[MS 0087.003.001.001]

With permission from IRCD BULLETIN, Eric Information Retrieval Center on the Disadvantaged, Horace Mann-Lincoln Institute, Teachers College, Columbia University, 525 West 120 Street, N. Y., N. Y. 10027

From Vol. VI, No. 5
December, 1970

Compensatory Education: Evaluation in Perspective

Edmund W. Gordon

In the mid 1960's the federal government began its large-scale intervention in the development and education of poor children. At the time, a preeminent educational research scientist—one of the most distinguished in the nation—reportedly declined to participate in an evaluation of the government's premier effort. Further clarification of activities and aspirations was necessary, this scholar contended, before criteria could be set and evaluation could occur. His pessimistic view did not deter an army of able, as well as not so able, investigators from rushing to evaluate the impact of Head Start and other programs of compensatory education. This is not to condemn them for their courage, or perhaps even their opportunism or recklessness. The author of this article marched along in the front ranks and even barked out a few of the orders for what proved to be rather futile skirmishes. However, it appears that the pessimistic prophet was by no means wrong—simply unheeded.

During the past five years, more than $10 billion has been invested in the education of poor and minority group members and at least $75 million has been spent on evaluations and special research projects. Despite this enormous expenditure, we are still not able to make definitive statements concerning the value of compensatory education. Even those of us who have been the most enthusiastic advocates of the need for such efforts have to concede that evidence of the value of our efforts is modest, if it exists at all. Some critics are far more harsh in their condemnation of the endeavor. A few have predictably asserted that compensatory education has not worked because it was practiced on a population which is genetically inferior and, hence, incapable of adequate response.

The sparsity of evidence in support of compensatory education may have little to do with its value. Some studies indicate that considerable slippage occurs between the designation of a program as compensatory and the actual implementation of compensating elements in a child's education. As in the case of ethnic integration in public schools, it may be incorrect to conclude that the programs have not worked when in most instances they have not been tried. Yet, it is probably correct that some compensatory education is not very effective. The traditional use of drill and repetition in remedial education is not likely to improve achievement for disadvantaged children. Similarly, increasing guidance contacts from one to two or three per year or even providing more intensive personal counseling as a solitary treatment seems to make little difference. Reducing class size without changing what teachers do seems unimportant, and, similarly, modest increments in available materials have hardly brought about radical improvements. But these and other observations are impressions, partially supported by data, but generally inconclusive. There are few intensive, qualitative and systematic evaluations of compensatory education. Hard data are needed; solid research studies are required as a basis for policy decisions. We have instead an abundance of indefinite, conflicting and confusing studies. The value of compensatory education may be obfuscated, in part, because the practice of evaluative educational research is poor.

The weaknesses in the application of evaluative research to compensatory education partially stem from the complex political and economic circumstances under which these programs were initiated and developed. From their inception, programs involved large expenditures—often made for other than purely experimental educational reasons. Foundations, local and federal governments channeled more than $10 billion into the education of poor and minority group children. Some of the foundation efforts unfortunately seemed also to reflect a desire to establish organizational leadership, a domain of action, or a model program which would be identified with the foundation. The federal programs that succeeded the work of the foundations were subject to a different set of pressures, mainly political concerns. Federal programs were in part responses to the rising demand for a social revolution, for the improvement of human rights, and for the increased development of underprivileged populations. For a while, it seemed more important, politically to act, to be identified with the effort to do something, than to act wisely. There was little time for planning. With large sums of money being spent, and with political objectives clearly the motive, "pork barreling" and politically determined distribution of funds naturally developed. To maintain some semblance of responsible government, the executive branch began to press for evaluation data—to prove favored programs successful and to provide the basis for reducing or eliminating unpopular activities. Initially the legislature was not greatly concerned with evaluation. Rather, the executive branch initiated the evaluation of the impact of compensatory education.

In this context, it is easy to see that large expenditures hastily appropriated for new programs, political pressures for change and a piece of the action, and the demand for immediate proof of impact have complicated the evaluation of the effectiveness of these programs. Evans (Office of Education) and Schiller (Office of Economic Opportunity) discuss the pressures they were under while designing and implementing Head Start:

> *Unfortunately, the political process is not orderly, scheduled, or rational. Crests of public and congressional support for social action programs often swell quickly and with little anticipation. Once legislation is enacted, the pressures on administrators for swift program implementation are intense. In these circumstances—which are the rule rather than the exception—pleas that the program should be implemented carefully, along the lines of a true experiment with random assignment of subjects so that we can confidently evaluate the program's effectiveness, are bound to be ignored.*

The results of such conditions were program and research designs based upon well intended but precipitous decisions. Often when evaluations were attempted after the fact, it was discovered that the original design had been inadequate.

