DOCUMENT RESUME

ED 088 867 SP 007 840

AUTHOR Cohen, David, Ed.; Malin, Maureen, Ed.
TITLE Proceedings of the Symposium on Education Research and Development (Washington, D.C., December 11-12, 1972).
INSTITUTION National Inst. of Education (DHEW), Washington, D.C.
PUB DATE 72
GRANT OEG-0-71-3480(520)
NOTE 212p.

EDRS PRICE MF-$0.75 HC-$10.20
DESCRIPTORS *Conference Reports; *Educational Development; Educational Innovation; *Educational Research; Information Dissemination; Research; *Symposia

ABSTRACT

This symposium report presents the text of papers delivered at the symposium and outlines of discussions. The contents are as follows: a) "Knowledge in Education" by Sheldon H. White, Joan S. Bissell, and John Golenski; c) Comments on White-Bissell-Golenski Paper by David K. Cohen; c) General Discussion of White-Bissell-Golenski Paper and Cohen Comments; d) "The Conduct of Development in Education" by Richard E. Schutz; e) Comments on Schutz Paper by David R. Krathwohl; f) General Discussion of Schutz and Krathwohl Comments; g) "The Dissemination of Educational R&D Products: Research and Policy Issues for the Federal Government" by Daniel Weiler; h) Comments on Weiler Paper by David L. Clark; i) "Incentives for Innovation in the Public Schools" by John Pincus; j) Comments on Pincus Paper by David S. Mundel; k) General Discussion of Weiler and Pincus Papers and Clark and Mundel Comments; l) "Evaluating the National Institute of Education" by Senta A. Raizen; m) Comments on Raizen Paper by Alice M. Rivlin; and n) General Discussion of Raizen Paper and Rivlin Comments. A list of conference participants is included. (JA)

Proceedings of the Symposium on
Education Research and Development,
December 11-12, 1972,
Washington, D.C.*

Edited by David Cohen and Maureen Malin
Center for Educational Policy Research
Harvard Graduate School of Education
Cambridge, Massachusetts 02138

*Sponsored under NIE Grant # OEG-0-713480(520)

LIST OF PARTICIPANTS

David K. Cohen, Chairman
Harvard University School of Education
Cambridge, Massachusetts

Dr. Robert C. Andringa
Committee on Education & Labor
Washington, D.C.

Ms. Joan Bissell
Harvard University
School of Education
Cambridge, Massachusetts

Dr. David Clark, Dean
School of Education
Indiana University
Bloomington, Indiana

Mr. Tom Clemmons
National Institute of Education
Washington, D.C.

Mr. Jack Duncan, Counsel
Select Subcommittee on Education
Washington, D.C.

Mr. John Egermier
National Institute of Education
Washington, D.C.

Mr. Emerson Elliott
National Institute of Education
Washington, D.C.

Mr. Thomas Glennan
National Institute of Education
Washington, D.C.

Dr. David Kratwohl, Dean
School of Education
Syracuse University
Syracuse, New York

Mr. Mark Lohman
HEW Office of Planning & Evaluation
Washington, D.C.

Maureen Malin
Harvard University
School of Education
Cambridge, Massachusetts

Mr. Bernie Martin
National Institute of Education
Washington, D.C.

Dr. David Mundel
Assistant Professor of Government
Kennedy School
Harvard University
Cambridge, Massachusetts

Mr. John Pincus
The RAND Corporation
Santa Monica, California

Mr. Paul Pottinger
National Institute of Education
Washington, D.C.

Ms. Senta Raizin
The RAND Corporation
Washington, D.C.

Ms. Alice Rivlin, Senior Fellow
The Brookings Institution
Washington, D.C.

Dr. Richard Schutz
Southwest Regional Lab.
Los Alamitos, California

Mr. Michael Timpane
The Brookings Institution
Washington, D.C.

Mr. Mark Tucker
National Institute of Education
Washington, D.C.

Mr. Dan Weiler
The RAND Corporation
Santa Monica, California

Professor Sheldon White
Harvard University
School of Education
Cambridge, Massachusetts

TABLE OF CONTENTS

KNOWLEDGE IN EDUCATION[1]

Sheldon H. White, Joan S. Bissell, and John Golenski
Laboratory of Human Development
Harvard University

We live at a time and in a society where a substantial amount of resources have been committed to research on human behavior. All this seems natural to us, because it has been so within our memory and because society's support of such research rests on an assumption that most of us share -- to a greater or lesser degree -- the belief that careful scientific analysis of human behavior will sooner or later have value for humanity. Probably, most of us do not share the visions of total scientific utopias suggested by some in the past. And many of us have had our confidence in human research considerably shaken in the last several years. But most of us continue to hold a modest faith in the ultimate benefit of human research. We believe that psychological, sociological and anthropological research should sooner or later make contact with everyday human activities and everyday human concerns. It should be helpful to people in some way.

