Psychological Bulletin
1986, Vol. 100, No. 2, 157-175

0033-2909/86/$00.75

Performance of Retarded and Nonretarded Persons on Information-Processing Tasks: Further Tests of the Similar Structure Hypothesis

Bahr Weiss, John R. Weisz, and Richard Bromfield
University of North Carolina at Chapel Hill

In this article we review information-processing studies applicable to the similar structure hypothesis. This hypothesis holds that when nonorganically impaired retarded and nonretarded persons are similar in general developmental level, they are also similar in the cognitive processes and concepts by which they reason. Results of this review were strikingly different from those of an earlier review focusing on Piagetian investigations. A meta-analysis revealed that the performance of retarded groups was significantly inferior to that of their nonretarded psychometric mental age peers. We also found a significant relation between the mental age of the subjects and the probability of finding a significant difference between retarded and nonretarded subjects. A number of alternative explanations are offered for these findings.

When a mentally retarded child and a younger nonretarded child are at the same level of cognitive development, say at the same mental age, how similar are the processes by which the two children reason? For the past two decades this question has been a central issue in a sometimes heated debate—one with significant implications for the study of both mental retardation specifically and cognitive development generally. This debate is called the developmental versus difference controversy.

On one side of the controversy is the developmental position, advanced by Zigler (1969) and elaborated by Weisz, Yeates, and Zigler (1982). This position, which applies only to individuals not suffering from organic impairment, holds that retarded and nonretarded people pass through cognitive developmental stages (e.g., the stages described by Piaget, 1964, 1970) in an identical order, but differ in rate and upper limit of development. Retarded children are said to traverse the stages more slowly and attain a lower developmental ceiling than nonretarded children. According to the developmental position, retarded and nonretarded children who are equated for level of development (most often operationally defined as psychometric mental age [MA]) will not differ in the cognitive processes by which they reason.

On the other side of the controversy is the difference position, held by several theorists. One tenet of this position is that the cognitive development of the retarded person differs from that of the nonretarded person in more ways than mere differences in developmental rate and upper limit. According to this position, retarded and nonretarded children, even when equated for level of development, will differ in the cognitive processes they use in reasoning. Spitz (1976), for example, has argued that retarded individuals are particularly deficient in cognitive abilities that developed relatively recently in the history of human evolution. Specifically, he posits that retarded individuals, even when matched on MA, will show a marked deficiency on tests requiring foresight and logical analysis, and to a lesser extent, on tests requiring verbal abstraction and conceptual ability.

As noted in earlier analyses (Weisz & Yeates, 1981; Weisz & Zigler, 1979), the developmental versus difference debate actually involves two separate hypotheses. One hypothesis is that retarded and nonretarded children pass through cognitive developmental stages in the same order. This has been labeled the *similar sequence hypothesis,* and evidence germane to it was examined in an earlier article in this journal (Weisz & Zigler, 1979). Most of that evidence, derived from Piagetian research, supported the hypothesis and thus the developmental position.

The second major tenet of the developmental position has been labeled the *similar structure hypothesis* (Weisz & Yeates, 1981). This hypothesis is based on the view that retarded and nonretarded children at similar levels of cognitive development are similar with respect to their cognitive structure. The term "cognitive structure" represents the organization of thinking and learning processes, such as, the logico-mathematical patterns that underlie human understanding, reasoning, and information acquisition and use. Because cognitive structure cannot be directly observed, it is inferred from such overtly observable measures as effectiveness of problem solving and speed of learning.[1] In cases where such measures are used with retarded and

The authors thank the following people for their assistance and helpful comments on drafts of this article: Alfred Baumeister, Jake Burak, David Caruso, J. P. Das, Geraldine Dawson, Bob Hodap, Betty J. House, Aviva Wasserman, Keith Owen Yeates, Edward Zigler, and two anonymous reviewers. We also thank the following persons who served as experts in the rating of the tasks regarding hemispheric processing: Richard Davidson, John Farber, Marcel Kinsbourne, Jerry Levy, Muriel Lezak, and Stephen Warrenburg. Finally, we thank the investigators, too numerous to name here, who responded to our inquires for information about their studies.

Correspondence concerning this article should be addressed to John R. Weisz, Department of Psychology, The University of North Carolina, Davie Hall 013 A, Chapel Hill, North Carolina 27514.

[1] As Baumeister (1967) has pointed out, equivalent scores on psychological tests do not necessarily indicate equivalent underlying processes. Certain studies we reviewed (e.g., Weisz & Achenbach, 1975) were specifically designed to explore processes underlying cognitive perfor-

nonretarded children of similar MA, and appropriate experimental control has been exercised (as described in the next section), the similar structure hypothesis holds that the two groups will not differ reliably.

In sharp contrast, difference theorists such as Milgram (1973) have argued that retarded children will prove inferior to their nonretarded MA peers in a number of reasoning processes. Arguing that IQ rather than MA is an index of a person's problem-solving ability, Milgram predicts that "given tasks without a ceiling or floor effect for the general MA of the subjects being used, the retarded will dependably demonstrate an equal-MA deficit" (1973, p. 209).

In arguing that retarded people will show inferior performance, Milgram takes what might be called the conventional difference position. A more unconventional position is taken by Kohlberg (1968). He believes that the development of reasoning skills requires massive doses of "general experience"; consequently, Kohlberg argues, retarded children should be cognitively more advanced than their higher IQ MA peers, because they have lived longer and have thus acquired a more extensive base of experience.

Weisz and Yeates (1981) reviewed research comparing the performance of retarded and nonretarded groups selected to be similar in general level of cognitive development. Their review focused on 30 studies, all involving Piagetian tasks, which together comprised 104 comparisons between retarded and nonretarded groups. In all of the studies, similarity of developmental level was approximated by selecting retarded and nonretarded groups who were matched for MA. Data from these studies were interpreted in the context of the developmental-difference controversy, and according to whether the studies controlled for organic impairment in the retarded sample. (See the section on *Organicity.*)

An overall analysis of the 104 separate tests (including those with samples not screened for organicity) found that 4% supported the unconventional difference position—the view (articulated by Kohlberg, 1968) that retarded subjects would perform at higher levels than their nonretarded MA peers. Another 24% of the comparisons supported the conventional difference position (e.g., Milgram, 1973)—that is, the view that nonretarded subjects would outperform their retarded MA peers. Some 72% of the comparisons supported the developmental hypothesis (Zigler, 1969)—that is, they reported no reliable difference between the performance of retarded and nonretarded subjects.

Although the majority of individual studies supported the similar structure hypothesis, Weisz and Yeates argued that the overall pattern could not be taken as definitive support for it. The similar structure hypothesis, they noted, is essentially a null hypothesis; it generates a prediction that group differences will occur only as a result of random variability; that is, across studies, significant results will be found only at a chance level. Thus, for the hypothesis to be supported, the number of comparisons showing significant differences between groups should be less than or equal to the total number of comparisons reviewed times the α level. Hence, when studies not explicitly controlling for organicity were included in the analysis, results favored rejection of the developmental hypothesis. However, when only studies that excluded organically impaired subjects were analyzed, this was not the case. Of the 33 comparisons involved in these studies, only 3 (9%) contradicted the similar structure hypothesis.

This support for the similar structure hypothesis appears strong but is limited in scope, as it was derived entirely from studies using Piagetian reasoning tasks. Noting this, Weisz and Yeates (1981) ended their review with a call for an examination of the literature using non-Piagetian comparisons of the cognitive behavior of retarded and nonretarded individuals. That body of literature is the focus of the present review.

Literature Search Procedures, Inclusion Criteria, and Rationale

Initially, we performed two computer searches to identify studies comparing retarded and nonretarded individuals on non-Piagetian cognitive tasks. Next, we searched all volumes from 1960–1983 of the *American Journal of Mental Deficiency, Journal of Mental Deficiency Research, International Review of Research in Mental Retardation,* and *Psychological Abstracts* to identify relevant articles overlooked in the computer search. Finally, in order to ensure comprehensiveness, we inspected the bibliographies of all articles, and any articles in them that appeared relevant were screened.

To be included in our review, studies had to meet three criteria. The retarded samples involved had to be (a) matched on MA with a nonretarded group, (b) screened to exclude organically impaired individuals, and (c) noninstitutionalized. (Further details of these criteria follow.)

If a particular study did not meet one or more of our criteria, we excluded it from further consideration.[2] If sufficient information to make a decision was not available in the published article, we sent a letter requesting clarification to the author(s). If no reply came, or if it did not indicate that the study met our criteria, we excluded the article and listed it in the supplementary bibliography. We mailed 87 letters, and received 68 replies (78%). Nine percent of the studies for which we received replies were included in our review (8 studies). Of the 227 studies in our initial pool, 24 (11%) met the criteria for inclusion.

Matching for Developmental Level

The similar structure hypothesis posits that retarded and nonretarded individuals of similar developmental level will show similar cognitive processes. Similar developmental level has generally been operationalized through matching subject groups on MA, as assessed by standardized intelligence tests. Most investigators appear to believe that groups should be matched on measures that tap a relatively broad array of cognitive functions and that are well standardized. At present, IQ tests that yield an MA appear to meet this dual standard, and

mance (e.g., use of hypotheses to solve discrimination learning problems). For other studies, Baumeister's criticism may be more relevant.

[2] A list of the studies we inspected that did not meet our inclusion criteria is available from the second author. Please include $5.00 for printing and postage fees.

most studies designed to compare retarded and nonretarded individuals of equivalent developmental level have used standardized IQ tests for this purpose. Our sample of studies included only those comparing retarded and MA-matched nonretarded subjects. Note, however, that other measures such as adaptive functioning scales have been proposed as a basis for matching (Brooks & Baumeister, 1977).

A number of studies reviewed for potential inclusion did not include IQ or MA assessments of nonretarded subjects. Instead, the mean IQ of these subjects was assumed to approximate 100, with MAs equal to chronological ages (CAs). Retarded subjects were then matched by equating the nonretarded subjects' CAs with the retarded subjects' MAs. We judged that the variability of the average "normal" IQ in small samples made this procedure unacceptable for our purposes. Therefore, articles that assumed MA = CA for nonretarded subjects were excluded.

Also excluded were studies that did not include procedures for MA matching, but relied instead on statistical control of MA differences through analysis of covariance. As Kappauf (1976) points out, when there are MA differences between groups that differ on IQ, MA and IQ will be correlated. If one statistically removes the effect of MA differences, one also partially removes the effect of IQ, resulting in a less powerful test and potentially misleading results.

IQs of the retarded groups. Ideally, all members of a retarded sample should have IQs below 70 (i.e., they should have IQs in the retarded range); the mean of the retarded group would thus be well below 70. Yet a number of studies that met our other criteria used a retarded group with a mean IQ somewhat greater than 70. Excluding these studies would have reduced the number of comparisons available, to the extent that it might not be possible to detect trends in the data. For this reason, studies in which the mean IQ of the retarded group was less than 75 were included. To determine whether setting a relatively high IQ cutoff (which would, in principle, favor the developmental hypothesis) affected our results, we performed a 2 × 2 chi-square analysis. In this analysis, one dimension divided the retarded/nonretarded comparisons into those using retarded groups with IQ < 70 and those with 70 < IQ < 75, whereas the other dimension divided the comparisons based on the significance or nonsignificance (and in the former case direction) of their results. The resulting chi-square value was nonsignificant, $\chi^2(2, n = 59) = 2.76, p > .25$, indicating that IQ level across the retarded groups was not significantly related to outcome.