In addition, as Caro observes, the clients of such programs can present a sensitive and difficult situation for the evaluative research. He continues:

> *Even though evaluative researchers may firmly believe that their efforts contribute ultimately to the cause of the poor, minority activists may confront them with great hostility . . . Preoccupied with the immediate, tangible, dramatic, and personal, the minority activist is likely to be impatient with the evaluator's concern with the future, abstract concepts, orderly procedures, and impersonal forces.*

Quite apart from the problems related to the conditions under which programs were initiated and conducted are the problems of evaluative research in general. Here one often finds a low level of expertise and inadequately developed methods. The best educational research scientists often choose to work with basic problems in areas such as child development, learning, linguistics, rather than with evaluative research. Evaluative and field research have only recently gained in respect and demand among educators and the public. Consequently, high demand has been suddenly created in a field with insufficient expertise. Although many good research scientists were drawn into evaluation, they could not readily transfer their research competence to the new situation. Indeed, given their experience in controlled laboratory settings, the problems of evaluative and field research may have been more difficult for them than for some less experienced investigators.

(Continued on page 3)

Compensatory Education (Continued from page 2)

In the conduct of evaluative research, one can distinguish three approaches or three levels of concern. The first attempts to discover whether or not a particular intervention program is effective: Are developmental and learning processes accelerated following the application of a particular teaching method, curriculum, etc.? The second level of concern is comparative: Is the particular intervention more effective than other known methods? The third level is explanatory: What is the nature of the relationship between specific intervention methods and specific associated changes in behavior? Most evaluative research has been directed at the first two levels of concern. The third level, however, is the most important. By answering questions on this level, one can establish a rational basis for action and begin to specify treatments in relation to known characteristics of the children to be served. At the third level the distinction between basic research and evaluative research collapses. The questions posed demand a quality of design which is appropriate to basic research but which can also serve the purposes of evaluation. Unfortunately, evaluative research of this quality has seldom been applied to compensatory education.

All of these approaches are made more complex by technical operational problems. The more compensatory education programs approach laboratory experimental conditions, the more one can discover what, how and why certain educational treatments alter educational underdevelopment. Yet, numerous obstacles stand in the way of establishing the necessary degree of precision and control in isolating variables and discovering the effectiveness of specific treatments.

One such obstacle involves difficulties in the utilization of an adequate method for selecting subjects. As Campbell and Erlebacher point out, "experimental" subjects are often not selected on a random basis. While the "control" group is selected to closely match the experimental group according to various indices, the control group is too often different from the experimental group in crucial aspects, however small a degree. Without random selection of subjects, the results of a program may reflect differences in the development of two populations—differences which are unrelated to the experimental treatment in question. In addition, matching procedures may produce regression artifacts. As for analysis of covariance and partial correlation, such biases may occur both where pretest scores are available and in after-the-fact studies. Campbell and Erlebacher propose true experiments in which randomization of subjects will avoid difficulties that previous quasi-experimental designs have encountered. However, parental objections, coupled with political pressures, have made large-scale application of random assignment of subjects impossible. Controlled comparative studies of this sort are often resisted by communities who will not accept arbitrary selection of subjects for experimentation when everyone wants the benefit of special treatment.

Another difficulty in establishing comparable experimental and control groups can be attributed to the influence of what has been called the radiation effect. Even if the two groups are initially "comparable," the effect of experimentation on the experimental subjects is radiated onto their families, siblings and eventually onto the control subjects if there is any contact, direct or indirect, between these several groups. Susan Grey (1966) reported the confounding impact of preschool on the experimental children's families and even on other members of the community in which they lived. Reporting on the Early Training Project, Grey found that at the end of each school year the controls caught up to the gains made during the summer by the experimental group. However, another control group in a town 60 miles away did not show such gains. In addition, untreated younger brothers and sisters of experimental subjects were observed to make unusual progress, no doubt as a result of the influence of the program on their parents or siblings (Kohlberg, 1966). Obviously, control subjects should be selected in a manner such that they can in no way be affected by the experimental treatment. However, this condition is increasingly difficult to maintain in large-scale field studies and demonstration projects.

In addition, investigators have discovered other effects that are associated with an intervention program—efforts which again are not direct results of the treatment itself. Rosenthal reported that a teacher's expectations can have an important influence on the performance of students. Shephard reported a similar experience in the early stages of his work in St. Louis. Where the teacher's expectation of the child's performance is high, the child is likely to show high achievement. Where expectations are low, achievement tends to be low. Consequently, in any compensatory education program, the expectations of the subjects' teachers may influence their subsequent performance. The Hawthorne effect, in which the mere fact of experimentation or altered learning conditions may cause a temporary change in performance, unrelated to the specific intervention method applied, can also color the results. In the evaluation of compensatory education, such interferences have not been identified or controlled for; hence the real consequences of the various treatments cannot be determined from these studies.

