or saying “Boom!”) to mirth responses to social and visual stimulation that is progressively more subtle and complex (e.g., the sight of mother sucking on baby’s bottle). At monthly intervals the infants sampled by Cicchetti and Sroufe were pre-
presented with 15 auditory and tactile items of the intense or vigorous type and 15 social and visual items of the more subtle and complex type. As in the earlier research with normal infants, the Down's syndrome infants laughed earliest in response to the auditory and tactile items and latest in response to "the more cognitively complicated social and visual items" (Cicchetti & Sroufe, 1976, p. 923). The responses of the Down's infants, of course, came months later (in CA) than did the corresponding responses of normal infants. Smile responses, a more sensitive index of positive affect in the Down's syndrome sample, showed the same pattern and revealed even more clearly than laughter responses the developmental decline in positive affect aroused by simpler auditory and tactile items as the infants matured beyond 13 months. In stressing the similarity of their findings with Down's infants to those with normal infants, Cicchetti and Sroufe (1976) pointed out that the laughter items were similarly ordered for both groups, "category by category and, in the main, item by item" (p. 923). Finally, to assess the merits of their claim that the affective responses they measured were closely related to cognitive development, Cicchetti and Sroufe calculated correlations of indices of affective expression (e.g., earliest laugh, total amount of smiling to all items, etc.) with the Bayley mental and motors scales and the Uzgiris-Hunt object permanence and operational causality scales. All 44 correlations were statistically significant.

Therefore, the Cicchetti-Sroufe investigation, unlike the Wohlhueter-Sindberg study, provides uniform support for the liberal version of the similar sequence hypothesis in which the hypothesis is applied to all retarded children regardless of etiology. The Cicchetti-Sroufe research deserves special attention because of its unusually careful methodology and its emphasis on the integrity of the developing infant. The research demonstrates a thoughtful means of assessing the Piagetian hypothesis that affective development and cognitive development are inter-dependent. In addition, it may help to point the way to tests of the similar sequence hypothesis in behavioral domains other than cognitive development as it has traditionally been construed.

Research on the preoperational-concrete operational transition—the Temple longitudinal study. The third longitudinal study reviewed here is by far the broadest in scope. In this ongoing investigation, Stephens and her colleagues (Stephens, 1974; Stephens, Mahaney, & McLaughlin, 1972; Stephens et al., 1974) have conducted biennial assessments of the performance of retarded and nonretarded persons on a variety of Piagetian tasks. The sample included 75 retarded subjects (IQs of 50-75) from special education classes and 75 nonretarded subjects (IQs of 90-110) from the same Philadelphia schools. In the first wave of testing, the age range in both subject groups was 6-18 years. Results from the first two waves of testing have now been reported and are discussed here within two content categories:

Moral judgment. The Temple battery included 11 measures designed to assess three aspects of moral judgment: (a) the relative weight assigned to intent versus consequences in judging the seriousness of a misdeed, (b) awareness of the injustice of punishing an entire group for the acts of only one or a few members, and (c) the ability to judge the relative fairness of various types of punishment including retributive and reciprocal justice. To determine whether judgment along these three dimensions follows the same developmental course in the retarded as in the nonretarded, Mahaney and Stephens (1974) examined changes in scores on the 11 component measures over the 2-year period from the first to the second wave of testing. They found that on 1 of the intent-versus-consequences measures the retarded group showed a nonsignificant decline in score (i.e., they made slightly less mature moral judgments according to the scoring criteria adopted by the authors) and that on 1 of the group punishment items nonretarded subjects showed a nonsignificant decline. On the 9 remaining items the direction of change was the same for both the retarded subjects and the nonretarded subjects; this similarity extended to 2 items on which both groups showed significant de-
clines in score, raising questions about the scoring of these particular items.