The ability of human research to deliver on that promissory note has recently come into serious question by government, by laymen, and by researchers themselves. The purpose of this essay is to address some of the questions that have been raised by analyzing the relationship of

---

[1]Paper presented at the NIE Symposium on Educational Research and Development, Washington, D.C., December 11 and 12, 1972.

research to education. The paper deals with issues concerning the character of knowledge in education, the channels of communication between research and education, and implications for the types of research activities in education that should be accorded high priority.

The importance of these issues is heightened by the recent creation of a new National Institute of Education. To be politically credible, the NIE must demonstrate an efficacy in research and development that, so far, has not been credited to previous arrangements for the support of R & D in education. There must be change in American education; the change should be facilitated through the R & D work of the NIE; the change should be taken to be progressive...to represent increases in educational effectiveness and efficiency.

The investment in educational R & D which the NIE represents is based on a faith that knowledge in education exists and that this knowledge can be extended selectively and directly through research support. It can be used as the basis for the introduction of rational innovations in education. The changes that can thereby be introduced into education will be surer, more powerful, and more demonstrable than those emanating from the traditional sources of testimony, debate, and advocacy.

The Types of Knowledge in Education

In examining the relationship of research to education, we must ask what is meant by research. We mean by research "investigation or experimentation aimed at the discovery and interpretation of facts, revision of accepted theories or laws in the light of new facts, or practical application of such new revised theories or laws."[2] Beginning with this

---

[2] Webster's Seventh New Collegiate Dictionary, Springfield, Mass.: G. & C. Merriam Co., 1963.

definition, we find that research in education consists of a three-part knowledge base: (a) Basic research -- findings and theories from a distributed mass of behavioral and social science research dealing with children's learning, thinking, motivation, emotion, socialization, abilities, and the like; (b) Problem and Program studies -- findings and theories from studies of the generalized and specialized educational needs of children, which provide evidence about the influence of various kinds of procedures, programs, and interventions on those needs; and (c) Analyses -- superordinate studies that try to draw together the knowledge of the knowledge base to bring forth issues and options for the management of American education at its various levels.

What are some examples of these different kinds of research -- these different elements of the educational knowledge base? Basic research consists of both theoretical and experimental endeavors. In this domain, we have the work of Freud, Lorenz, Piaget and Skinner; the work of Dewey, Bruner, Kohlberg, Erikson, Rogers, DeVore, Whiting and Cole. What, in contrast, are some examples of problem and program studies? James Coleman's work and the subsequent Moynihan, Mosteller et al. research on equality of educational opportunity and the Low and Spindler study of the child care arrangements of working mothers are two clear examples of problem studies. Each of these was a study of a major educational or social problem in which an attempt was made to break the problem into its component parts, to analyze its magnitude, its causes, its consequences, and the obstacles to solving it. What do we mean by program studies? These studies give evidence about the influence of various kinds of procedures, programs, and

interventions in relation to children's educational needs. Examples are Jeanne Chall's work on reading, the work of Suppes and Atkinson on Computer Aided Instruction, the activities that led to the development of PSSC Physics, Project Physics, Sesame Street, The Electric Company, and the Planned Variation exerpiments in Head Start and Follow Through. What about analyses? What are some examples of these superordinate studies that try to draw together elements of the knowledge base and apply them to decision options for the management of American education at its various levels? Among the best examples of this kind of work are the analyses by Harvey Averch et al., "How Effective is Schooling? A Critical Review and Synthesis of Research Findings;" the "Child Care" analysis by Alice Rivlin in Schultze et al., Setting National Priorities: The 1973 Budget; the HEW 'white paper,' "The Effectiveness of Compensatory Education;" and the analysis by S. White et al., Federal Programs for Young Children: Review and Recommendations.