There are still more problems referrable to evaluative research design which confuse, distort or limit the initial data as well as subsequent findings. Most evaluations of compensatory education studies depend excessively on static variables and quantitative measures to the neglect of the process variables and the qualitative analysis of behavior, circumstances and conditions. This dependence on quantitative measures of status to the neglect of qualitative study of process not only opens these works to questions related to the validity of the measurement instruments; it also ignores the growing appreciation of situational and transactional factors as determinants of function. Compensatory education programs under study include and affect a wide variety of independent and dependent variables which are insufficiently accounted for in the more narrowly designed evaluation studies that have dominated the field to date.

This rather static approach to assessment has led investigators to view pupil characteristics which differ from some presumed norm as negative, as well as to consider any correlation between these negative characteristics and learning dysfunction as support for a deficits theory of intervention. In practice this has meant that researchers see all differences between the target populations and the standard group as deficits to be overcome rather than characteristics to be utilized and developed.

Relationships between stereotypical and fairly static input and output variables (usually isolated in pairs) are investigated; no attention is paid to the complex dialectic relationships between patterns of dependent variables and patterns of independent variables, many of which may be idiosyncratic to individuals and situations. These inadequate attempts at the assessment and treatment of pupil characteristics are often accompanied by an even less adequate appraisal of program variables. In practically all of the so-called national impact studies and most of the evaluation of specific programs little or no attention is paid to the fact that intervention treatment is uneven and control of that treatment almost nonexistent. When national impact data are pooled we could easily have results which show no effect, if the effect of specified programs with positive impact is cancelled out by other programs with no positive effect. Even more serious is the apparent disregard of our growing conviction that individual pupils respond differentially to treatments. When mean changes in status are used as the indices to outcome, again we may have negative responders cancelling positive responders to indicate no effect-even though the treatment may be highly effective for specific individuals under specific circumstances.

Several possible explanations have been advanced to illustrate how these confusing data can be interpreted to demonstrate the programs' ineffectiveness. The most extreme is the theory that the subjects involved are simply genetically inferior and not able to be brought up to hoped-for standards. Those who have attempted to advance such hypotheses have been blasted from all sides for the extremely questionable nature of their scientific "support," as well, of course, as for the dubious social value of advancing such theories at this point in the society's development, when they cannot be adequately proven.

However, whatever the range of possible interpretations of apparently discouraging data, what cannot be ignored is that far too many children from economically or ethnically disadvantaged groups are failing to master the traditional learning tasks of schooling. The problem

(Continued on page 4)

Compensatory Education (Continued from page 3)

is not only tragic, but staggeringly complex. Perhaps the most important response to the discouraging data presented by many evaluation reports, after allowing for many of the research problems already discussed, is a rigorous examination of the suitability of what is actually taking place in the schools.

Public schools as social institutions have never had to assume responsibility for their failures. Only recently have observers begun to view and describe objectively some of the horrors that are perpetrated in the name of public education. We must come to grips with the problem of the utterly stultifying atmosphere of many classrooms, with the way in which rote learning and repetition discourage *real* learning; and we must also realize that discipline for discipline's sake serves the purpose of creating artificial order, but at the same time produces dull automatons instead of eager students, or turns the inmates of public schools against education, to their lifelong detriment.

Even where extraordinary programs of compensatory education have brought about some beneficial results, larger social factors may negate these results in the long run. Outside the classroom, disadvantaged children confront a society that is hostile to their healthy development. Learning in structured situations may be irrelevant in the context of their life outside the school. There is some evidence to suggest that ethnic, economic, or social integration does have beneficial effects on children whose background results in such school problems. Achievement levels have been shown to rise after desegregation in many schools, although the exact interplay of reactions leading to this result has not been conclusively determined. For example, improved teacher morale or other improved conditions brought about by the process of desegregation may result in an overall increase in the quality of education throughout the system. Other evidence points to the conclusion that integration on a social status group basis has beneficial effects for disadvantaged children when the majority of their peers in the school are from higher status groups. Even these results, however, are not sufficiently conclusive to provide a legitimate basis for large-scale generalizations. The problem is further complicated by the new renaissance in cultural nationalism among ethnic minorities, a movement which affects any assumptions to be made about ethnic integration and education. In a society which has alternately pushed ethnic separation or ethnic amalgamation and which has never truly accepted cultural and ethnic pluralism, blacks, chicanos, Puerto Ricans, and native Americans are insisting that the traditional public school is guilty not only of intellectual and social but also of cultural genocide for their children. There are class and caste conflicts to which insufficient attention has been given in the organization and delivery of educational services. If cultural and ethnic identification are important components of the learning experience, to ignore or demean them is poor education, at best. Even if these factors are sufficiently taken into account in the school, we are far from any guarantees that the society will honor such values outside the classroom. It is not at all clear that intensive, short-term in-school treatment can counter the negative, external forces working upon disadvantaged populations.

