Preventing and Curing Mental Retardation by Behavioral Intervention: An Evaluation of Some Claims

HERMAN H. SPITZ
Edward R. Johnstone Training & Research Center
Bordentown, NJ

Statements are being made that by the judicious application of known behavioral and cognitive techniques we now have the means to raise intelligence and prevent or cure most mental retardation. As one example, it has been claimed that this goal can be achieved by combining infant, preschool, and elementary school intervention programs that have already proven effective. However, examination of the cited programs fails to reveal any evidence that they are effective. In view of the cycle of optimism and disillusionment that pervades the history of this field, caution and restraint are recommended.

A recent book titled Curative Aspects of Mental Retardation: Biomedical and Behavioral Advances (Menolascino, Neman, & Stark, 1983) recounts the proceedings of a national conference cosponsored by the Association for Retarded Citizens of the United States and the President’s Committee on Mental Retardation. In the section on behavioral advances, a number of optimistic statements are made about the possibility not only of preventing, but also of curing, mental retardation by application of current knowledge accumulated by behavioral and cognitive psychologists. According to one of the participants, 75% to 85% of all mental retardation “is a function of limited opportunities stemming primarily from sociocultural handicaps,” and rather than being the province of the biomedical sciences, “the strategy for the prevention of mild and moderate retardation must be founded on knowledge from the behavioral sciences” (Bijou, 1983, p. 230).

Another participant, allowing that “wishfulness is sometimes remarkably heuristic . . . presents an optimistic case for the heady possibility that behavioral science is on the threshold of delivering ways to improve thinking to the extent that we will be able to claim cures for the cognitive deficits that lead to mental retardation” (Butterfield, 1983, p. 203). A discussant agrees that “modern training techniques do work. There is increasing acceptance of the view that

Correspondence and requests for reprints should be addressed to Herman H. Spitz, Edward R. Johnstone Training & Research Center, Bordentown, NJ 08505.
these considerations fully employ the posture of the cure of mental retardation" (Berkson, 1983, p. 245).

In his introduction to the behavioral section, Stark (1983) notes that "intervention programs (with mildly regarded youngsters) have often resulted in essentially normal levels of functioning" and that therefore "if the procedures are applied at a young enough age, we are now able to cure this very large grouping (75%–85%) of the entire population of mentally retarded individuals! That is, we currently have the behavioral technology and knowledge to move individuals from the 'retarded functioning range' to an average intellectual and social-adaptive range via environmental enrichment and engineering" (p. 202, original italics).

Is this true? Do we possess the means to both prevent and cure most mental retardation by the use of behavioral engineering? In his contribution, Sidney Bijou (1983) suggests that this can be accomplished by exposing high-risk children to a combination of three programs that have proven successful: (1) a parent training program starting when children enroll in the preschool at 18 months of age, and designed to teach the parents behavioral management techniques that will enhance the children's development; (2) a preschool program to prepare the children for first grade; and (3) a special elementary school program to help the children master basic academic skills and acquire the motivation to use these skills independently.

The purpose of this paper is to look carefully at the three programs that serve as models for the three components of Dr. Bijou's (1983) integrated format and that, according to him,

have clearly demonstrated their effectiveness: the Portage Project for training parents to train their handicapped child in the home, the behavior analysis preschool model for compensatory education for the handicapped, and the Direct Instruction Model for comprehensive educational intervention with the disadvantaged. (p. 238)

THE PORTAGE PROJECT

Great success has been reported for the Portage Project, initiated in 1969 with a grant from the U.S. Office of Education, and replicated in numerous sites in the United States and throughout the world. Consider the following quotation:

The Portage data show that the average Cattell Infant Scale and Stanford-Binet Intelligence Scale IQ of the children upon enrollment was 75. After an 8-month period of training there was an average gain of 18.3 IQ points. (Bijou, 1983, p. 231)

A gain of 18 IQ points in eight months is remarkable and surely requires confirmation and careful examination of the data. The cited reference is to Shearer and Shearer (1976) and, a few sentences later, to Peniston (1972). The 1976 Shearer
and Shearer paper noted that the "greater gains made by the Portage Project children [when compared with randomly selected children from the local programs for the disadvantaged] in the areas of mental age, IQ, language, academic, and socialization skills were statistically significant" (p. 348), and cited Peniston (1972) as support for this statement. To back their statement that "the mean gain in IQ scores on the Stanford-Binet was 18.3" (p. 348), Shearer and Shearer cite their own earlier paper (1972). But the earlier paper merely repeats the above claims, again referring the reader to Peniston (1972). It is, then, to the Peniston paper that the reader must turn for the necessary information to make an informed judgment.