Educational Research and Educational Practice

How good is this knowledge base as a source of vitality in American education? There is, today, some danger that the value of the existing knowlege base in this capacity can be both overestimated and underestimated. The creation of an NIE seems to represent a prima facie indication of an overestimation of the value of the knowledge base...because it seems to be widely believed that an NIE under good and intelligent management will bring forth what has for so long been missing, positive and conspicuous development in American education. There can be little doubt that an improvement

in the management of educational R & D is in order. Testimony is abundant that the previously-existing R & D management was of poor quality -- confused, poorly organized, tending to be insensitive to rudimentary issues of research quality. Still, improvements in the quality of the R & D management may bring about large changes in the productivity of the R & D system without necessarily bringing about a great increase in manifest, positive, conspicuous change in American education. We suspect that enlightened management of the NIE will have to be concerned to bring forth positive and conspicuous changes, while maintaining a concern for other R & D efforts that form the nexus for more subtle, but quite possibly more powerful, developments. The problem is that R & D in its most understood and visible form, both in general and with respect to education, is seen as a matter of invention and dissemination of procedural change. It is not understood as having influence through another extremely important process -- one of diffusion. Probably, we overestimate the ability of the knowledge base to engender change through the dissemination of procedures. Probably, we underestimate the ability of the knowledge base to engender change through diffusion.

Dissemination versus Diffusion

Our usual understanding of R & D centers on the creation and export of tangible goods or procedures that are of unequivocal -- or, at the very least, marketable -- benefit. A Salk vaccine is invented and moves out. Old goods and procedures are incessantly supplanted by better goods and procedures: the black-and-white TV by the color TV, the vacuum tube TV by the transistor TV and that by the circuit chip TV. Some may have long thoughts about the larger or ultimate value of the dynamism of

changes through R & D, but few would disagree that R & D as it is usually understood brings definite changes and beneficial changes within some pretty consensual and robust general understanding of what a benefit is.

It is not hard to find inventions and disseminations occurring in American education, but this kind of R & D dynamism in American education has usually seemed pallid and of somewhat uncertain value. At the turn of the century, there was the invention of psychometric procedures and a great flowering of child testing. Are tests worthwhile? Do they help education? Do they do more harm than good? In the 1950's there was the flowering of the curriculum-building movement and successive waves of new curricula were disseminated towards the educational system. Have the new curricula really changed anything? Is the New Math any better than the Old Math? Some believe that our educational methods have become better for some children, and that we must continue to search for new methods that are appropriate to the learning styles of other children, thereby diversifying American education. We agree with this. We believe that many significant options can be invented and disseminated, thereby widening the menu of alternatives available to children, teachers, parents, and others. Nevertheless, even with the creation and dissemination of alternatives in the American educational system, there will probably continue to be considerable feeling that these kinds of changes have not gotten us anywhere or are not getting us where we imagine we might go.

Probably, the basic uncertainty about procedural change stems ultimately from the lack of a differentiated, consensual, clear understanding of what the benefits of American education really are.

Everyone agrees that American education has value and gives value, but there is underneath this agreement an almost astonishing uncertainty and/or disagreement about wherein the benefits and values lie. We all agree that education is a good thing for the individual and for society but we do not agree about why it is good, how much it is good, or even about what kinds of changes in education will enhance what is good about it. This lack of agreement is fundamental for educational R & D, because it is not a disagreement based on an opposition of clearly different views of education -- say, a right-wing theory of education versus a left-wing theory of education. It is a lack of agreement resting on multiple vague notions about what education is and does, these notions distributed among tens of millions of significantly influential people, the disagreements entangled in emotions that make many of these people extremely sensitive and extremely concerned about any attempts to move or change educational practices.

Why do we not have any clear understanding of the benefits of American education? Probably because, for a long time, we did not need to have a very clear understanding. From the turn of the century to pretty near the present, it seemed sufficient and perhaps it was sufficient to understand that "More is better." The basic political question was pretty close to a pocketbook question: not whether to support increases in education but how much, from one year to the next, to bring about the increase. Under the "More is better" principle American society extended education increasingly to all children and to all classes of children. "Moreness" was supported in diverse senses -- more years of schooling for each child,

more higher education, more professionalization of the teacher, more extracurricular activities, new buildings, better teacher-pupil ratios.

The movement towards R & D in education is recent, since the second World War, and probably it has begun to arise and to be taken seriously because the "More is better" principle will no longer serve. Something... rising cost-pressures, the contemporary outcry for educational reform, the disaffection of the young with schools, a larger concern about issues of quality of life...has now caused us to ask, "What is good and what is bad?" We do not know. The question is, literally, somewhat new. Unless and until we can form some opinions about the question...at the very least, some reasonably focused disagreements...we cannot begin to try to make non-trivial changes in educational procedures.