The schools face a difficult challenge if they are to make learning an exciting and stimulating experience, relevant and effective, for all their students from all cultural and social backgrounds. However, even meeting these criteria will not be enough. Educators still face the problem of matching the developmental patterns, learning styles and temperamental traits of individual learners to the educational experiences to which they are exposed. Many researchers have concentrated on differences in level of intellectual function, a concern reflected in the heavy emphasis on intelligence testing and the placement, even "tracking", of pupils based on these tests. This tradition has emphasized quantitative measurement, classification, and prediction to the neglect of qualitative measurement, description and prescription. These latter processes are clearly essential to the effective teaching of children who come to the schools with characteristics different from those of both their teachers and the other children to whom most teachers are accustomed. Research data indicate wide variations in patterns of intellectual and social function across and within sub-populations. Variations in function within privileged groups may be less important because of a variety of environmental factors which support adequate development and learning; however, among disadvantaged populations—where traditional forms of environmental support may be absent—attention to differential learning patterns may be crucial to adequate development. Understanding the role of one set of behaviors as facilitators of more comprehensive behaviors is at the heart of differential analysis of learner characteristics and differential design of learning experiences. Schooling for disadvantaged children—indeed, for all children in our schools—comes nowhere near meeting these implied criteria. Assessment technology has not seriously engaged the problem. Curriculum specialists are just beginning to face the task, in some of their work in individually prescribed learning.

The problems of social disadvantage in the society at large, and the failure of the schools to mold their practices to cultural differences and individual learning styles are not the only obstacles to successful compensatory education. Social disadvantagement gives rise to a variety of harmful health and nutritional problems which militate against healthy development and adequate utilization of educational opportunities. It is becoming increasingly recognized that low income results in poor health care and frequent malnutrition; these disadvantages are related to high risks for the pregnant mother and fetus, and for the child after birth, in terms of mortality or maldevelopment. Poor health conditions may result in either a direct impairment of the nervous system or an indirect interference with the learning process by a low level of energy or high level of distractibility. Such health-related conditions probably have a crucial effect on school and general social adjustment. It has now been shown that impaired health or organic dysfunction can influence school attendance, learning efficiency, developmental rate as well as personality development. Clearly, adequacy of health status and adequacy of health care in our society are influenced by adequacy of income. Thus poverty results in a number of conditions directly referrable to school success and to development in general.

Despite the many problems in the design, implementation and evaluation of compensatory education programs and the equivocal status of much of the evaluation effort, we are nonetheless constantly called upon to make judgments and policy decisions based upon the experiences so far. There are useful insights to be drawn from these experiences:

1. The search for the best or the generic treatment is clearly a futile search. Problems of human development and learning are so complex and conditions of life so varied that the chances of finding a curriculum which is universally superior are quite modest. In well designed and conducted studies comparing different approaches to early childhood education, differences in curriculum orientation seemed less important than the following factors: systematic planning, clear objectives, intensity of treatment, attention to individual needs and learning patterns, opportunities for individual and small group interaction, support in the home environment for the learning experiences provided at school and the presence of personnel committed to the pedagogical procedures prescribed. It seems that as these conditions are approached, no matter what the content or method, personal development and content mastery are advanced. Hard data in support of these conclusions are scarce since few studies have been designed to be particularly sensitive to this constellation of variables. Nonetheless logical and impressionistic evidence mounts in support of the validity of these observations.

2. Although the concept of individual differences has been with us for a long time, individualization is underrepresented in programs of treatment and evaluation of programs. Confusing interpretation of evaluation data may occur because of this neglect and the countertendency to generalize too freely. In a few longitudinal studies where impact on individuals (or on youngsters identified as having been exposed to known treatments over time) has been investigated, emerging achievement patterns are encouraging. There appear to be insufficient studies of highly sophisticated programs of individually prescribed learning experiences to draw definitive conclusions. Yet some of the more generalized individually prescribed instructional programs do seem to be widening the range of achievement among pupils so exposed. These generalized IPI programs are probably not the answer even though they represent an advance in educational technology. The

(Continued on page 5)

Compensatory Education (Continued from page 4)

true matching of pace, content and conditions of learning to the specific characteristics of each learner is not yet a part of even our highly experimental work. Insufficient progress in the qualitative analysis of learning behavior may be partially responsible for this situation. Such analysis is clearly prerequisite to any serious effort at achieving sophistication in the individualization of instruction and learning.