The Peniston paper (1972) is an unpublished manuscript. Among other results, the mean gain in mental age and IQ of 36 children randomly selected from the Portage Project (described as including multiply handicapped preschool children) are compared with the scores of 27 randomly selected, physically normal children enrolled in Head Start. All 63 children were given the "Cattell Infant Scale of [sic] Stanford-Binet Intelligence Scale, the Alpern-Boll Developmental Skills Age Inventory and the Gesell Developmental Schedule" (p. 23). Multiple analysis of covariance was applied to the data because "there was no assurance that the experimental and control pupils would be similar at the beginning of the study" (p. 23).

Discussing the limitations of the study, Peniston "wishes to caution the reader against the validity of the test scores in this research project due to the use of Home Trainers as competent examiners in the administration and scoring procedures of psychological tests, who in essence, are not psychometrists or psychologists" (p. 25). Elsewhere in the paper, Home Trainers are described as teachers who "teach the mothers how to teach developmental skills" (p. 11). In other words, the tests were administered by the teachers.

The reference to the "Cattell Infant Scale of Stanford Binet" is apparently not entirely a misprint. The dependent variable was in fact "the Combined Form of the Cattell-Binet (Form L-M, Stanford-Binet, and the Cattell Infant Scale)" (p. 33). How and why these scores were combined, and how this combination takes into account the differences in each test’s standardization, is nowhere mentioned.

There is one other limitation mentioned by Peniston that is especially important. The "investigator did not use a longitudinal . . . design following the same individuals over time (number of years) to measure change. Instead, the investigator is taking cross-sectional aspects of different multiply handicapped preschool children at different times and inferring change" (p. 26).

Just as perplexing are the results that Peniston presented. In Table 1, in which the results of the covariance analysis are applied to the posttest combined Cattell-Binet mental age scores of the two groups, the mean score for the Portage children is 383.00 (whatever that means) and for the comparison group, 475.26; the comparison group scored higher than the Portage group. But in the next column, "Adjusted Means" are presented that are 468.09 for the Portage group
and 361.81 for the comparison group, completely reversing the results. An adjustment that changes group differences from 92 points in one direction to 106 in the other is certainly suspect, for there is no analysis of covariance that can legitimately rectify such extreme sampling disparities, even if analysis of covariance was legitimate for data that are not longitudinal.

Similarly, in Table 2, the unadjusted, combined mean posttest IQs are given as 91.5 for the Portage and 105.4 for the comparison groups, whereas the adjusted means are reversed: 104.3 and 88.3, respectively. This, despite the fact that Peniston had stated that "the IQ's of all students were comparable, in both groups" (p. 19). Most of the other results are similar: Differences that favor the comparison group are reversed by "adjusted means."

In a 1984 paper, Cochran and Shearer present a table showing a mean pretest Stanford-Binet IQ of 77.1 and, after 9 months, a mean posttest IQ of 95.4 for "57 of the 60 enrolled children" (p. 105). This frequently cited 18.3-point gain in IQ is from the evaluation of the first year of operation, 1970–1971, whereas Peniston's evaluation had been for 1971–1972, apparently with a new group of children. In any case, for the first group of Portage children there was no comparison group and no information on who did the pre- and posttesting. The pretest mean IQ of 77 indicates that most of the Portage children were not retarded. In fact, the project is described as serving children diagnosed as having behavior problems, being emotionally disturbed, mentally retarded, physically handicapped, culturally deprived, or vision, speech, language, or hearing impaired, including children with any combination of these disabilities. Peniston had described the intelligence of both groups in his comparative study as ranging from mental defective to above normal. In order to know whether the project improved the intelligence of mentally retarded children, their scores would have to be extracted from the total group, something that to my knowledge has not been done. In the analyses of other years of program operations, no comparison groups were used.