One function of the NIE must be to create sharpened, more analyzed conceptions of the operations and consequences of American education and then to create, through diffusion, a sharper sense of the "What is good and what is bad" choices that are made in support, nonsupport, or re-direction of the American school system.

But there is a second sense, a more fundemental sense, in which the NIE must operate through diffusion rather than dissemination. To a significant extent, American education is vested in the belief systems and perceptions and intuitions of American teachers...checked and balanced by constraints vested in the belief systems of peers, administrators, and parents. Books, materials, and curricula are distributed to the teachers. There are achievement tests and other less obvious accountabilities. But there is a saying in American education that "Curriculum is what happens

after the teacher closes the door." It may be true that the matrix of intuitions and belief systems that surround the child -- the teacher and the parents and the principal -- form the most powerful determinant of what happens and what is possible in education. All these parties operate on normative, consensual ideas about what learning is, what education is supposed to bring about, what a child is like. It may be true that the single most powerful lever towards change in education comes in the shift of these normative ideas. One can suspect, though one cannot at this time prove, that these shifts come about through diffusions from the knowledge base. Unhappily, one can also suspect that these shifts as they occur may be almost undetectible and imperceptible. One has a shift from what was "obvious" in 1920 to what is "obvious" in 1940. Forget the fads... the "adjustment" fad of the 1930's subsiding to give way to the "creativity" fad of the 1960's. The fads may simply show on a surface level what may be significant on a deeper level...that opinions and attitudes about learning and education and children change and, as they do, education changes.

If our view of this diffusion mechanism is correct, it is a somewhat frustrating one. There is not much for an educational manager to get his hands on. There may be a benefit from educational R & D that may be virtually undetectible as a direct benefit from such R & D. But development by diffusion may be fundamental for development by dissemination.

Multiparadigmatic Knowledge of Education

We have said, thus far, that the NIE as an R & D institution will develop and exploit its knowledge base for two kinds of movement into practice: (1) invention and dissemination, and (2) diffusion. Invention and dissemination are well understood. Diffusion is poorly understood. Yet diffusion is surely of some importance and, perhaps, it is of very great importance. In the most basic of senses, teachers, parents, administrators, and others still need to "make up their minds" about education. They need to "make up their minds" about a set of interrelated issues:

Private versus public functions. American education seems to serve, not one, but a set of public functions. The public functions of education co-exist with private utilities and benefits. But the thrusts of the public functions are not identical with the thrusts of the private functions. The public functions of education have something to do with quality and quantity of labor supply, labor regulation, and political socialization. The private functions offered by education have something to do with pleasure, the feeling of growing competence, the winning of a competitive position in society. We need a clearer understanding of the public and the private functions.

Plurality of public functions. The set of public functions served by American education are not all purely educational. There is good reason to believe that the educational system was created in some part to serve other functions -- some having to do with custodial care, others having to do with socialization, others with "sorting" children into

various streams for adult life, and the like. If not all that looks like education is education, this creates significant limits and complications for an R & D function.

> "The continued assertion on the part of the educational research establishment that it is research and development which can improve and reform education, and the attendant pleas for more funds and better institutional structures, imply that that establishment is still taking the ascription of 'educational' problems at face value. That is, it is assuming that problems in the schools are, in fact, primarily 'educational' problems, and as such are amenable to 'solution' through 'educational' research.
>
> "However, it may well be that the major functions which schools perform in this society are not wholly or strictly 'educational' -- that they are primarily custodial, serving baby sitting needs for parents and keeping large numbers of young adults off the labor market; that they provide an acceptable rationale for sorting people into a purported meritocracy's analogue for a class structure; that they socialize youth for the myths of our culture if not for the realities, etc. If the schools are serving any of these purposes, then the fundamental dynamics which determine what happens in schools may be only partially 'educational.' Research and development designed to improve 'educational' inputs or the 'educational' operations of the schools will have limited power to affect the overall 'schooling process.'"[3]

Of course, one can do R & D on American education as a multipurposed system...attempting to optimize each purpose. To date, however, R & D in education has assumed what we have all tended to assume, that all education is education.