3. The absence of broader representation and utilization of the social sciences in the evaluation of compensatory education has contributed to the neglect of social psychological, social and political factors in these programs. Yet as important as the strictly pedagogical problems are, the politics of education delivery systems, the social psychology and political economy of education and the sociology of knowledge and learning share the stage with pedagogy in accounting for the success or failure of compensatory education. Whether we are considering the role of pupils in directing their own learning or the roles of parents and community in directing school policy, the influence of involvement, participation, commitment and values is so critical as to render much of our evaluation and our treatment useless unless we give these factors greater consideration. In the very inadequate studies of several informal schooling situations (storefront academies and the adult education programs of groups like the Black Panthers, Black Muslims, Young Lords, etc.) the blending of control, participation, politics, values and demonstrated change in opportunity structure begin to appear as important factors in educational rehabilitation. Unfortunately, the research and evaluation data that we have are not sufficient to erect guidelines or to draw firm conclusions but again impression and logic suggest that we should look to these concerns in our programs and evaluation.

This **Bulletin** was prepared pursuant to a contract with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their judgment in professional and technical matters. Points of view or opinions do not, therefore, necessarily represent official Office of Education position or policy.

NEW PRICE SCHEDULE
CURRENT INDEX TO JOURNALS IN EDUCATION

Since January 1, 1971, the following price schedule has been in effect for CIJE:

Monthly (12 issues) $39.00
Semiannual and Annual $40.00
Monthly, Semiannual & Annual $74.00
Annual (purchased singly) $29.00
Single monthly copies continue to be . . . . . $ 3.50 each

CIJE is a monthly cataloging and indexing publication for journal and periodical literature in the field of education. It is available from:

CCM Information Corporation
(A subsidiary of Crowell Collier and Macmillan, Inc.)
866 Third Avenue, Room 1126
New York, New York 10022.

HOW TO ORDER ERIC DOCUMENT REPRODUCTIONS

RIE BACK COLLECTIONS

Name	Total Fiche	Unit Price	Collection Price
Reports in Research in Education for 1966 & 67	4,426	$0.89	$ 394.00
Reports in Research in Education for 1968	13,326	$0.89	$1,187.00
Reports in Research in Education for 1969	15,899	$0.89	$1,416.00
Reports in Research in Education for 1970	16,188	$0.89	$1,441.00

SPECIAL COLLECTIONS

Name	Total Fiche	Unit Price	Collection Price
ERIC Catalog of Selected Documents on the Disadvantaged	2,740	$0.14	$ 384.00
Office of Education Research Reports, 1956-65	3,315	$0.14	$ 465.00
Selected Documents in Higher Education	1,258	$0.14	$ 177.00
Pacesetters in Innovation, Fiscal Year 1966	1,185	$0.14	$ 166.00
Pacesetters in Innovation, Fiscal Year 1967	1,437	$0.14	$ 202.00
Pacesetters in Innovation, Fiscal Year 1968	919	$0.14	$ 129.00
Manpower Research, Inventory for Fiscal Years 1966 & 67	653	$0.14	$ 92.00
Manpower Research, Inventory for Fiscal Year 1968	364	$0.14	$ 51.00
Manpower Research, Inventory for Fiscal Year 1969	473	$0.14	$ 67.00

Individual Orders

Microfiche (MF)

Microfiche for all reports announced in **Research in Education** as available from EDRS are available regardless of document size at $0.65 per title.

Hard Copy (HC)

Full Size paper copies are available according to the following graduated pricing table:

No. of Pages	Price
1-100	$3.29
101-200	6.58
201-300	9.87
Each additional 1-100 page increment	3.29

There is no handling charge. However, payment must accompany all orders under $10.00. Orders must be in writing, stating the ED numbers, type of reproduction (MF or HC), and the number of copies desired.

Address all orders to:

ERIC Document Reproduction Service
P.O. Drawer O
Bethesda, Maryland 20014

The above information represents a price change for all ERIC documents which became effective on February 21, 1971. All documents cited in the ERIC system in the past as well as those which will be cited in the future are governed by the new pricing. Appropriate adjustments should be made in all prices listed prior to the May 1971 issue of **Research in Education**.

(Continued on page 6)