Many questions must be raised about evaluations that continue to cite evidence from early studies that cannot meet even minimal standards of established procedures for testing the effects of intervention. The Portage Project may be a fine program for early intervention, but I can find no good, objective evidence that it prevents or cures mental retardation.

THE PRESCHOOL PROGRAMS

The second component of the intervention triad consists of a variation of four preschool programs. Their relevance for the claims that mental retardation can be prevented or cured will be summarized.

The Illinois and Arizona Projects

This program was originally developed by Dr. Bijou at the University of Illinois, and then extended by him to the University of Arizona (Bijou, 1972; 1977).
Designed for 2- to 6-year-olds, it provides typical nursery school activities plus individualized teaching in self-care, language, and "preacademic skills." Its distinguishing feature is its use of operant conditioning principles to shape the children's behavior. In the 1972 paper, in a section titled Formal Research, readers are referred to a study in which a retarded boy's social and eating behavior was improved, and to another study in which "contingent management techniques" increased the length of time that children of average and above-average intelligence worked productively in the classroom. The 1977 papers describes an "interactive model of child development." There are references to Project Follow Through, and a reference to Shearer and Shearer (1972), but no original data are presented. On the basis of the cited references, this program provided no evidence that it can either prevent or cure mental retardation.

The Kansas Preschool Program
This program also used operant conditioning to remedy the behavior problems of children labelled retarded, autistic, or schizophrenic, and to allow them to complete their schooling (Baer et al., 1976). A token-mediated reinforcement system was used to induce four children (whose characteristics are not given) to complete their academic activities. The aggressive behavior of an 8-year-old Down's syndrome child was eliminated, and the nervous glancing behavior of a young boy with autistic mannerisms was reduced. Again, however, there was no attempt to assess changes in the intellectual level of retarded children.

It might be argued that behaviorists are not interested in assessing gains on intelligence tests, if it were not for the fact that they have not always objected to their use.

While a functional-behavioral approach emphasizes descriptive observational methods of assessing individual behavior, it may also utilize information from standardized psychological tests. . . . The results of an intelligence test . . . may be used as indicators of the cognitive repertoires of the average child in a given elementary school grade. (Bijou & Peterson, 1971, p. 78)

More importantly, if assertions are being made that techniques are available that will move retarded persons into the average intellectual range, then an assessment of intellectual functioning is mandatory.

The Oregon Program
The Data Based Classroom Model of the Teaching Research Infant and Child Center in Monmouth, OR, also uses operant conditioning principles (Fredericks et al., 1982. Note that this is the fourth edition, whereas Bijou cited the second edition). It was originally designed to serve children up to 12 years of age who are moderately or severely handicapped, including children who are retarded, autistic, multiply handicapped (such as deaf/blind), or who have cerebral palsy. The teacher ""is utilized primarily as a manager of the learning environment"" (p.
and reinforcements are selectively given to shape a range of skills from self-help to academic.

Most of the Fredericks et al. (1982) monograph is a description of the program, but the last chapter is titled "Results." The first of three results is from a statewide evaluation of seven classes for moderately and severely handicapped children. The total mean gains made by children in the Teaching Research classes exceeded those of the six other classes on seven skills (receptive language, expressive language, feeding, dressing, personal hygiene, motor skills, and physical fitness), although only in expressive language did the gains made by the Teaching Research classes lead all other groups. They had the second largest gains in four of the other skills. Note, however, that we do not know the characteristics of the children in the seven groups, and consequently cannot gauge the extent to which the gains may have reflected unequal distributions of children with physical, rather than mental, disabilities.