<u>Plurality of conceptions of education</u>. If now, we single out the purely educational components of education, we reach an important next

---

[3]Barbara Scott Nelson, Federally-funded research and development organizations for education: An analysis of the early years. Harvard Graduate School of Education, Qualifying Paper, 1972, p. 61.

level of the working agenda. Significant disagreements exist about what education as education is supposed to be bringing about. There is a multiplicity of vague theorizing about the goals and processes of education. A series of ideas about education coexist in our society, beliefs or belief-systems. Few are spelled out. Some seem to contradict others. Some are ideas about what education is doing, others notions about what it ought to be doing. Consider some of the present principles of education being advocated:

--Education prepares a child, or ought to be made to prepare a child, for a vocational place in society.

--Education should prepare a child for a world of change. Education should not so much prepare a child for a vocation, but prepare him to learn and to change and to grow to adjust in a changing society.

--Education should humanize the child, give him a rich appreciation of the world around him, sensitize him to the joys of learning. Through an ideal education, children should be brought towards literature, science, music, and the arts, towards being civilized in the deepest sense of the world.

--Education should teach children to think.

--Education should stimulate children to develop habits of mental discipline, effort, to attack and solve problems. Education should foster the qualities of mind on which a highly developed society depends. We do not so much teach children history and science as we teach them to think like a historian or think like a scientist.

--Education should stimulate creativity in children or, at all costs, not crush it.

--Education should transmit to children basic symbolic skills, and higher order skills with which to frame and solve problems.

--Education should foster the cognitive development of children... if not to accelerate it, then at least to broaden and deepen the movement of cognitive development in growth.

--Education should be concerned with the total personal development of children, not just his cognitive resources but his full set of personal resources upon which his cognitive ability ultimately depends.

One could go on and on listing principles or slogans or belief-systems. There is no official list. While there are contradictions in the above list, even the contradictions are elusive. There is, probably, a way of stating each of the above principles so that it is in harmony with, or in disagreement with, any other. As "soft" and as insubstantial as all these principles are, they are all "out there" today, being used as rationales for one or another kind of advocacy in education. What would a full set of such principles look like? The mind staggers. But it is an important fact about education today that the private and social and political acceptability of educational R & D today hinges on variegated matches and mismatches of innovations as they are perceived to go with or against such principles as these. One cannot isolate constituencies for the various beliefs. Probably all the beliefs are in coexistence in the average PTA meeting. Probably, any reasonably intelligent individual in confrontation with education and what it means is a constituency for many

for all of them. He holds them all. He knows that some contradict others, or should contradict others, but he does not quite know how. He is not so much ignorant of education, as he is complex in his thinking about it. We would argue that he is "state of the art" so far as the knowledge base about education goes today.

Scientific Paradigms in Education

We will argue that the "knowledge base" for education today is 'multiparadigmatic' at each of the several significant levels earlier discussed -- the basic level, the level of problem and program research and the level of analysis. There is knowledge, but the knowledge is of a special primitive form in which limited consensus is possible, and only limited acceptability for R & D efforts exists or can exist.

There is a radical redefinition of the nature of science and scientific thought afoot today. The revision has important implications for our contemporary understanding of education and children's learning. Indeed, one of the most important interpreters of this new view of scientific movement is Jean Piaget, whose 'genetic epistemology' is at one and at the same time an analysis of the forward movement of children's thought and an analysis of the forward movement of the history of science. But the bellwether for American discussions of the new conception of scientific movement has been Thomas Kuhn's important book, The Structure of Scientific Revolutions. It is Kuhn's argument that science moves forward only in part by the traditionally accepted evolutionary processes of research, analysis, fact-finding and consensual verification...what he terms the processes of "normal science." At times science moves forward by

revolution rather than evolution. There is a "paradigm shift" through which the very reality with which the community of science deals is changed.

> "Let us, therefore, now take it for granted that the differences between successive paradigms are both necessary and irreconcilable. Can we then say more explicitly what sorts of differences these are? The most apparent type has already been illustrated repeatedly. Successive paradigms tell us different things about the population of the universe and about that population's behavior. They differ, that is, about such questions as the existence of subatomic particles, the materiality of light, and the conservation of heat or of energy. These are the substantive differences between successive paradigms, and they require no further illustration. But paradigms differ in more than substance, for they are directed not only to nature but also back upon the science that produced them. They are the source of the methods, problem-field, and standards of solution accepted by any mature scientific community at any given time. As a result, the reception of a new paradigm often necessitates a redefinition of the corresponding science. Some old problems may be relegated to another science or declared entirely 'unscientific.' Others that were previously non-existent or trivial may, with a new paradigm become the very archetypes of significant scientific achievement. And as the problems change, so, often, does the standard that distinguishes a real scientific solution from a mere metaphysical speculation, word game, or mathematical play. The normal-scientific tradition that emerges from a scientific revolution is not only incompatible but often actually incommensurable with that which has gone before."[4]