Organicity

The current developmental position (e.g., Zigler & Balla, 1982) specifically excludes retarded individuals with organic etiologies from the similar structure and similar sequence hypotheses. As Zigler and Balla point out, retarded individuals with known brain damage do indeed possess a "difference" relative to nonretarded individuals. Moreover, the Weisz and Yeates (1981) review yielded significantly different results depending on whether organically impaired subjects were included or excluded from the analyses. Thus, we decided that to be included in this review, a study must contain at least one retarded group that was not organically impaired.

A number of important questions have been raised about the exclusion of organically impaired subjects from research on the developmental hypothesis. For instance, A. A. Baumeister (personal communication, August 4, 1984) has noted that in some instances, the "diagnosis" of organicity is made based on impaired performance on psychological tests. Yet the exclusion of such "organically" impaired subjects might bias results in favor of the developmental hypothesis, in that it screens out subjects who perform poorly on tests. Further, we recognize that the available methods of diagnosing organicity are relatively crude, and that excluding organically impaired individuals from the subject sample with absolute certainty is not possible. Given these uncertainties, we decided that any one of the following criteria was sufficient evidence, for our purposes, that the retarded sample did not contain inordinate numbers of organically impaired individuals.

1. The article stated that subjects were free from or showed no indications of organic impairment, neurological involvement, or brain damage.
2. The article stated that the subjects were "cultural-familial" or "familial" retarded.
3. The author(s) of the article responded to our inquiry, indicating that organically impaired individuals were not included in the subject sample.

Clearly, studies differ widely in the amount of effort directed toward screening for etiology. Hetherington and Banta (1962), for example, required that to be considered familially retarded, subjects have at least one other family member in an institution or special class for retarded individuals, and that there be no known organic brain damage. Few investigators have been this explicit or careful. In choosing our organicity criteria, we tried to be as conservative as possible, without excessively limiting the number of studies passing the criteria.

Institutionalization and Noncognitive Factors

The developmental view does not specifically exclude institutionalized individuals from its hypotheses. However, certain factors (e.g., motivation, expectancy of failure) resulting from differential life experiences can have an important impact on individuals' performance (Zigler & Balla, 1982). For example, a number of studies have found that institutionalization apparently has a significant impact on the behavior of retarded individuals (see e.g., Zigler & Balla, 1977); these effects are diverse and difficult to predict. In order to minimize the effects of "extracognitive" factors on retarded–nonretarded comparisons, we decided to exclude studies with only institutionalized retarded subjects from the present review.

There is almost certainly a multitude of noncognitive factors on which retarded and nonretarded groups differ, even when samples are confined to noninstitutionalized individuals. However, few investigators have tried explicitly to control for such factors (e.g., motivation and expectancy of success). Thus, we were not able to select studies on the basis of their control over noncognitive factors per se. Given the current state of research, excluding studies that did not include a noninstitutionalized retarded group was the most conservative approach feasible.

We regard the criteria detailed previously as appropriately stringent, given our objectives. We recognize, though, that different investigators with different goals might prefer less stringent but equally valid criteria.

Potential Sources of Bias in the Literature

This review's results should be interpreted in the light of two potential sources of bias. One, called the "file drawer problem" (Rosenthal, 1979), grows out of the apparent tendency of journal reviewers and editors to favor publishing studies that show statistically significant group differences. An extreme view is that studies published in the journals represent the 5% of all studies performed that showed significant effects by chance ($\alpha = .05$), whereas the 95% of the studies that failed to find significant results remain in university file drawers unpublished. If the file drawer problem exists to even a small degree in the areas spanned by this review, there would be an increased probability that studies whose results favored the difference position (i.e., studies showing significant group differences) would appear in the published literature. This might have been particularly true of studies submitted for publication prior to Zigler's (1969) articulation of the developmental position, which provided a theoretical framework for explaining—and thus a justification for publishing—"no difference" findings.[3]

We should note, though, that studies have often included groups other than the standard retarded and MA-matched control groups (e.g., CA control groups, institutionalized retarded subjects). Such group comparisons might be more likely to show significant differences that would provide sufficient justification for publication, and thus save the study from the oblivion of the file cabinet.

The similar structure hypothesis is favored because it is essentially a null hypothesis; that is, it involves a prediction of no difference between groups. A failure to reject the null hypothesis may result from modest differences between groups, small sample size, and/or large variability on the dependent measure(s). Following the reasoning of Weisz and Yeates (1981), for studies reporting no effect for IQ group, we included information about other significant effects between the groups (e.g., differences due to MA level) to give some indication of whether the finding of no difference may have been due to small samples and/or high variability.

Organization of Studies in This Review

We might have organized the studies in this review around a particular theoretical model of information processing, adopted a priori. However, we judged that such an a priori assumption might blind us to unexpected patterns in the data. So instead, we chose a relatively atheoretical approach, arranging studies into broad categories. We do not claim that these categories are in any absolute sense "correct." Further, the results of some studies bear on more than one cognitive area; in a few cases, the results of different facets of studies are discussed in different sections. Our overall goal was to provide an organizational framework for an initial examination of the data—a framework that may have heuristic value.

One final organizational caveat is in order. Many of the studies reviewed, although quite relevant to the similar structure hypothesis, were not originally designed as tests of this hypothesis. Portions of the results of these studies that are not directly relevant to the topic at hand are not detailed here. We have also generally omitted discussing secondary or repeated analyses that were not independent of the primary analyses.

Meta-Analysis

Two forms of data analysis were used. First, we examined our findings of the various retarded–nonretarded group comparisons from a categorical standpoint, using a chi-square analysis. Findings were organized according to direction and overall significance of group differences. This facilitated a direct comparison with the results of Weisz and Yeates (1981), who used a chi-square approach. In addition, we performed a meta-analysis, pooling data of the various studies in a way designed to maximize use of information and statistical power.

Meta-analyses of psychotherapy outcome studies frequently have used a mean effect-size procedure (e.g., Smith & Glass, 1977; Steinbrueck, Maxwell, & Howard, 1983). Though such a procedure is very useful when available data permit it, a number of the studies we reviewed did not include the means and standard deviations necessary for computing effect sizes. Consequently, we chose the mean z score procedure suggested by Rosenthal (1978), as it relies on F and t values rather than means and standard deviations. As Rosenthal and Rubin (1979, 1982) have pointed out, there is in general a "substantial correlation between level of significance and magnitude of effect" (1979, p. 1165), and the comparison of ps across studies is roughly equivalent to the comparison of effect sizes across studies, as long as differences between sample sizes are not extreme. Finally, we felt that the mean Z score procedure was appropriate because, unlike psychotherapy analyses in which the *size* of group differences is all important, the present analysis was concerned with whether there were overall significant group differences.

Therefore, following Rosenthal (1978) we computed the Z score equivalent for each test of group differences. In addition, we coded the MA, CA, and IQ of the retarded group, whether the experimental procedure included a manipulation designed to equate the motivation of the subjects, and whether the study met our initial criteria for organicity (i.e., did not require a letter). We discuss the rationale for choosing these variables, and the coding procedures used later in the article.

Review of Evidence

Intelligence Tests: IQ Profiles

Whether retarded and nonretarded children show similar subtest patterns on IQ tests is a question of particular importance to this review: All the studies we review used IQ tests for

---

[3] As a rough test of this hypothesis, we tested whether studies published subsequent to 1970 showed lower significance levels than studies published prior to 1970, which would support the view that there had been a (greater) bias in the journal selection process prior to 1970. In fact, however, the effect was the opposite of this prediction: There was a marginally significant trend ($p = .10$) for the more recent studies to show smaller significance levels.

matching on developmental level. Although matching on MA insures that the various groups have averaged approximately the same number of total points on the tests involved, it does not guarantee that the groups have similar subtest or "profile" patterns. Thus, we begin our review with the question of profile patterns.

Groff and Linden (1982). Though a number of studies have compared the intelligence test profiles of retarded and nonretarded children (e.g., Achenbach, 1970, 1971; Magaret & Thompson, 1950; Thompson & Magaret, 1947), we found only one such study satisfying our subject selection criteria. In this study, Groff and Linden (1982) compared the Wechsler Intelligence Scale for Children–Revised (WISC-R) factor profiles of retarded and nonretarded children. Included in the profiles were the Verbal Comprehension, Perceptual Organization, and Freedom From Distractibility factors that have been identified previously in a number of factor analytic studies of the WISC–R (e.g., Kaufman, 1975). A test of the Groups × Factor Score interaction indicated that the shapes of the retarded and nonretarded childrens' factor profiles were not significantly different.

Intelligence Tests: Progressive Matrices

Nesbit and Chambers (1976). In a study focusing on spatial ability, Nesbit and Chambers (1976) found the performance of nonretarded children superior to that of retarded children on Raven's Coloured Progressive Matrices. They concluded that the spatial abilities of the retarded children in their sample were inferior to those of the MA-matched nonretarded subjects, despite an attempt to control for motivational differences between the groups. They suggested that the retarded children's poor performance may have resulted from their cognitive style—a style that appears to have been less analytical, and more field dependent.

Tests of Memory

Fagan (1968). Interest in the memory processes of retarded persons dates at least back to Galton (1883) and has produced a number of studies satisfying our criteria. In one of the earliest of these, Fagan (1968) compared the acquisition and retention components of short-term memory for digit strings in retarded and nonretarded children. Both immediate and delayed recall were tested. During delayed trials, children experienced either low retroactive interference (white noise) or high retroactive interference (repeating color words).

Under all conditions, the retarded subjects' memory acquisition (i.e., immediate recall) scores were significantly lower than those of the nonretarded group. However, under conditions of unordered recall the two groups did not differ significantly in retention—that is, the difference between acquisition and delayed recall. Finally, the ordered recall retention of the retarded group was significantly superior to that of the nonretarded group ($p < .05$) under the low-interference condition, and marginally superior ($p < .10$) under the high-interference condition. In sum, this study suggests an acquisition deficit for the retarded group, equivalence between retarded and nonretarded groups in unordered-recall retention, and a trend toward superiority in ordered-recall retention for the retarded group.

Brown (1974). In another study using digit strings, Brown (1974) assessed the effects of recall order. On half the trials, subjects recalled the digits in the order presented, on the other half, they recalled the digits backward. Subjects were involved in a distraction task during the delay interval. Dependent measures were total recall (independent of order) and ordered recall.

Brown found no significant main or interactive effect for IQ, although the main effect for IQ on ordered recall did approach significance ($p < .10$). There were significant effects, though, for length of delay, serial position, and several interactions not involving IQ.

Winters and Attlee (1974). Winters and Attlee (1974) investigated the von Restorff effect—that is, the improved recall for an item that differs from others along a salient dimension (e.g., a single high-volume item embedded in a series of low-volume items).

Retarded and nonretarded subjects listened to a series of tape-recorded nouns, then recalled as many items as possible. During one session, the volume of all the words was equal; during another session, the volume of the sixth word was louder than the others. Overall, the nonretarded subjects recalled more stimulus words than the retarded subjects. Both groups showed a marked von Restorff effect, recalling the high-volume word especially well. Finally, there were indications that the two groups used somewhat different recall strategies. The retarded group tended to start recall with the most salient item; the nonretarded group more often recalled the list in serial order.

Hornstein and Mosley (1979). In an iconic memory study Hornstein and Mosley (1979) exposed retarded and nonretarded subjects to slides of six Chinese characters arranged in a circle. After a delay, an indicator pointed toward one of the characters, or a circle surrounded one of the characters. Then another Chinese character appeared, and the subject indicated whether this character was the character indicated in the previous slide.