Inhelder (1943/1968), in a study described earlier, referred to certain "oscillations" in the reasoning of the retarded. It is of some interest to note that Mahaney and Stephens (who was the translator of the Inhelder book) reported oscillations, that is, instances when "the improvement which occurred in one area of moral judgment was not maintained when opinions were solicited on another, but similar, situation" (Mahaney & Stephens, p. 137), in both retarded and nonretarded subjects. There is some indication that such oscillations may have been somewhat more frequent in the retarded group. One line of evidence that suggests this possibility comes from Mahaney and Stephens's analysis of change scores for three separate age levels within both the retarded and the nonretarded samples. Of the 29 change scores reported for the nonretarded groups, 20 were increases (11 significant), 7 were decreases (3 significant), and 2 showed no change. Among the retarded subjects, 17 change scores were increases (7 significant), 11 were decreases (4 significant), and 1 involved no change. However, when retarded comparison groups were staggered in order to broaden the age difference involved in age group comparisons, for example, by comparing Phase 1 6–10-year-olds with Phase 2 12–16-year-olds, the only items showing a decrease with age were the two that showed a decrease in nonretarded subjects as well. Overall, the report by Mahaney and Stephens (1974) suggests that although growth in moral judgment concepts among the retarded may be "torporific and sporadic" (p. 141), the direction of development is the same for both retarded and nonretarded persons.

Conservation, classification, symbolic imagery, and formal operations in the Temple study. The Temple investigation (Stephens, 1974; Stephens et al., 1972; Stephens et al., 1974) also included 29 measures of cognitive development across four broad conceptual domains: (a) conservation (of substance, length, weight, continuous quantity, and volume, as well as term-to-term correspondence), (b) logic classification (class inclusion and class intersection, and relative thinking measured by the brothers–sisters and right–left tests), (c) operativity and symbolic imagery (tests involving imagined rotations of objects through space, transferring from two to three dimensions, and changing one's perspective on a stimulus), and (d) combinatory logic (Piaget & Inhelder's 1956, combination of liquids task). Explanations given by subjects on each of the 29 items were scored on a 9-point scale that took into account, among other things, the degree to which the subject wavered between a correct and an incorrect answer and the degree to which reversibility was shown. Stephens and McLaughlin (1974) reported on changes in scores on the 29 measures over the 2-year period separating the two waves of testing. They found that the nonretarded group showed improvement on all 29 measures, with 25 statistically significant; the retarded group also improved on all 29 measures, with 26 statistically significant. This finding indicates that the direction of development on these Piagetian reasoning tasks was similar in the retarded and nonretarded groups. In another report following the second wave of the Temple study (Stephens et al., 1972), the Piagetian reasoning tasks were rank ordered with respect to the MAs at which 50% of the subjects (in the retarded and nonretarded groups separately) made correct responses. As Stephens et al. indicated, the order of difficulty for both subject groups was generally consistent with previous findings that conservation of substance precedes conservation of weight.

Whether this apparent similarity in the direction of development is actually a function of an invariant, stagelike progression is thus far an open question because of the questionable nature of the measures themselves. In a critique of this portion of the Temple longitudinal study, Kohlberg (1974) maintained that Piagetian moral judgment measures used in the Temple study do not even warrant detailed longitudinal analysis, because

Piaget himself does not consider that his moral judgment measures yield genuine stages, nor do they pair up with his logical stages in ways compatible with his current thinking about cognitive stages . . . .

Empirical research confirms the fact that Piaget's moral stage measures do not meet the criteria of structural stages which his logical stages do meet. (p. 142)

This being the case, it is appropriate to be cautious about what one concludes with respect to the moral judgment portion of the Temple investigation.
and that conservation of weight precedes conservation of volume. In addition, although the ranks were rather crude because MA levels were listed in whole years and because not even the most ardent Piagetian would expect all 29 items to form an orderly developmental scale, it is interesting to note that our own calculations yielded a Spearman rho of .634 between the rank order given for the retarded group and that given for the nonretarded group.