In another comparison, 20 students were randomly selected from the Testing Research preschool classrooms, and 10 from replication sites (other comparable programs not using the Teacher Research Model). "An average of 6.44 skills per month were acquired by the children at Teaching Research and an average of 9.01 skills per month were acquired by the children in the replication sites" (Fredericks et al., 1982, p. 242), an obvious improvement over the less than 1 skill a month that had been acquired by the children from the time they entered the program until they started the instructional program. The larger gains made by children in replication sites were attributed to the smaller percentage of severely and profoundly handicapped in that group than in the Teaching Research group. The largest gains were in areas labelled Language, Motor, and Self-Help. Very few children made any gains in the Cognitive area (knowing personal information such as name, age, birthdate, telephone number, address, and so on; reading; writing; and number concepts), perhaps because the students were too young (their ages are not given).

A third assessment consisted of a random sample of 141 children from classrooms where teachers had been trained in the Teaching Research Model. Each child was measured on the Student Progress Record, a nonstandardized instrument that assesses student progress in 13 curricular areas. They were compared with a randomly selected group of 141 students from classes with teachers who were not trained in the Teaching Research Model. The former group gained 2.28 points more than did the latter group.

This is the extent of the assessment provided by Fredericks et al. (1982). There is no evidence that his program has had any significant impact on the intelligence of retarded persons.

The Washington Program
The final preschool project to which Dr. Bijou referred us is the Down's Syndrome Program of the Model Preschool Center at the University of Washington in Seattle (Hayden & Haring, 1976). Infants and children ranging in age from
birth to 6 years participated in year-round classes, the goal being to make the
development of Down's syndrome children more nearly approximate the sequen-
tial development of nonretarded children. According to the report, the program
was remarkably successful. "Data for the 1972–1973 school year show that all
the infants enrolled during that year met and maintained developmental objec-
tives for their various ages and that four of the children were above age level in
some skills" (Hayden & Haring, 1976, p. 593). In addition, children in the
preschool and advanced preschool programs also made excellent gains. Over the
course of a year, the mean Peabody Picture Vocabulary Test IQ of children rose
from 75 to 83.

These are well above the IQs to be expected from a group of Down's syn-
drome children and, to add to the optimism, Hayden and Haring (1977) also
reported that, on the Down's syndrome Performance Inventory (an unstandar-
dized instrument), children who had been in both the preschool and primary level
programs were, by the time they were 8 years old, developing at a normal rate.
But, as the authors noted, "the question concerning Down's syndrome chil-
dren's intellectual progress is whether the early gains that so many of them show
in response to educational programming can be maintained" (Hayden & Haring,
1976, p. 597, original italics).

By way of an answer, in her dissertation at the University of Washington, du
Verglas (1984/1985) did an 8-year follow up of 11 children who had attended the
Model Preschool Program and the special grade school program. (The 5 other
children were omitted for various reasons, including age discrepancies and miss-
ing scores.) When they had a mean age of 13.6 years, she gave them a number of
tests, including the Stanford-Binet, on which their average IQ was 49. This was
8 points higher than the mean IQ of a control group of 11 Down's syndrome
children who had not attended the University of Washington early-intervention
program. However, for only 7 children in each group was there a Stanford-Binet
IQ available from 1975, when the children were 5 and 6 years of age. Based on
these data, the control children rose from a mean IQ of 34 to a mean of 42 over
that 8-year period, whereas the experimental group dropped from 52 to 50. It is
possible, then, that the differences in favor of the experimental group in 1983
merely reflected initial differences between the groups, but in any case it seems
clear that the earlier hope that early intervention followed by special school
programs would allow the children to develop at a normal maturational rate has
not been realized. Of course, Bijou (1983) could not have known about du
Verglas's results when he included the University of Washington program in his
list of exemplary models.

THE DIRECT INSTRUCTION MODEL

The third component of the intervention triad is the elementary school program
designed to teach children from first to fourth grade "the social and academic
skills and knowledge . . . that constitute intelligent behavior" (Bijou, 1983, p.
The prototype program is the Direct Instruction Model, originally developed by Carl Bereiter and Siegfried Engelmann (1966) at the University of Illinois as a preschool teaching program based on Skinnerian reinforcement techniques, and later—after Bereiter left Illinois—developed into the DISTAR (Direct Instruction System in Arithmetic and Reading) and other Direct Instruction Systems by Engelmann and Wesley Becker, both of whom later moved to the University of Oregon. Economically disadvantaged children were the target population.