The retarded subjects' performance did not differ significantly from that of the MA controls. However, the IQ factor did enter into an interaction. Although the MA controls were equally accurate regardless of whether the correct answer was "yes" or "no," the retarded subjects were more accurate when the correct answer was "yes." Evidently, this significant Group × Response Type interaction reflected the use of different response strategies by the two groups, rather than basic differences in sensory processing. The retarded subjects were apparently trying to maximize the number of hits by frequently saying "yes," at the risk of false alarms.

Mosley (1980). In another iconic memory study, Mosley (1980) tachistoscopically presented retarded adults and nonretarded MA controls with either familiar stimuli (letters) or unfamiliar stimuli (Chinese characters); a masking stimulus (a pattern of crisscrossed diagonal lines superimposed on the index stimulus) was presented on random trials. After seeing the index stimulus, subjects were asked to name the letter, or pick the correct Chinese character from an array of four.

Averaged across length of masking-stimulus onset delay, the retarded group performed significantly worse than the nonretarded group when unfamiliar stimuli were used, but not when familiar stimuli were used. However, for both familiar and unfamiliar stimuli under the minimal or maximal masking delay, the retarded group was as able as the MA controls to correctly identify the stimuli. On the other hand, under conditions of intermediate masking delay the retarded group performed more poorly, and this effect was more pronounced when the stimuli were unfamiliar. One must consider the possibility that this pattern of findings reflects ceiling and floor effects.

Mosley (1981). Mosley (1981) attempted to determine whether the "readout process" in iconic memory (the transfer of information out of iconic memory for further processing) is equivalent in retarded and nonretarded individuals. Subjects were tachistoscopically presented with various sized linear arrays of letters. Averaged across different masking-stimulus onset delays, the retarded group recalled significantly fewer of the letters than the MA controls. There were also several significant interactions involving group: Essentially, the shorter the masking delay interval and the greater the information load (i.e., array size), the more the performance of the retarded subjects deteriorated relative to the MA controls. Mosley interpreted these findings as indicating that retarded individuals suffered from an iconic memory deficit that resulted from slower iconic memory-store readout.

Paired-Associate Learning

Blue (1970). Blue (1970) assessed the effect of IQ and sensory mode of input on paired-associate learning. Children were required to learn an association between the index stimulus, a black and white nonsense design, and the response stimulus, a color. The color was presented either visually, auditorially, or via both modalities. Averaged across conditions, the MA controls reached the learning criterion significantly faster than the retarded subjects. No interaction was reported. However, individual cell means (p. 530) indicate that with auditory presentation, the retarded subjects did marginally better than the MA controls.

Winters, Attlee, and Harvey (1974). Winters et al. (1974) assessed the effects of method of presentation and pretraining on paired-associate learning, using a paired-associate list of eight pairs of pictures of common objects (e.g., a hat and book, a pencil and hammer). They found that the nonretarded group required significantly fewer trials to reach the criterion than an institutionalized retarded group. The performance of the noninstitutionalized retarded group was intermediate, but not significantly different from either group. Because of the possibility of a ceiling effect, the data were reanalyzed focusing on only the number of items paired correctly during the first 10 trials. Although several small interaction effects involved the institutionalized group, once again the nonretarded children performed better overall than the institutionalized retarded group, but not better than the noninstitutionalized retarded group. Further, neither the Group × Training nor the Group × Method of Presentation interactions were significant.

Input Organization

Some researchers have suggested that a key problem of mentally retarded individuals is their inadequate organization of input for subsequent cognitive processing. One investigator examined this possibility in a way that met our inclusion criteria.

August (1980). The effect of input organization on memory was studied by August (1980). Subjects were given a pack of file cards with a single noun printed on each. Half the subjects received strongly related nouns in six taxonomic categories (e.g., vehicles, vegetables); the other half received groups of weakly related nouns (e.g., train-wheel-road-ship). Subjects were told to read each word aloud, place the card into one of eight different boxes, and when finished, say aloud all the words they could remember. Subjects received one of four sets of sorting instructions, differing in the extent to which sorting according to the experimenter-defined categories was encouraged.

Overall, retarded subjects correctly recalled fewer of the nouns. There was also a significant Group × Stimulus Relations × Presentation Method interaction; the nonretarded group performed particularly well relative to the retarded group when weakly related stimuli were used, as compared with when strongly related stimuli were used. Nonretarded subjects also recalled words from more categories, worked faster, were more consistent in sorting words into categories they created under two of the instruction sets, and created categories more similar to those devised by the norm group when sorting weakly related stimuli. Recall organization also was more consistently related to input categorization for the nonretarded subjects. Thus, for all six dependent variables reported, the nonretarded subjects showed superior performance.

Although the results of the preceding studies have not been completely consistent, they certainly suggest the possibility of a performance deficit among retarded individuals in the area of memory. Of the 21 comparisons between retarded and MA-matched nonretarded subjects, 67% found that the nonretarded subjects performed significantly better. The somewhat contradictory results of some studies makes it difficult to specify the exact nature of this deficit, however.

Selective Attention

The ability to selectively attend to stimuli is critical to success in many social, academic, and vocational pursuits. Crosby and Blatt (1968) have suggested that mentally retarded individuals are less able to focus their attention than nonretarded individuals, even when matched for MA. We found one study that tests this notion.

Crosby (1972). Crosby (1972) assessed retarded and nonretarded children's maintenance of selective attention in the face of distracting stimuli using the Continuous Performance Test (CPT; Rosvold, Mirsky, Sarason, Bransome, & Beck, 1956). Subjects were presented visually a series of *X*s and *A*s. During the *X* task series, subjects were to respond to *X*s but not *A*s, whereas the *AX* task series required a response to *X*s only if they were preceded by an *A*. Four distraction conditions were used: no distraction, auditory, visual, or auditory and visual distraction. Crosby's institutionalized, organically impaired retarded group was least able to sustain attention, whereas the noninstitutionalized, nonorganically retarded subjects were most able to sustain attention; the performance of an institutionalized, nonorganically impaired retarded group and a nonretarded group was intermediate. There was a very wide range of distractibility among the retarded subjects, but on average their performance on the CPT was no more disrupted by the auditory and/or visual distractions than was the performance of the nonretarded subjects.

Discrimination Learning and Learning Set

Several investigators have explored the possibility that information-processing deficits will be found among the mentally retarded on tasks involving discrimination learning or formation of learning set. In this section we examine their work.

Libkuman (1972). Libkuman (1972) assessed the effect of stimulus familiarity on verbal discrimination learning. Half of his subjects read a word list later used in a verbal discrimination task; the other subjects read a similar list that was not used. He found that nonretarded subjects required significantly fewer trials to reach the learning criterion than the retarded subjects. Though there was a significant main effect for stimulus familiarity, the interaction with IQ was not significant.

Harter, Brown, and Zigler (1971). Harter et al. (1971) assessed the effect of task difficulty on discrimination learning. Subjects tried to solve either a simple three-choice picture discrimination task or a more complicated "oddity" discrimination task. Children also were placed in either a social reinforcement condition or a no-reinforcement condition. Subjects who reached the learning criterion were asked to explain how they succeeded.

When number of correct responses was the dependent variable, the main effect for group was highly significant ($p < .001$); nonretarded subjects made more correct responses than either an institutionalized or a noninstitutionalized retarded group. Of the subjects who reached the learning criterion, significantly more of the nonretarded subjects versus either of the nonretarded groups were able to verbalize the rule. Secondary analyses indicated that although reinforcement condition had no effect on nonretarded subjects for either task, noninstitutionalized subjects performed better on the simple three-choice task under the social reinforcement condition, whereas institutionalized subjects performed better under the standard condition. If one accepts Balla and Zigler's (1982) conjecture that institutionalized retarded individuals have suffered social deprivation, this latter finding is congruent with Harter's (1967) suggestion that subjects may perform less efficiently under the social condition because they are focusing on social aspects of the situation, rather than on the task itself.

Harter (1965). In this study subjects tried to learn under which of two differently shaped Styrofoam blocks a marble was hidden. An analysis of backward learning curves (Hayes, 1953) revealed that the retarded and nonretarded groups showed similar patterns of learning-set acquisition: A relatively long period of poor performance followed by a rapid acceleration to a high level of performance as the set was acquired. Despite this similarity, highly significant main effects ($p < .001$) were found for both IQ and MA: The higher the IQ or the MA, the quicker subjects reached criterion. The ability to verbalize the solution increased with MA, but retarded children were as able to articulate their reason for choosing the correct block as their MA-matched peers.

Harter (1967). In this study of motivational factors and learning set acquisition, Harter (1967) used a methodology similar to the preceding study, with two exceptions. An institutionalized retarded group was included, and half the subjects received social reinforcement during testing. Once again, though the acquistion process (assessed via backward learning curves) appeared similar for all groups, the nonretarded subjects reached the criterion quicker than their MA-matched retarded peers. Within the retarded groups, the noninstitutionalized subjects reached criterion before the institutionalized subjects. A Group × Condition interaction revealed that nonretarded subjects performed best in the social condition, although the retarded subjects performed best without social reinforcement. Harter hypothesized that the retarded subjects took longer to reach the criterion under the social reinforcement condition because their attention was focused on interacting with the experimenter, rather than on the task.

Richman, Adams, Nida, and Richman (1978). Richman et al. (1978) investigated complex aspects of set discrimination learning, including the ability to make intra- and extradimensional shifts. Children were presented with variously shaped and colored pieces of Styrofoam. For half the children, the correct stimulus dimension was color, for the other half, shape. Three different sets of instructions were used. Children in the conventional group were told to press a button indicating which of the two stimuli they thought was correct. Children in the imagery group were told to think of the Styrofoam blocks as people, one of whom lived in a gray Styrofoam "house" in the middle of the test apparatus; they were told to press the button to indicate which "person" they thought lived in the house. In the subject-manipulation group children were given the same instructions as the imagery group, except they were also told to pick up the person they thought lived in the house, and if their choice was correct, to place the person inside the house. On reaching the learning criterion, the correct stimulus was shifted either intradimensionally (e.g., from red to brown) or extradimensionally (e.g., from red to triangle). After this shift, subjects were again run to criterion.

There were main and interactive effects for Instruction and Dimension; color discriminations were easier than form discriminations, and when the correct dimension was form, the subject-manipulation group made fewer errors to criterion than the other groups. Retarded and nonretarded children did not differ significantly in how quickly they learned the initial discrimination. However, when intra- and extradimensional shifts were made, the nonretarded subjects learned the new correct stimulus faster than the retarded subjects. This finding might seem to suggest a kind of cognitive "rigidity" in retarded individuals (cf. Kounin, 1941). However, secondary analyses indicated that a subgroup of the retarded sample showed excessive flexibility; they appeared to have been distracted by the novelty of the shift stimuli. Thus, in this study at least, retarded individuals showed a high degree of variability in how rigidly (or flexibly) they responded to changing stimuli.

The trend in this section appears clear: The bulk of the evidence shows retarded individuals performing less well on discrimination learning tests than their MA peers. Of the six comparisons of discrimination performance, five favored the nonretarded children.

Incidental Learning

Various investigators (e.g., Denny, 1966) have suggested that retarded individuals suffer from a deficit in incidental learning.

In this section we review several studies that have tested this hypothesis.

Libkuman (1972). While investigating verbal discrimination learning (see the section on *Discrimination Learning*), Libkuman (1972) also assessed his subjects' incidental learning. He presented each subject with a list of words used during the discrimination phase of the study, and asked them to draw a line connecting the words that had been paired during the discrimination trials. On this associative task, none of the main or interactive effects approached significance: The retarded and nonretarded groups were not reliably different.