The preceding data are consistent with the similar sequence hypothesis, as far as they have been taken, but they could be taken considerably further. With the one exception mentioned in the preceding paragraph, there has been no apparent effort thus far by the Temple investigators to check their findings against specific developmental stage sequences such as the horizontal decalages reported for nonretarded subjects in previous Piagetian research (Kohlberg & DeVries, 1971; Nassafat, 1963; Siegelman & Block, 1969; Smedslund, 1964; Uzgiris, 1968). Moreover, there has been a persistent inclination to report data in terms of group means, rather than in terms of the number of individuals (retarded and non-retarded) who show specific developmental patterns. This latter type of analysis is the unique province of longitudinal research and can only be approximated indirectly by scaling procedures in research of the nonlongitudinal variety. There is some indication (see Stephens, 1974) that efforts to profile individual performance changes over time and to cross-validate specific vertical and horizontal decalages will be forthcoming from the Temple investigators. Such efforts are needed if the investigators are to fully capitalize on the power of their longitudinal design.

Status of Evidence on the Similar Sequence Hypothesis

Only 1 of the 3 longitudinal studies reviewed—the Wohlhueter and Sindberg (1975) investigation of object concept substages—produced findings inconsistent with the similar sequence hypothesis. In that investigation a distinct subgroup of 10 (out of 49) children showed apparent atypical developmental sequences, and most of these children showed anomalous EEG patterns. This finding may indicate that brain wave anomalies can be associated with atypical developmental patterns. Alternatively, the EEG abnormalities may simply have been associated with attentional and other deficits that interfered with accurate assessment of substage levels in children whose actual development was consistent with the similar sequence hypothesis. The latter interpretation has special credence in the area of object concept assessment, in which procedures demand that the subject sustain attention to an object long enough to seek after it once it has been removed from the perceptual field.

Of the 28 nonlongitudinal studies reviewed, only 4 contained a finding inconsistent with the similar sequence hypothesis, and in each of these studies the inconsistent finding was relatively minor and was alone among a number of findings supporting the hypothesis. Furthermore, the questions raised generally concerned rather fine-grained steps or substages within horizontal decalages on which studies of nonretarded subjects alone have not always agreed.

These facts, plus the measurement problems inherent in these experimental procedures, make the degree of consistency in the findings of these 31 studies rather striking. Positive findings have now been reported in conceptual areas that include sensorimotor spatial concepts, object permanence, causality, imitation, affective responding, identity and equivalence conservation (of many properties), seriation, transitivity, moral reasoning, comparison processes (or gross, intensive, and extensive quantities), time, space, relative thinking, role taking, mental imagery, geometric concepts, and classification and class inclusion. For the 31 studies spanning this list of conceptual areas, the great preponderance of the evidence is consistent with the hypothesis that retarded and nonretarded persons traverse the same stages of development in the same order, differing only in the rate at which they progress and in the ultimate ceiling they attain. The hypothesis seems to be generally supported in studies of retarded individuals, regardless of etiology, with the possible exception of individuals suffering from pronounced EEG abnormalities.
Quality of the Evidence

Having said this, we believe it is important to comment on the quality of the available evidence and to offer suggestions for improving it. Cross-sectional data relevant to the similar sequence hypothesis have most often been presented in ways that provide only the weakest inferential power. A table displaying the percentage of subjects at each age level who pass each Piagetian item can yield only a rather faint glimmer of developmental sequence compared with the information generated when each child is classified with respect to specific pass-fail patterns, that is, with respect to response scale types. When such a scaling analysis is combined with calculation of scalogram summary statistics (e.g., Green, 1956; Guttman, 1950), the potential power of the nonlongitudinal design is more fully utilized.

Similarly, the bulk of the longitudinal data we found was presented only in terms of mean difference between experimental groups or changes in group means or percentages over time. As Hunt (1974) has noted, reporting only group summary statistics at Time 1 and Time 2 can mask the fact that some individuals progressed while others regressed over time. Thus, although it is useful to know that the Time 1 and Time 2 means differed in the same direction for retarded and nonretarded groups, such information is no substitute for an analysis of the number of individuals in each group showing specific developmental patterns over time. In both longitudinal and nonlongitudinal research aimed at testing the similar sequence hypothesis, it makes little sense to invest the time and energy necessary to gather potentially relevant data and then analyze the data in ways that fail to capitalize on their full potential.

Suggestions Toward Improved Research

These problems, and others to which we referred earlier in the text, suggest three principles that if widely adopted would substantially improve the quality of evidence on the similar sequence hypothesis.