The Direct Instruction program, like most of the cited programs, is strictly behavioristic.

Intelligent behavior is operant behavior; therefore, it is learned and can be taught. . . . Learning rate is a function of teaching technology. . . . Thinking and related covert cognitive processes can be taught first as overt (usually verbal) processes. . . . If the student fails, do not blame the student; diagnose the teaching history. The teaching sequences control what can be learned. (Becker, Engelmann, Carnine, & Maggs, 1982, p. 154)

A number of sites using the Direct Instruction Model were included in the well-known Abt Associates, Inc. evaluation of children in Project Follow Through who had completed third grade (e.g., Bock, Stebbins, & Proper, 1977). The evaluation noted that children in the Direct Instruction Model fared better on measures of basic skills than did children in other Follow Through programs. However, results were not uniform across the different sites where the Direct Instruction Model was used. Indeed, differences in performance at the different sites had the greatest influence on the results, a finding also applicable to the other Follow Through models. Furthermore, on the only test of general intelligence used in the Abt evaluation (Raven's Coloured Progressive Matrices Test), at most sites the Direct Instruction groups scored either no better or reliably worse than did similarly disadvantaged nontreatment comparison groups.

Becker and Carnine (1981) took issue with a number of the Abt report's conclusions. Among other things, they noted that significant gains on the Slosson Intelligence Test, given under the supervision of the project staff, were maintained through the third grade. However, although short-term gains in the mean IQ of children in Direct Instruction programs have been reported, these gains have not been maintained in later testing (Karnes, Shwedel, & Williams, 1983; Miller & Bizzell, 1983). Similarly, after the children left the Direct Instruction programs and reentered regular classrooms, their academic achievement scores dropped appreciably as they progressed through school, an outcome that Becker, Engelmann, Carnine, and Rhine (1981) attributed to ineffective teaching.

Most of the children who participated in the above studies were not mentally retarded (i.e., did not have IQ < 70). Maggs and his colleagues, on the other hand, have applied the DISTAR programs specifically to retarded groups. In one
study, 14 6- to 14-year-old moderately and severely retarded children received the program 1 hr a day for 2 years (Maggs & Morath, 1976). Their mean mental age (MA) rose 22.5 months, compared to a control group's gain of only 7.5 months, which translates into a gain of about 10 points for the treatment relative to the control group. In another study, 12 6- to 12-year-old children, with a mean Stanford-Binet IQ of 45 (adjusted for regression artifacts), gained 6 IQ points following a 5-year application of DISTAR Reading and Language programs (Gersten & Maggs, 1982). These gains are typical of early intervention programs, and in the past have all but evaporated in follow-up tests.

CONCLUDING REMARKS

From the time of Itard’s celebrated encounter with the Wild Boy there have been repeated attempts to use special pedagogical and psychological techniques to eliminate the intellectual deficiency that characterizes mental retardation. This fascinating history is peppered with exaggerated claims of success by both well-meaning and not so well-meaning individuals (Spitz, 1986). Each new generation believes that new knowledge and new techniques permit them to succeed where “Seguin and so many others doing a much more thorough job of special training and often in the best of home environments failed so dismally” (Kuhlmann, 1940, p. 19). Thus far, these claims have led only to frustration and disillusionment.

Few people have the time, the facilities, or the inclination to search carefully for the basis of similar claims made today by many respected psychologists (primarily, but not only, those associated with the behaviorist philosophy). This brief review is an example of these claims. It is conceivable that, although no single program has yet proven to be successful, a combination of programs applied from infancy through elementary school will finally succeed. But the history of the care and treatment of mentally retarded persons should be warning enough that extra caution should be taken before announcing that the integration of existing programs will alter parent and teacher practices “so that by the completion of the fourth grade the children involved will be able to demonstrate on objective tests intellectual, personal, and academic competencies within the normal range of development” (Bijou, 1983, p. 238). Unverified claims of success do nothing to enhance the reputation of psychology as a science, whereas they add to the burden of frustration and guilt of parents who compare the unrealistic pronouncements of experts to the actual progress of their own retarded children.

REFERENCES