The correlation between number of trials to criterion and associative matching scores was computed for all 16 subgroups, and was significant for only 1 group. Thus, the seeming equality of retarded and nonretarded groups in incidental learning in this study was apparently not an artifact of the slower initial learning by the retarded group.

Hetherington and Banta (1962). This study deserves special note for the care with which subjects were selected despite the fact that it was carried out before the developmental hypothesis had been articulated (cf. Zigler, 1966, 1969). Hetherington and Banta (1962) classified noninstitutionalized retarded children as familially retarded (and hence selected them for inclusion in the study) by the absence of known organic brain damage, and by the presence of one or more immediate family members in an institution or special class for retarded individuals.

All subjects took part in four sessions. The first session assessed incidental learning using a memory for colored objects test presented to the children as a color-naming test. Two days later, the subjects were tested on the same material for incidental retention (delayed free recall of incidentally learned objects), and incidental recognition (delayed recognition of incidentally learned objects). In the third session intentional immediate recall was assessed, and in the fourth session intentional delayed retention was tested.

Both the nonretarded children and a group of familially retarded children scored higher on incidental learning than a group of organically impaired children, but the first two groups did not differ significantly from each other. None of the groups differed on incidental recall. On incidental recognition, the nonretarded and familially retarded groups both had higher scores than the organically impaired group, but once again did not differ significantly from each other. There were no significant differences between any of the groups on intentional learning, or on intentional recall or recognition.

Forehand, Calhoun, Peed, and Yoder (1973). Forehand et al. (1973) studied the effect of distracting stimuli on incidental learning. Subjects were shown a series of colored pictures and were told to name the color and that they would later be tested to see how many of the colors they could remember. The children were assigned to either a no-distraction condition, a white noise condition, or a varied noise condition involving a tape-recorded child's story. After viewing the pictures, subjects were asked to recall the objects in the pictures, rather than the colors.

With number of correct responses as the dependent variable and number of total responses as a covariate, an analysis of covariance revealed no significant effects for group, noise condition, or their interaction. Under all conditions, then, the retarded and nonretarded subjects did equally well. However, retarded subjects gave significantly more incorrect (and hence total) responses than the nonretarded subjects. This latter finding was interpreted as an instance of Zigler's (1966) "maximizing behavior." This concept posits that retarded individuals have a higher expectancy of failure than nonretarded individuals, and that as a result, they attempt to maximize correct responses by emitting a large number of responses, without due regard for the number of errors they may make.

Though once again the small number of studies prevents reaching firm conclusions, the available evidence seems to contradict the hypothesis that retarded persons suffer from a deficit in incidental learning. We now turn to our last section before reviewing the evidence.

Complex Processes

The five studies we review next focus on higher-order or integrated cognitive functioning—functioning that appears to involve the interplay of multiple processes. Included in this section are investigations of concept usage, hypothesis testing, and humor.

Concept Usage and Matching

Two studies in the area of concept usage and matching satisfied our inclusion criteria.

Blount (1971). Blount (1971) asked retarded and nonretarded children to name everything they could that fit various common concept categories (e.g., animals, "sticky"). Responses were rated by judges as to whether they were appropriate examples of the concepts. The subjects were later asked which three of five pictures constructed from their earlier responses went together. After subjects made their choices, the experimenter covered up the two irrelevant pictures and asked why the three related pictures went together. Subjects also were administered Hammill and Irwin's (1966) test of abstraction ability.

The two groups of subjects did not differ significantly in abstraction ability, total number, number of appropriate responses to the concept-member naming task, or (most important) the number of correct pictures chosen in the concept-choice task. However, the nonretarded group was significantly better at verbalizing the relation between the three relevant pictures. Blount concluded that the retarded group was specifically deficient in their ability to verbalize relations.

Levitt (1975). In a somewhat different approach, Levitt (1975) used a concept-matching task that required subjects to match an index object to another object (the "response" object) presented within an array of "distractor" objects. The index and response objects were related either through common or independent function (e.g., respectively, a comb and brush; a match and candle). Subjects performed the concept-matching series once using three-dimensional objects, and once using drawings of the objects; children were asked to identify the objects or drawings before beginning the concept-matching task.

Levitt found that the retarded and nonretarded groups did not differ significantly on either the number of items correctly identified or concept matching, when performance was averaged across dimensional condition. However, although the performance of the two groups was equivalent on the three-dimensional concept-matching test, the performance of the retarded subjects on the two-dimensional concept-matching test was "more adequate" than that of the nonretarded children. Levitt hypothesized that the relative superiority of the retarded group under the two-dimensional condition may have been due to their greater CA.

Hypothesis Testing Behavior

Weisz and Achenbach (1975). An important aspect of cognitive functioning is the use of hypotheses and strategies in problem solving. Although several studies have included questions asking the subject how he or she knew the answer, Weisz and Achenbach (1975) used a relatively nonverbal, blank trial discrimination learning procedure to assess the use of hypotheses. Subjects were shown a series of Plexiglas "cards" that contained two different shapes side by side. Under the "2D" (i.e., two-dimensional) condition, the shapes varied in terms of shape and position on the card (i.e., left- or right-hand side). Under the "4D" condition, the stimuli varied in shape, position, color, and size. Children were instructed to figure out what the "correct" dimension was for a series of nine cards. On some trials the children received feedback about the correctness of their responses, whereas on other trials the feedback was withheld (i.e., "blank" trials). Response patterns during these blank trials revealed the use of hypotheses (e.g., consistent choice of the blue stimulus across several no-feedback trials).

Neither IQ nor MA had a main effect on hypothesis usage, but there was a significant IQ × Problem Type interaction. Retarded and nonretarded subjects were about equally likely to use hypotheses under the 2D condition; but under the 4D condition the retarded subjects used hypotheses less often than the nonretarded group. The higher IQ group was also superior in maintaining the same hypothesis when told that an answer was correct. However, the two IQ groups did not differ in their ability to switch to a new hypothesis when told that a response was incorrect. There was a main effect for MA on this latter variable, however, with higher MA subjects using "wrong" feedback more effectively than lower MA subjects.

Two related explanations were offered for the differences between the retarded and nonretarded groups. It was suggested that the results might support Harter's (1965) and Weir's (1967) hypothesis that IQ effects are found only when a task is relatively complex. It was also suggested that the retarded children's general lack of confidence in their problem-solving ability might have been a factor. As the difficulty of the problem increased, the retarded subjects' confidence in their solution and their readiness to stick to it may have decreased.

Weisz (1977). A follow-up study by Weisz (1977) involved similar blank-trials methodology. In this study, though, the procedure controlled for the effects of certain response sets (e.g., stereotyped left–right–left–right responding), in part by substituting a letter dimension (*X* or *O*) for the position dimension. There was also a "rotating" condition, in which the shapes were placed on a rotating turntable (thus eliminating any stable stimulus position), and a "stationary" condition, in which the stimuli were stationary. In both conditions, only the 4D condition was included.

Results indicated that the IQ groups did not significantly differ from each other in the amount of help they needed during practice problems, though the higher MA groups required less help. Hypothesis usage during the blank-trial blocks (i.e., trials with no feedback) also was unrelated to IQ, but was more frequent for the higher MA groups. Similarly, position responding (in the stationary condition) was not affected by IQ, but decreased with increasing MA. The effect of "wrong" feedback was not related to IQ, but was related to MA. Finally, there was a significant MA × IQ interaction after "right" feedback. At the highest MA only, retarded subjects scored lower than the nonretarded subjects; at the other two MA levels the IQ effect was not significant. Weisz concluded that the apparent effects for IQ found by Weisz and Achenbach (1975) under the 4D condition may have been the result of specific experimental methodology (e.g., a failure to control for position response sets), rather than a deficit invariably inherent in low-IQ persons.

Although the preceding four studies revealed a few differences (in both directions) between retarded and nonretarded groups, they generally indicate a rough equivalence between retarded and nonretarded children in the areas of concept usage and matching, and hypothesis testing behavior. However, as in earlier sections of our review, the small number of studies precludes firm conclusions.

Humor

Zigler, Levine, and Gould (1966). Humor is a complex response involving motivational, emotional, and experiental as well as cognitive components. It thus is a challenging context for a test of the developmental hypothesis. Zigler et al. (1966) structured such a test, focusing on children's responses to cartoons. All children were administered the Children's Mirth Response Test, in which children read (or are read) a series of cartoons, then are asked to explain why the cartoon is funny. The nonretarded group in this study showed significantly higher comprehension of the cartoons than both an institutionalized and a noninstitutionalized retarded group, who did not differ significantly from each other.

Status of Evidence on the Similar Structure Hypothesis

We have reviewed the findings of 24 studies involving 59 carefully structured group comparisons, each directly relevant to the similar structure hypothesis. What do these comparisons, taken together, tell us about the status of the hypothesis? To answer this question, let us first examine Table 1, which summarizes the findings. Table 1 shows that the similar structure hypothesis is supported by slightly more than half of the group comparisons. Whereas 52% of the comparisons resulted in no significant retarded–nonretarded group differences, another 45% revealed significant differences, with the nonretarded group outperforming their mentally retarded MA peers. Only 3% (two comparisons) supported Kohlberg's (1968) unconventional difference position, with retarded subjects outperforming the nonretarded. At first blush, the overall pattern might appear to be essentially an empirical standoff, with half the evidence supporting the similar structure hypothesis and half supporting one of the difference positions. However, as Weisz and Yeates

Table 1
Summary of Comparisons of MA-Matched Retarded and Nonretarded Persons on Information-Processing Tasks

Study	MA	IQ	Special features	R > N	N = R	R < N
August (1980)	11.0	74				Item recall (5.5), category recall (2.1), speed (2.5), sorting consistency (2.7), similarity to norm-group categories (1.9), similarity in input/recall categorization (5.5)
Blount (1971)	7.7	65			Abstraction ability (−0.2). Concept naming: total number (A, 0.9), number appropriate (A, 0.8), and concept matching (0.7)	Verbalization of relationship (2.1)
Blue (1970)	11.2	71				Visual paired-associate learning (1.8)
Brown (1974)	8.2	74			STM: ordered recall (1.4), unordered recall (0.5)	
Crosby (1972)	9.2	66	IO, INO	Ability to sustain attention (−3.1)	Resistance to distraction (0.2)	
Fagan (1968)	8.7	74		STM retention: ordered recall (−1.8)	STM retention: unordered recall (−0.5)	STM acquisition: unordered recall (3.1), ordered recall (2.9)
Forehand, Calhoun, Peed, & Yoder (1973)	6.9	64			Effect of distracting stimuli (0.0)[a] incidental learning (−0.1)	
Groff & Linden (1982)	9.5	67			WISC–R factor pattern (0.5)	
Harter (1965)	5, 7, and 9	70	MOT		Verbalize solution (0.3)	Object discrimination (5.5)
Harter (1967)	5, 7, and 9	67	MOT, INO			Object discrimination (3.0)
Harter, Brown, & Zigler (1971)	8.0	67	MOT, INO			Picture discrimination (3.2); verbalize rule (1.0)
Hetherington & Banta (1962)	5.9	61	ORG		Incidental learning (0.3); delayed incidental recall (A, −0.2); delayed incidental recognition (A, −0.2); intentional learning (0.4); delayed intentional recall (B, −0.5); delayed incidental recognition (B, −0.1)	
Hornstein & Mosley (1979)	9.2	63			Iconic memory (0.2)	
Levitt (1975)	4.7	65			Concept matching (0.2); item identification (0.2)	
Libkuman (1972)	9.2	66			Incidental learning (0.5)	Verbal discrimination (1.7)
Mosley (1980)	9.9	66			Iconic memory (familiar stimuli) (0.5)	Iconic memory (unfamiliar stimuli; 2.7); effect of masking (NA)
Mosley (1981)	11.4	73				Letter iconic memory (4.8); effect of masking (NA)
Nesbit & Chambers (1976)	8.9	65				Raven's PCM (3.1)[b]
Richman, Adams, Nida, & Richman (1978)	5.0	58			Object discrimination (0.2)	Dimensional shifts (2.4)
Weisz (1977)	5.5, 7.5, and 9.5	70	MOT		Help needed during discrimination practice (0.15); use of hypothesis-testing behavior (0.15); position responding (0.15); use of "wrong" feedback (0.15)	Use of "right" feedback (NA)