Structuring Direct Comparisons

A problem with many of the studies reviewed in earlier sections is that their samples included only mentally retarded subjects. In those few instances in which findings of these studies disagree with findings of other studies sampling only nonretarded subjects, the discrepancies are difficult to interpret. This is because of uncertainty over whether the discrepancies reflect actual process differences between the retarded and the nonretarded or whether the differences in experimental methodology across studies are responsible. An obvious way to prevent such difficulties is to expose retarded and nonretarded children within a similar cognitive developmental range to precisely the same procedure by including both groups in the same study. To fail to do so is to risk uninterpretable findings.

Attending to Etiology

The marked heterogeneity of many of the mentally retarded samples described earlier suggests a somewhat opportunistic approach to subject selection or perhaps an approach in which the etiology of retardation is simply not regarded as an important factor. Yet, theoretical considerations discussed early in this article (see also, Weisz, 1976; Zigler, 1969, 1971) point to the need to give special attention to familial retarded children as opposed to those suffering from organic impairment or genetic disorder. Furthermore, Wohlhueter and Sindberg’s (1975) report of atypical development in a group of children with a high incidence of EEG anomalies suggests the potential importance of efforts to identify developmentally distinct subgroups within the nonfamilial population. Their analysis illustrates that subgroup analyses can be useful even when they are post hoc.

Promoting Uniformity

Finally, there is a clear need for increased uniformity across studies in the kinds of statistical analyses carried out and in the way statistics are reported. Toward this end, we suggest that every cross-sectional study addressing the similar sequence hypothesis should yield data bearing on the following threefold question:

1. Within the retarded and nonretarded groups do the task items form the same scale, and does this scale show high reproducibility
(à la Guttman, 1950) and a high index of consistency (à la Green, 1956)?

2. Do mean scale scores increase with level of cognitive maturity in each separate subject group (see Kohlberg, 1969)?

3. Do mean levels of success on each individual item increase with levels of cognitive maturity in each separate subject group (see Kohlberg, 1969)?

In longitudinal research, Questions 2 and 3 should also be asked in a way that only longitudinal investigation permits: Over the period spanned by the longitudinal study, in what percentage of individual subjects (from retarded and nonretarded groups) do scale scores and individual item scores (a) increase either smoothly or monotonically, (b) remain stable throughout, and (c) show at least some declines.

These recommended questions are designed to promote greater uniformity, and thus greater comparability, among studies addressing the similar sequence hypothesis. In opposition to such uniformity, one might argue that the degree of consistency in the findings of the numerous studies reviewed here is all the more impressive precisely because of the methodological diversity of the studies. There is some truth to this argument, but only in those cases in which findings support the similar sequence hypothesis. However, we have argued that even in those cases apparent group similarities in developmental sequence may result from a failure to ask the most probing questions of one's data. It seems clear from our review that evidence from the 31 studies currently available offers rather consistent support for the similar sequence hypothesis; it also seems likely that the best evidence has yet to be gathered.

Reference Notes


References


Braine, M. D. S., & Shanks, B. The conservation of a shape property and a proposal about the origins of the conservations. Canadian Journal of Psychology, 1965, 19, 197–207. (a)


Lane, E. B., & Kinder, E. F. Relativism in the thinking of subnormal subjects as measured by certain of Piaget's tests. *Journal of Genetic Psychology*, 1939, 54, 107–118.


McManis, D. Comparison of gross, intensive, and extensive quantities by retardates. *Journal of Genetic Psychology*, 1969, 115, 229–236. (a)

McManis, D. Comparisons of gross, intensive, and extensive quantities by normals and retardates. *Child Development*, 1969, 30, 237–244. (b)


McManis, D. Conservation of mass, weight, and volume by normal and retarded children. *American Journal of Mental Deficiency*, 1969, 73, 762–767. (d)

McManis, D. Conservation and transitivity of weight and length by normals and retardates. *Developmental Psychology*, 1969, 1, 373–382. (e)


Woodward, M. The behavior of idiots interpreted by


Received April 3, 1978