Table 1 (*continued*)

Study	MA	IQ	Special features	R > N	N = R	R < N
Weisz & Achenbach (1975)	5.5, 7.5, and 9.5	68	MOT		Hypothesis use (2D condition; 0.15); use of "wrong" feedback (0.15)	Hypothesis use (4D condition; 1.9); use of "right" feedback (2.2)
Winters & Attlee (1974)	7.9	67				Word STM (4.6)
Winters, Atlee, & Harvey (1974)	7.9	64	MOT, INO		Picture paired-associate (0.2)	
Zigler, Levine, & Gould (1966)	9.9	66	INO			Comprehension of cartoons (2.6)[c]

Note. The value in parentheses is the Z score for the effect's probability level. Letters in parentheses refer to effects within a given study that were collapsed for the meta-analysis: Effects with the same letter were collapsed together. Retarded IQs are for nonorganically, noninstitutionalized retarded subjects, collapsed across groups where necessary. IO = institutionalized, organically retarded group included. INO = institutionalized, nonorganically retarded group included. MA = mental age. MOT = study included motivation manipulation. NA = Z score is not applicable because the effect involved an interaction. ORG = noninstitutionalized, organically retarded group included. PCM = Progressive Coloured Matrices. R > N = retarded subjects' performance was superior to nonretarded subjects' performance. R = N = retarded and nonretarded subjects' performances were equivalent. R < N = retarded subjects' performance was inferior to nonretarded subjects' performance. STM = short-term memory. WISC–R = Wechsler Intelligence Scale for Children–Revised. 2D = two-dimensional condition. 4D = four-dimensional condition.
[a] The retarded group made significantly more incorrect and total responses. [b] Retarded and nonretarded subjects also differed on the Rod and Frame Test and the Embedded Figures Test. [c] The retarded and nonretarded groups did not differ significantly on the number of cartoons they said were funny.

(1981) noted, the similar structure hypothesis is essentially a null hypothesis and is thus most appropriately tested against an expected normal distribution of group differences—that is, the distribution expected if the similar structure hypothesis is in fact valid. We carried out such a test with these data, first using a chi-square analysis, and then a meta-analysis. Before reporting the results we describe the meta-analytic procedures we followed.

Meta-Analytic Procedures

For each of the studies reviewed, the Z-score equivalent (Rosenthal, 1978) for the probability of the group effect (the retarded–nonretarded comparison) was computed: F and t values for each group effect were converted via computer (to minimize rounding errors) to p values, and hence to Z scores. (These Z scores are included in Table 1). To minimize interdependence, similar dependent measures within studies (e.g., incidental recall and incidental recognition in Hetherington & Banta, 1962) were collapsed.[4] Where more than two groups contributed to the group effect (e.g., when there was also a CA-matched group), a separate effect for the two groups of interest was obtained by working backward from the group means to obtain the sum of squares between (for the three groups). The mean square between was then computed and hence mean square error. Finally, the F ratio was recomputed, using the pooled mean square error, mean square between for the two groups of interest, and adjusted degrees of freedom.

When the published article did not contain enough information to perform this conversion (e.g., "effect was not significant"), we tried to obtain the actual p or F value from the investigator. If we could not obtain the information in this manner, the average reported nonsignificant effect for group was taken as an estimate for nonreported, nonsignificant group effects. However, because different nonsignificant effect sizes were noted when the experimental stimuli were verbal (e.g., words, letters, numbers) or nonverbal (e.g., pictures, drawings), separate mean effects for the two types of stimuli were computed. Then, for each nonsignificant, nonreported t or F, the appropriate estimate was substituted.

The second part of our meta-analytic procedures consisted of coding independent variables. We describe these various factors in following sections.

Summary Comparisons: Chi-Square and Meta-Analysis Results

Table 2 is a summary of the findings of our review together with those of an earlier review focused on Piagetian research (Weisz & Yeates, 1981). To maximize comparability with the present review, we include only those studies from the Piagetian review that included procedures for screening organically impaired subjects from the mentally retarded samples involved. In the Piagetian review, Weisz and Yeates (1981) used only a chi-square analysis; we used a chi-square analysis to provide a comparable, categorical look at the data, then applied a more powerful meta-analytic technique (Rosenthal, 1978), as described previously. As Table 2 shows, the distribution of findings of the Piagetian studies did not differ significantly from the hypothetical normal distribution that would be expected if the similar structure hypothesis were valid. By contrast, our chi-square analysis showed a highly significant deviation from the distribution expected under the null hypothesis: $\chi^2(2, n = 59) = 185.77, p < .0001$.

[4] For the sake of descriptive completeness, the listings in Table 1 were not collapsed. Thus, the degrees of freedom in the meta-analysis do not equal the number of comparisons in this table. To ascertain whether collapsing might have biased results, we reran the meta-analysis without collapsing effects. Results were essentially unchanged.

Table 2
Comparisons of MA-Matched Retarded and Nonretarded Groups on Piagetian and Information-Processing Tasks

	Direction of difference			
Tasks	MR > NMR	MR = NMR	MR < NMR	Statistical test
Piagetian[a]				
Actual totals	0	30	3	$\chi^2(2, n = 33) = 2.76, p > .25$
Expected totals	1.65	29.70	1.65	
Information processing				
Chi-square analysis[b]				
Acual totals	2	31	26	$\chi^2(2, n = 59) = 185.77, p < .001$
Expected totals	2.95	53.10	2.95	
Meta-analysis[c]	$\bar{Z} = 1.32$			$t(52) = 5.37, p < .0001$
	$\bar{Z} = 1.37$			$t(46) = 5.23, p < .0001$

Note. MA = mental age. MR = mentally retarded. NMR = non-mentally retarded.
[a] These data are from Weisz and Yeates (1981). They are the results of 33 retarded–nonretarded group comparisons in which the groups were matched for MA and in which investigators tried to exclude organically impaired subjects from the sample. [b] These data represent the 59 retarded–nonretarded group comparisons we surveyed in this review. These comparisons, likewise, involved MA-matching and efforts to exclude organically impaired subjects. [c] The first set of values includes all comparisons; the second set excludes comparisons that did not pass our organicity criteria based on the published article.

Similarly, our meta-analysis revealed a highly significant difference between the performance of the retarded and nonretarded groups, $\bar{Z} = 1.32$; test that $\bar{Z} = 0$: $t(52) = 5.37, p < .0001$. Retarded subjects performed significantly worse than nonretarded subjects, *averaged* across tasks and subject samples. Note that certain task or subject variables may be related to these performance differences in theoretically important ways; such relations may suggest factors that aid in explaining the group differences. One such relation, an interaction between age and IQ, has been proposed by several investigators (e.g., J. P. Das, personal communication, July 24, 1984; Spitz, 1976, 1982; Weisz, 1979).

To test this prediction in our data, we included the MA and CA of the retarded subjects as variables in our meta-analysis. Although CA was not significantly correlated with the retarded–nonretarded differences (r = 0.27, *ns*), MA was (r = 0.44, $p < .005$)—that is, the higher the subjects' MA, the higher the probability that the nonretarded subjects' performance was superior to that of the retarded subjects. We discuss various interpretations of this finding in the next section.

It is important to first note, however, one possible artifactual source of this relation. If researchers found it easier to work with more cognitively advanced subjects (i.e., those with higher MAs), there might be a positive correlation between MA and sample size. Such a relation would mean smaller standard errors at higher MA levels (because of the larger sample size), and a greater likelihood of finding significant differences between the groups. However, the correlation between MA and sample size (as well as the square root of the sample size) was nonsignificant ($p < .25$).

Group Differences as a Function of Cognitive Domain

Focusing again on Table 1, it would be useful to delineate the cognitive domains where retarded and nonretarded groups differed most significantly. However, results in some categories were inconsistent, making interpretation of these results problematic. Consequently, along with the mean Z statistic (Rosenthal, 1978) for each category, we also computed a chi-square statistic that evaluated homogeneity of significance within a category (Rosenthal & Rubin, 1979) to warn us when significant heterogeneity might make it necessary to approach our conclusions with caution.[5] We shall note those categories in which this is the case.

As Table 3 shows, the performance of mentally retarded subjects was most deficient in the area of discrimination. Evidence exists supporting a deficit for retarded persons for verbal (printed word) discrimination, picture discrimination, and three-dimensional object discrimination. However, the chi-square statistic indicated that this deficiency was not consistent across studies within this group. Comparison of the studies showing the largest differences (Harter, 1965, 1967; Harter et al., 1971) with the study showing the smallest (i.e., nonsignificant) difference (Richman et al., 1978) revealed the following: First, all studies in the former group had the same first author. Thus the three studies might have showed similar large group differences because they were designed by the same investigator, who used similar methodologies across studies. However, although the designs of Harter's 1965 and 1967 studies were quite similar, the design of her 1971 study differed from both. Further inspection of the clearest and most consistent difference among these four studies was that those studies showing the largest differences all included a motivational manipulation, whereas the study that failed to show a significant difference did not. This might suggest that *nonretarded* persons benefit most from attempts to raise motivation. However, collapsing across all

[5] As Rosenthal and Rubin (1979) noted, this statistic is affected by the sample size, as well as the raw effect size. However, inspection of our chi-squares indicate that our results reflect more than sample size, in that the relation between degrees of freedom and probability is not consistent or linear (i.e., categories with lower degrees of freedom in some cases have higher p values; e.g., incidental learning, the third largest category, has the least significant p value).

Table 3
Summary of Cognitive Domains

Domain	No. of comparisons	Mean z	p	$\chi^{2,a}$	P
Concept usage	3	0.58	.16	0.29(2, n = 3)	.87
Discrimination and discrimination set-learning	6	2.64	.0001	15.9(5, n = 6)	.007
Distractibility and selective attention	3	−0.97	.05	7.07(2, n = 2)	.03
Verbal explanations	3	1.09	.03	1.65(2, n = 2)	.44
Hypothesis-testing behavior	3	0.73	.10	2.04(2, n = 2)	.36
Incidental learning	4	0.14	.39	0.30(3, n = 4)	.96
Memory	21	1.93	.00001	82.3(20, n = 21)	.0001

Note. We excluded categories with only one comparison (e.g., cartoon comprehension; Zigler, Levine, & Gould, 1966). In reading this table, it is important to remember that the significance level is a function of both the mean probability level, and the number of comparisons.
[a] Chi-square statistic tests for homogeneity within category (Rosenthal & Rubin, 1979). A large chi-square and small probability indicates significant heterogeneity with category (i.e., probability level varies significantly across studies within the particular category).

comparisons in our review, we found that when a motivational manipulation was included, the performance deficit of retarded persons was *reduced,* though not significantly. (We discuss this finding more fully in the section on *Extracognitive variables.*)

Moving to the next category, we found that a strong deficit apparently exists in the area of memory. Deficits for retarded individuals have been found in serial and nonserial auditory short-term memory (STM), visual iconic memory, visual STM, cross-modal STM, and visual paired-associate learning. However, we found nonsignificant differences for serial STM, susceptibility to the von Restroff effect, visual iconic memory, and verbal paired-associate learning, and the chi-square statistic revealed a significant heterogeneity within this category. Therefore, it probably is safest to conclude that this apparent deficit is specific to certain aspects of memory.

One area in which nonretarded subjects were consistently superior was in their ability to explain the strategies they used in the various experiments. In some cognitive areas, however, the performance of retarded subjects was comparable to that of their nonretarded MA peers. The evidence indicates that retarded children and their MA peers perform equally well on tests of incidental learning. Interestingly, retarded subjects also consistently failed to show a performance deficit on most tests of more integrated or higher level cognitive behaviors, such as, concept usage and matching, and hypothesis testing behavior. This finding seems to argue against Spitz's (1976) hypothesis that retarded individual's are particularly deficient in cognitive areas that have developed relatively recently in human evolution. (However, see our discussion of a difference interpretation of the MA × IQ interaction).

Finally, in the area of distractibility and selective attention, the performance of retarded subjects, averaged across studies was superior to that of nonretarded subjects. However, this finding was not consistent; two of the three tests in this area showed no significant differences.

Within-Study Interaction Effects

We also tried to further our understanding of these data by focusing on significant interaction effects within the studies. In fact, such effects have often been considered by information-processing researchers to be particularly meaningful in investigations of process differences. Therefore, we include a table of significant interaction effects (Table 4). As seemed to occur with main effects (i.e., group differences), there was no consistent pattern across studies. For instance, in some cases the complexity of the task seemed to hinder retarded subjects (e.g., Weisz & Achenbach, 1975), whereas in other instances it appeared to favor them (e.g., Fagan, 1968). Similarly, there was an inconsistent interaction involving reinforcement condition. In one study (Harter, 1967) retarded subjects performed better under a standard reinforcement condition (i.e., no social reinforcement), whereas in another (Harter et al., 1971) retarded subjects' performance was superior under the social reinforcement condition.

It is notable, however, that only in two interactions did the retarded subjects perform better than the nonretarded subjects (when such a comparison was meaningful). This seems to be further evidence against the similar structure hypothesis, in that if random variability were responsible for the significant interactions, one would expect the performance differences to be equally distributed across groups.

Perspectives on the Obtained Group Differences

What then are we to conclude from the findings just reviewed? Perhaps that the information-processing evidence argues for rejecting the similar structure hypothesis, and accepting the general difference position on mental retardation (e.g., Milgram, 1973)—at least for some important cognitive processes. Such a conclusion might be premature, however, as several alternative explanations are plausible enough to warrent scrutiny and testing. Among such alternatives, perhaps the most plausible is one that could (a) explain why the similar structure hypothesis seems to hold for Piagetian tasks (Weisz & Yeates, 1981) but not for information-processing tasks, and (b) encompass the study-to-study inconsistencies within the present review by identifying some variable or variables that explain the pattern of results we obtained. In the next section we consider several perspectives from which to view our data—perspectives that may help us move toward these objectives.

Artifactual explanations. Inconsistent results within our review in areas such as memory might be a function of the subject

Table 4
Within-Study Interactions

Study	Interaction
August (1980)	Three complex interactions indicated that the effect of the different sorting instructions was not equivalent for retarded and nonretarded subjects. This was true for the number of nouns recalled, the number of categories recalled, and recall organization. A fourth interaction indicated that the effect of stimulus relations (high or low) on sorting time was not consistent across group.
Fagan (1968)	Recall of digit strings: For ordered (but not unordered) recall, the performance of retarded subjects was less impaired by retroactive interference than the performance of nonretarded subjects, particularly under the low-interference condition.
Harter (1965)	Discrimination learning: There was an MA × IQ interaction, apparently the result of a ceiling effect at the upper MA and IQ levels.
Harter (1967)	Discrimination learning: Nonretarded subjects learned faster under the social reinforcement condition, whereas retarded subjects learned faster under the standard condition.
Harter, Brown, & Zigler (1971)	Simple three-choice discrimination task (but not the more complex oddity discrimination task): (a) Nonretarded subjects reached asymptotic responding faster than retarded subjects. (b) Retarded subjects performed better under the social reinforcement condition, whereas nonretarded subjects performed equally well under both conditions.
Hornstein & Mosley (1979)	Iconic memory: The Group × Response Type (yes/no) interaction was significant. Retarded subjects were more accurate when the correct answer was "yes," whereas nonretarded subjects were equally accurate under both conditions.
Levitt (1975)	Concept matching: Nonretarded children performed significantly worse on the two-dimensional task than the three-dimensional task, whereas retarded children did equally well under both conditions. The correlation between identification and concept matching was nonsignificant for nonretarded children but highly significant for retarded children, under the two-dimensional condition only.
Mosley (1980)	Iconic memory: Whereas retarded and nonretarded subjects performed equivalently under minimal or maximal masking delay for both familiar and unfamiliar stimuli, under conditions of intermediate masking delay the retarded group did worse, and this effect was more pronounced when the stimuli were unfamiliar.
Mosley (1981)	Iconic memory: There was a significant Group × Masking-Delay Interval × Information Load (array size) interaction. Essentially, the shorter the masking-delay interval and the greater the information load, the more the performance of the retarded subjects deteriorated relative to the nonretarded subjects.
Weisz (1977)	Following the feedback "right," but not the feedback "wrong": (a) Retarded subjects were more likely to repeat hypotheses when the stimuli were stationary than when they were rotating, whereas the nonretarded subjects showed the reverse effect. (b) There was a significant IQ × MA interaction; at the highest MA only, retarded subjects were less likely to use hypotheses than nonretarded subjects.
Weisz & Achenbach (1975)	Hypothesis use: (a) Retarded children performed worse than nonretarded children under the four-dimensional condition but not under the two-dimensional condition. (b) Nonretarded children's use of the feedback "right" was superior to that of retarded children, whereas the two groups did not differ in their use of the feedback "wrong."
Winters & Attlee (1974)	Recall of nouns: Nonretarded subjects performed better than retarded subjects at all serial positions except 6 (the von Restorff position) and 10 (the last position).

Note. MA = mental age.

population sampled,[6] or of the design and methodology of the studies. Though we tried to create a homogenous population through our selection criteria, across studies the retarded populations are undoubtedly still heterogeneous. Some retarded individuals, with a certain set of experiences, may evidence cognitive deficits, whereas others with different experiences may not. These differences may be nonrandom (e.g., related to selection methods, geography, quality of special education in public schools, etc.), and may result in different patterns of results coming from different investigators.

Moreover, two tasks that appear to tap similar cognitive domains via somewhat different methodologies may not in fact be assessing the same ability. For instance, the apparent contradiction between the results obtained by Winters et al. (1974), and Blue (1970) may have been a result of differences in their methodologies. Though both studies involved paired-associate paradigms, specific differences in the procedures (e.g., the use of nonsense shapes vs. common objects) may have resulted in their assessing different abilities.

The disparity between our results and those of Weisz and Yeates (1981) might also be due to an artifact. For example, inconsistencies could result partly from experimenter selectivity and bias. Investigators who perform Piagetian studies may be more likely than those who conduct information-processing studies to favor the developmental hypothesis, which is closely linked to Piagetian theory (cf. Zigler, 1969; Zigler & Balla, 1982). Through various kinds of subtle and inadvertent experi-

[6] In our earlier description of inclusion criteria, we noted that past evidence indicated institutionalized and noninstitutionalized retarded individuals do not always show equivalent cognitive functioning. Our results support this view. Of the eight studies reviewed that contained both kinds of retarded groups, all but two showed a significant difference between the groups on at least one dependent variable we reported. Of the two that did not, Zigler, Levine, and Gould (1966) did find such differences on variables we did not report. In addition, though Winters, Attlee, and Harvey (1974) did not find a significant difference between the performance of their noninstitutionalized and institutionalized retarded groups, the latter group did differ significantly from the nonretarded group, whereas the former group did not.

menter bias (Rosenthal, 1966) on sampling, task administration, and so on, investigators in the two groups might find results supporting their point of view. Such an experimenter effect might explain the inconsistency between reviews. However, although it may be reasonable to attribute some inconsistencies across studies to an experimenter effect, it seems unlikely that an experimenter effect could be so powerful as to fully account for the strong consistency within Weisz and Yeates's (1981) review.

A variety of alternative artifactual explanations could be advanced. For example, it might be argued that Piagetian tasks are more susceptible to ceiling and floor effects (cf. Spitz, 1983) than information-processing tasks, or that the information-processing studies used inferior procedures for organicity screening. These and the other artifactual explanations we have considered seem at least moderately plausible. Yet each one also seems ill-equipped to account for both the strong consistency of the Piagetian findings and the marked variability within the present review. An appealing alternative is that the pattern of results across the two reviews is due to factors inherent to the tasks themselves.[7]

Interpretations Favoring the Developmental Perspective

Without direct data, it would be difficult to predict whether artifactual factors bias results toward the developmental or the difference position. These factors could attenuate differences between groups (biasing results toward the developmental position), or they could produce spurious differences between groups (biasing results toward the difference position). The hypotheses that follow, unlike the artifactual explanations, support either one position or the other. They have been broken down into two groups: those that support the developmental position, detailed immediately below, and those that support the difference position, described in the section immediately following the developmental explanations.

Ecological validity. One task-based explanation for the differences between our results and those of Weisz and Yeates (1981) is that the Piagetian tasks reviewed by Weisz and Yeates (e.g., making moral judgments) are more "lifelike" or "ecologically valid" than the tasks used in studies included in our review (e.g., discrimination learning of colored shapes). This may be true of the skills themselves, or of the manner in which they were assessed. The manner in which the Piagetian skills have typically been assessed tends to be more lifelike, conversational, and "natural" than the typical means of assessing discrimination skills, such as paired-associate learning, and so on.

In fact, Brooks and Baumeister (1977) have suggested that all too often experimental tasks used in retardation research may tell us little about the nature of mental retardation and even less about the "real-life" adaptation of retarded persons. This criticism may be more applicable to information-processing tasks than to Piagetian tasks. Thus, our findings may differ from those of Weisz and Yeates (1981) partly because the two domains of research surveyed may differ in level of ecological validity (but see also House, 1977, for a well-developed opposing viewpoint, and Ellis, Harris, & Baker, 1983, for evidence contrary to Brooks & Baumeister's hypothesis).

Extracognitive variables. Another explanation from the developmental position is that retarded and nonretarded children differ significantly on some extracognitive variable (e.g., motivation, expectancy for success) that may influence some types of tasks (e.g., some information-processing measures) more than others (e.g., Piagetian). Some reviews (e.g., Zigler & Balla, 1982) indicate that retarded and nonretarded individuals of similar MA appear to differ on various other extracognitive factors. Of course, this evidence alone is not sufficient to validate a motivational explanation for the observed group differences on cognitive tasks. Motivational differences do not preclude the existence of group differences in cognitive ability. To rule this out, what needs to be shown is that when extracognitive differences are controlled, differences between retarded and nonretarded subjects on cognitive task performance are no longer statistically significant.

Information contained in the present review permitted one rough test of this hypothesis. In several of the studies we reviewed, subjects were told they would receive some form of reward for good performance; the intent was to insure that for all subjects, the motivation for success on the experimental task would roughly be equated at a relatively high level. As part of our meta-analysis, we tested whether the presence or absence of a motivational manipulation affected the retarded–nonretarded group comparisons. Comparing studies that included a motivational manipulation to those that did not, we found a Z of only 0.37 for the differential between the two groups of studies. The resultant t was nonsignificant, $t(51) = 0.69, p > .25$, indicating that controlling for motivation made little difference. However, whether such short-term manipulations could effectively counter a lifetime of experience is a question that this analysis does not answer.

Questions such as this suggest a need for theorists who emphasize extracognitive factors to formulate more specific hypotheses. Many different variables have been suggested as being responsible for retarded/MA-matched nonretarded differences (e.g., motivation—Zigler, 1969; outer-directedness—Yando & Zigler, 1971). Specific hypotheses need to be advanced as to which of these variables are most important for which subject groups, and how the effects of these variables can best be gauged and controlled within an experimental situation. Moreover, to be of maximal value, the motivational hypotheses must yield testable predictions of the specific effects of motivational variables on cognitive performance, predictions that preclude difference theories as rival explanations, and are subject to experimental disconfirmation.

IQ × MA interaction: Learned helplessness? As we noted in our description of the meta-analysis, a significant interaction between IQ and MA was discovered. J. P. Das (personal communication, July 24, 1984) based his prediction of this relation on the fact that school is often a primary site of failure for retarded children, and that with age, a child's experience becomes more and more subjectively centered on school. Thus, he suggests, as age increases, the performance of retarded persons on information-processing tasks (on which school performance is

[7] We did, however, test one possible artifactual explanation. We reran our overall meta-analysis, dropping studies that did not pass our organicity criteria on the basis of the published article. This did not affect our results (see Table 2).

heavily based) may deteriorate relative to that of nonretarded persons.

What factor or factors change with advancement in school and MA that might be related to increasing performance deficits? Some such factors may be directly tied to cognitive ability (and are discussed in the next section), but others may involve expectancy and motivational variables. For example, it has been suggested that the propensity to develop learned helplessness (and concomitant performance deficits) is positively related to cognitive developmental level or MA (Weisz, 1979). A few findings do appear to be consistent with this hypothesis. Rholes, Blackwell, Jordan, and Walters (1980) found that susceptibility to learned helplessness increased among nonretarded children as the age of the children increased from the kindergarten through late elementary years. In addition, in a developmental analysis of learned helplessness among retarded and nonretarded children, Weisz (1979) found that retarded children were more helpless, relative to MA-matched nonretarded groups, at higher MA levels than at lower levels. That is, as MA increased, the helplessness of the retarded children relative to that of the nonretarded children increased. Such a process might underlie the significant relation we found between MA and group differences, and in part explain the inferior performance of the retarded subjects in our review.

Interpretations Favoring the Difference Perspective

It is quite possible that even if all of the abovementioned factors (i.e., artifactual effects, motivation, ecological validity, learned helplessness) were controlled, we would still find that the development of retarded children within some cognitive domains is inferior to that of their nonretarded MA peers. Our findings may merely reflect what they appear to reflect: Retarded persons suffer from various cognitive deficits that are more than a simple developmental delay. If this were true, one could argue—independent of the evidence reviewed here—that the cognitive domains most likely to reveal differences are those within the information-processing tradition. Information-processing research, after all, developed originally out of a search for individual and group differences. The Piagetian tradition, by contrast, grew out of a search for the commonalities that underlie development. Thus, Piagetian research may have evolved toward the use of tasks sensitive to developmental change, but not to differences between individuals or groups at similar developmental levels. Regardless of this argument's validity however, the inconsistencies across the tasks we reviewed remain to be explained.

IQ × MA interaction: Cognitive complexity? Spitz (1976, 1982), as we noted in the introduction to the difference position, has suggested that retarded individuals are particularly deficient in relatively complex cognitive processes (e.g., abstraction, problem solving, foresight). Following this reasoning, one would expect that at low-MA levels retarded and nonretarded children might perform equivalently because the complex abilities on which they (will ultimately) differ are equally undeveloped in both groups. In contrast, at higher MA levels the nonretarded children would begin to outperform their retarded MA peers as the former group began to develop their more sophisticated abilities and strategies. This reasoning led Spitz to predict an MA × IQ interaction, which we did in fact find.

J. P. Das's (personal communication, July 24, 1985) hypothesis can be interpreted similarly. As children advance in school, the complexity of the tasks they are required to solve increases rapidly. Consequently, decreasing performance of the retarded subjects relative to the nonretarded subjects could be the result of the increasing complexity of the cognitive processes required for success in school, as well as in the laboratory.

Verbal or left-hemisphere deficit. It may prove useful to compare the cognitive processes involved in the tasks that showed substantial retarded–nonretarded differences with those that did not show such differences. An initial impression that we and others (e.g., K. O. Yeates, personal communication, July 10, 1984) shared was that tasks requiring substantial amounts of verbal processing were especially likely to yield significant retarded–nonretarded group differences. This, in turn, suggested that tasks requiring substantial activity in the left cerebral hemisphere might reveal the most pronounced deficit in mentally retarded persons. This reasoning sounds plausible, given the rather widely held view that retarded individuals are particularly deficient in verbal skills (see e.g., Clarke & Clarke, 1965; Ingalls, 1978). Further, it is generally accepted that the left hemisphere is more efficient at verbal processing (Bradshaw & Nettleton, 1981), and that the right hemisphere is more efficient at nonverbal processing (e.g., form perception, spatial orientation, and localization; Gordon, 1983).

Thus, we hypothesized that the retarded subjects might be particularly deficient in left-hemisphere processing. Because retarded–nonretarded differences are not found on Piagetian tasks (Weisz & Yeates, 1981; Weisz & Zigler, 1979), then the processing of Piagetian tasks should occur in the right hemisphere (if the abovementioned reasoning were true). There is, in fact, evidence suggesting that Piagetian processing occurs in the right hemisphere (Kraft, Mitchell, Languis, & Wheatley, 1980).

Following the preceding reasoning, we planned to include in our meta-analysis variables quantifying the degree to which the tasks in the studies we reviewed were dependent on verbal or left-hemisphere processing. We therefore sent descriptions of the tasks to a number of researchers considered to be experts in the area of hemispheric differentiation, and asked them to make several ratings (e.g., the extent to which a left-hemisphere deficit would impair performance) for each task. However, half of our raters spontaneously stated that they felt the complexity of the interaction between the hemispheres and the ability of each of the hemispheres to compensate for impairment in the other made post hoc ratings problematic. Consequently, we decided that using these ratings would be of uncertain validity; therefore, we did not include them in the formal meta-analysis.[8]

[8] Because our categorization of the stimuli used in the experiments seemed reliable, we performed one rough test of our hypothesis. We divided the comparisons by type of stimuli (i.e., verbal vs. nonverbal) and direction of effect. Because the two "retarded performance superior" cells had a total of only two observations, they were not included in the analysis. The likelihood chi-square ratio approached significance, $\chi^2(1, n = 49) = 3.30, p < .07$), with the frequencies in the predicted direction. That is, a higher proportion of the within-study comparisons

A nonunitary model of mental retardation? In an attempt to explain our results from a difference position, we considered several comprehensive difference theories. The verbal mediation theory (cf. Borkowski & Wanschura, 1974) is one such theory. Briefly, it proposes that retarded persons produce fewer verbal mediators, and that their inferior performance on many tasks is a consequence of this deficiency. Although some evidence supports this position (e.g., Borkowski & Wanschura, 1974), our data are not entirely consistent with this theory. For instance, it is unlikely that iconic memory deficits (e.g., Mosley, 1981) result from a verbal mediation deficit, neither does it seem probable that subjects suffering from a verbal mediation deficit would perform well on complex discrimination tasks (e.g., Richman et al., 1978).

We considered other explanations, such as inhibition theory (e.g., Evans, 1982; Heal & Johnson, 1970), which proposes that retarded persons are unable to properly inhibit learned responses or suppress extraneous stimulus input when such actions would be adaptive. This theory also seems consistent with some but not all of the findings reviewed.

After weighing several promising theoretical accounts, we concluded that no single theory or explanation explained our results in a truly adequate, comprehensive fashion. Our failure might lead a reader unfamiliar with this research domain to conclude that it suffers from inadequate or improper development. We do not believe that to be true, however, and favor a different interpretation. What we would suggest, on a heuristic basis, is that retarded individuals may be deficient in more than one fundamental process, and that the differences observed across studies in this review may be the result of several separate factors. Thus, attempts to fit a model based on a specific unitary (albeit broad) deficit may produce results that appear inconsistent or contradictory because multiple processes are operating. The possibility that retarded persons may have multiple deficiencies has been noted previously (e.g., Fisher & Zeaman, 1973; Heal & Johnson, 1970). However, it may be necessary to explicitly develop theories in which the effects of multiple deficits are integrated into a coherent whole, for these factors may interact with each other in ways that produce patterns of abilities and deficits not predictable from the individual factors alone.

Conclusions and Future Directions

In conclusion, it is evident that in many cognitive areas, significant differences may be found between the performance of retarded individuals and their nonretarded MA peers. It may be valuable to now focus research on determining why these differences seem to exist in some areas (e.g., discrimination learning) but not in others (e.g., conservation, incidental learning). Answers to such questions could greatly expand our knowledge about the cognitive functioning of both retarded and nonretarded individuals.

It will also be important in future research to gauge the range, magnitude, and impact of retarded–nonretarded differences in such extracognitive factors as motivation. It would be difficult to overstate the importance of determining whether our results can be explained as, say, motivation or expectancy effects rather than as true differences in cognitive abilities. If the differences we have noted can be explained in terms of such extracognitive differences, it would suggest that retarded individuals may not differ cognitively from nonretarded persons of similar developmental level (a basic tenet of the developmental position). The cognitive deficits of retarded individuals might then be construed as limited to a delay in development and a relatively depressed upper limit. This in turn would have important implications for our intervention strategies and goals (Weisz & Yeates, 1981). Moreover, differences in cognitive performance between retarded and nonretarded persons of similar developmental level might then be attributable to the experiences that shape personality and motivation—experiences that are modifiable, at least in principle.

On the other hand, if the group differences we observed cannot be explained in terms of differential motivation and expectancies, a very different conceptualization would be indicated. For at least some cognitive processes, the nature of the basic deficit(s) in mental retardation should be construed as something more profound than merely a slowed pace and lower ceiling of development. This in turn would mean that the task of understanding and treating mental retardation might well be more complex than many of us have thought. Whatever interpretation of the present findings is ultimately shown to be most appropriate, these findings and the work to which they will lead, seem likely to enrich our understanding of both mental retardation and cognitive development.

showed significant differences when the stimuli were verbal than when they were nonverbal.

References

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

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

August, G. J. (1980). Input organization as a mediating factor in memory: A comparison of educable mentally retarded and nonretarded individuals. *Journal of Experimental Child Psychology, 30,* 125–143.

Balla, D., & Zigler, E. (1982). Impact of institutional experience on the behavior and development of retarded persons. In E. Zigler & D. Balla (Eds.), *Mental retardation: The developmental-difference controversy* (pp. 41–58). Hillsdale, NJ: Erlbaum.

Baumeister, A. A. (1967). Problems in comparative studies of mental retardates and normals. *American Journal of Mental Deficiency, 71,* 869–875.

Blount, W. R. (1971). Concept-usage performance: Abstraction ability, number of referents, and item familiarity. *American Journal of Mental Deficiency, 76,* 125–129.

Blue, C. M. (1970). Influence of mode of presentation, age, and intelligence on paired-associates learning. *American Journal of Mental Deficiency, 74,* 527–532.

Borkowski, J. G., & Wanschura, P. B. (1974). Mediational processes in the retarded. In N. R. Ellis (Ed.), *International review of research in mental retardation* (Vol. 7, pp. 1–54). New York: Academic Press.

Bradshaw, J. L., & Nettleton, N. C. (1981). The nature of hemispheric specialization in man. *Behavioral and Brain Sciences, 4,* 51–91.

Brooks, P. H., & Baumeister, A. A. (1977). A plea for consideration of ecological validity in the experimental psychology of mental retardation: A guest editorial. *American Journal of Mental Deficiency, 81,* 407–416.

Brown, R. M. (1974). Effects of recall order, cue placement, and retention interval on short-term memory of normal and retarded children. *Perceptual & Motor Skills, 39,* 167–178.

Clarke, A. M., & Clarke, A. D. B. (1965). *Mental deficiency: The changing outlook.* New York: Free Press.

Crosby, K. G. (1972). Attention and distractibility in mentally retarded and intellectually average children. *American Journal of Mental Deficiency, 77,* 46–53.

Crosby, K. G., & Blatt, B. (1968). Attention and mental retardation. *Journal of Education, 150,* 67–81.

Denny, M. R. (1966). A theoretical analysis and its application to training the mentally retarded. In N. R. Ellis (Ed.), *International review of research in mental retardation* (Vol. 2, pp. 1–28). New York: Academic Press.

Ellis, N. R., Harris, L. A., & Baker, H. R. (1983). Ecological validity of laboratory-type discrimination-learning tasks. *American Journal of Mental Deficiency, 88,* 106–108.

Evans, P. L. C. (1982). Inhibition and individual differences in inhibitory processes in retarded children. In N. R. Ellis (Ed.), *International review of research in mental retardation* (Vol. 11, pp. 80–123). New York: Academic Press.

Fagan, J. F. (1968). Short-term memory processes in normal and retarded children. *Journal of Experimental Child Psychology, 6,* 279–296.

Fisher, M. A., & Zeaman, D. (1973). An attention–retention theory of retardate discrimination learning. In N. R. Ellis (Ed.), *International review of research in mental retardation* (Vol. 6, pp. 171–257). New York: Academic Press.

Forehand, R., Calhoun, K., Peed, S., & Yoder, P. (1973). The effects of intelligence quotient and extraneous stimulation on incidental learning. *Journal of Mental Deficiency Research, 17,* 24–27.

Galton, F. (1883). *Inquiries into human faculty and its development.* London: Macmillan.

Gordon, H. W. (1983). Dyslexia. In R. E. Tarter & G. Goldstein (Eds.), *Neuropsychology of childhood.* New York: Plenum Press.

Groff, M. G., & Linden, K. W. (1982). The WISC-R factor score profiles of cultural–familial mentally retarded and nonretarded youth. *American Journal of Mental Deficiency, 87,* 147–152.

Hammill, D. D., & Irwin, O. C. (1966). An abstraction test adopted for use with mentally retarded children. *American Journal of Mental Deficiency, 70,* 807–812.

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

Harter, S. (1967). Mental age, IQ, and motivational factors in the discrimination learning set performance of normal and retarded children. *Journal of Experimental Child Psychology, 5,* 123–141.

Harter, S., Brown, L., & Zigler, E. (1971). Discrimination learning in retarded and nonretarded children as a function of task difficulty and social reinforcement. *American Journal of Mental Deficiency, 76,* 275–283.

Hayes, K. J. (1953). The backward curve: A method for the study of learning. *Psychological Review, 60,* 269–275.

Heal, L. W., & Johnson, J. T., Jr. (1970). Inhibition deficits in retardate learning and attention. In N. R. Ellis (Ed.), *International review of research in mental retardation* (Vol. 4, pp. 107–150). New York: Academic Press.

Hetherington, E. M., & Banta, T. J. (1962). Incidental and intentional learning in normal and mentally retarded children. *Journal of Comparative and Physiological Psychology, 55,* 402–404.

Hornstein, H. A., & Mosley, J. L. (1979). Iconic-memory processing of unfamiliar stimuli by retarded and nonretarded individuals. *American Journal of Mental Deficiency, 84,* 40–48.

House, B. J. (1977). Scientific explanation and ecological validity: A reply to Brooks and Baumeister. *American Journal of Mental Deficiency, 81,* 534–542.

Ingalls, R. P. (1978). *Mental retardation: The changing outlook.* New York: Wiley.

Kappauf, W. E. (1976). Critique of the use of covariance adjustments for CA and MA in comparative studies in retarded and nonretarded persons. *American Journal of Mental Deficiency, 81,* 240–247.

Kaufman, A. S. (1975). Factor analysis of the WISC-R at 11 age levels between 6½ and 16½ years. *Journal of Consulting and Clinical Psychology, 43,* 135–147.

Kohlberg, L. (1968). Early education: A cognitive-developmental view. *Child Development, 39,* 1013–1062.

Kounin, J. (1941). Experimental studies of rigidity: I. The measurement of rigidity in normal and feeble-minded persons. *Character and Personality, 9,* 251–272.

Kraft, R. H., Mitchell, O. R., Languis, M. L., & Wheatley, G. H. (1980). Hemispheric asymmetries during six- to eight-year-olds' performance of Piagetian conservation and reading tasks. *Neuropsychologia, 18,* 637–643.

Levitt, E. (1975). The effects of identifying objects on a concept-matching task performed by four preschool groups. *Genetic Psychology Monographs, 91,* 227–256.

Libkuman, T. M. (1972). Word frequency and pronunciation and the verbal-discrimination learning of nonretarded and retarded children. *American Journal of Mental Deficiency, 77,* 322–327.

Magaret, A., & Thompson, C. W. (1950). Differential test responses of normal, superior and mentally defective subjects. *Journal of Abnormal and Social Psychology, 45,* 163–167.

Milgram, N. A. (1973). Cognition and language in mental retardation: Distinctions and implications. In D. K. Routh (Ed.), *The experimental psychology of mental retardation* (pp. 157–230). Chicago: Adline.

Mosley, J. L. (1980). Selective attention of mildly mentally retarded and nonretarded individuals. *American Journal of Mental Deficiency, 84,* 568–576.

Mosley, J. L. (1981). Iconic store readout of mildly mentally retarded and nonretarded individuals. *American Journal of Mental Deficiency, 86,* 60–66.

Nesbit, W. C., & Chambers, J. (1976). Performance of MA-matched nonretarded and retarded children on measures of field-dependence. *American Journal of Mental Deficiency, 80,* 469–472.

Piaget, J. (1964). Cognitive development in children: The Piaget papers. In R. E. Ripple & V. N. Rockcastle (Eds.), *Piaget rediscovered: A report of the Conference on Cognitive Studies and Curriculum Development* (pp. 47–132). Ithaca, NY: Cornell University, School of Education.

Piaget, J. (1970). Piaget's theory. In P. H. Mussen (Ed.), *Carmichaels's manual of child psychology* (3rd ed, pp. 703–732). New York: Wiley.

Rholes, W. S., Blackwell, J., Jordan, C., & Walters, C. (1980). A developmental study of learned helplessness. *Developmental Psychology, 16,* 616–624.

Richman, C. L., Adams, K. A., Nida, S. A., & Richman, J. (1978). Performance of MA-matched nonretarded and retarded children on discrimination learning and transfer-shift tasks. *American Journal of Mental Deficiency, 83,* 262–269.

Rosenthal, R. (1966). *Experimenter effects in behavioral research.* New York: Appleton-Century-Crofts.

Rosenthal, R. (1978). Combining results of independent studies. *Psychological Bulletin, 85,* 185–193.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin, 86,* 638–641.

Rosenthal, R., & Rubin, D. B. (1979). Comparing significance levels of independent studies. *Psychological Bulletin, 86,* 1165–1168.

Rosenthal, R., & Rubin, D. B. (1982). Comparing effect sizes of independent studies. *Psychological Bulletin, 92,* 500–504.

Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome, S. D., & Beck, L. H. (1956). A continuous performance test of brain damage. *Journal of Consulting Psychology, 202,*343–350.

Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist, 32,* 752–760.

Spitz, H. H. (1976). Toward a relative psychology of mental retardation, with a special emphasis on evolution. In N. R. Ellis (Ed.), *International review of research in mental retardation* (Vol. 8, pp. 35–56). New York: Academic Press.

Spitz, H. H. (1982). Intellectual extremes, mental age, and the nature of human intelligence. *Merrill-Palmer Quarterly, 28,* 167–192.

Spitz, H. H. (1983). Critique of the developmental position in mental-retardation research. *Journal of Special Education, 17,* 261–294.

Steinbrueck, S. M., Maxwell, S. E., & Howard, G. S. (1983). A meta-analysis of psychotherapy and drug therapy in the treatment of unipolar depression with adults. *Journal of Consulting and Clinical Psychology, 51,* 856–863.

Thompson, C. W., & Magaret, A. (1947). Differential test responses of normals and mental defectives. *Journal of Abnormal and Social Psychology, 42,* 285–293.

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

Weisz, J. R. (1977). A follow-up developmental study of hypothesis behavior among mentally retarded and nonretarded children. *Journal of Experimental Child Psychology, 24,* 108–122.

Weisz, J. R. (1979). Perceived control and learned helplessness among mentally retarded and nonretarded children: A developmental analysis. *Developmental Psychology, 15,* 311–319.

Weisz, J. R., & Achenbach, T. M. (1975). Effects of IQ and mental age on hypothesis behavior in normal and retarded children. *Developmental Psychology, 11,* 304–310.

Weisz, J. R., & Yeates, K. O. (1981). Cognitive development in retarded and nonretarded persons: Piagetian tests of the similar structure hypothesis. *Psychological Bulletin, 90,* 153–178.

Weisz, J. R., Yeates, K. O., & Zigler, E. (1982). Piagetian evidence and the developmental-difference controversy. In E. Zigler & D. Balla (Eds.), *Mental retardation: The developmental-difference controversy* (pp. 213–276). Hillsdale, NJ: Erlbaum.

Weisz, J. R., & Zigler, E. (1979). Cognitive development in retarded and nonretarded persons: Piagetian tests of the similar sequence hypothesis. *Psychological Bulletin, 86,* 831–851.

Winters, J. J., & Attlee, L. C. (1974). The von Restorff effect and measures of organization in normals and retardates. *Memory & Cognition, 2,* 301–305.

Winters, J. J., Attlee, L. C., & Harvey, F. (1974). Paired-associate learning of EMR adolescents and nonretarded children as a function of methods of presentation and training. *American Journal of Mental Deficiency, 79,* 70–76.

Yando, R., & Zigler, E. (1971). Outerdirectedness in the problem-solving of institutionalized and noninstitutionalized normal and retarded children. *Developmental Psychology, 4,* 277–288.

Zigler, E. (1966). Research on personality structure in the retardate. In N. R. Ellis (Ed.), *International review of research in mental retardation* (Vol. 1, pp. 77–108). New York: Academic Press.

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

Zigler, E., & Balla, D. A. (1977). Impact of institutional experience on the behavior and development of retarded persons. *American Journal of Mental Deficiency, 82,* 1–11.

Zigler, E., & Balla, D. A. (1982). *Mental retardation: The developmental-difference controversy.* Hillsdale, NJ: Erlbaum.

Zigler, E., Levine, J., & Gould L. (1966). The humor response of normal, institutionalized retarded, and noninstitutionalized retarded children. *American Journal of Mental Deficiency, 71,* 472–480.

Received September 10, 1985
Revision received December 11, 1985 ■