Kuhn refers to some areas of science as being "pre-paradigmatic" that is, as awaiting their first general conception. It would seem as though a better conception of the state of Psychology -- and, indeed, of each segment of the knowledge base on which the NIE rests -- would be "multiparadigmatic." There is not one paradigm undergoing evolution

---

[4] Kuhn, T.S. The structure of scientific revolutions. Chicago: University of Chicago Press-Phoenix Edition, 1962, p. 102.

and eventual revolution. A number of paradigms coexist simultaneously with 'revolutionary' relationships to one another. They each hold different facts, knowledge, arguments, and methods as relevant. They each capture a different picture of the reality of education. They form, in more ordinary terms, a set of Schools.

Research and Development With a Multiparadigmatic Knowledge Base

Suppose we argue that at every relevant level of educational discussion -- the basic level, the level of problem and program studies, the level of analysis, and the everyday level of social and political debate -- there is multiparadigmatic knowledge. What does this imply?

1. It implies that one must deal with not one but a number of conceptions of the nature of education, the techniques and resources and methods to optimize education, the direction of R & D efforts, and the goals to be maximized by R & D work. People sometimes characterize the disagreements about education in public debate today as to be so extreme as to be chaotic. "Nobody agrees on anything." "Nothing is known." But this is probably too extreme. Multiparadigmatic knowledge is a valid form of knowledge...that is, it is basically factual knowledge and it does represent a reduction of uncertainty...but it is not knowledge rationalized into an overarching paradigm on which there is consensus. In a multiparadigmatic system one has "islands" of knowledge although one cannot map the space in which the knowledge exists. (For example, one has maturationist, interactionist and environmentalist conceptions of goals and methods in education coexisting, without having an understanding

of the boundaries and intersections among them.)

2. It implies that normal processes of scientific work -- research, evaluation, measurement -- will have limited impact on our understanding of educational processes although not necessarily negligible impact. "Normal science" develops within paradigms. One will be able to mount research and development efforts within the everyday conceptions of R & D, but the impact of those efforts will be tangible only within one or an adjacent family of relevant paradigms. One can compare the behavior modification efforts of two proponents of the environmentalist view that education is engineered behavior change. But the interactionist or maturationist who is a proponent of the view that education should heighten the self-concept or mental health of children will be uninterested in the comparison, and his program will not be commensurate in any "normal science" comparison with that of the behavior modifiers.

3. It implies that an important component of the NIE's efforts should be to act on the multiparadigmatic belief systems themselves. How? There is no worked-out answer. However, one suspects that the key function of the basic level of scientific work in education is to promote the evolution of the coexisting system of schools of thought about children's learning and development and, indirectly, about education itself. The nucleus of the knowledge base is a set of conceptions of the nature of human learning -- the learning theory conception, the information-processing conception, the psychoanalytic conception, the comparative-developmental conception, the ethological conception -- each part-fact, part-speculation, each a vision of the pattern of the phenomena of learning coming out of looks

through a different window, each true, each false. The schools come and they go. One imagines that as they come and go, succeeding schools are somewhat more powerful, somewhat more central, somewhat more subsumptive, than the ones preceding. One imagines reconciliations and a simplification of the multiparadigm system.

4. Let us imagine something like the following picture. Suppose there is a kind of nucleus to our conception of education. There are schools of thought about human learning...within each school, techniques of observation and research more or less rigorous, procedures by which two people can look at the same things and resolve their differences by more or less rough empiricism. Between schools, there are debates... usually bootless and sustained because there are not procedures for resolving these kinds of debates empirically. Each school is imperialistic. It extends beyond what it can see and prove, tries to explain 'everything' or, at least, many things. Coalescences occur, most often speculatively, at times genuinely. Berlyne speculatively tries to explain Piaget in S-R terms. Bowlby reformulates some basic child development notions of Freud in ethological terms and the reformulation seems interesting; it seems to fit; it seems worth working on. Basic research hunts its unifying paradigm and, at the very least, finds integrations among its coexisting paradigms.

Extending out from this nuclear system -- in the behavioral and the social sciences -- there are multiparadigmatic conceptions of educational processes and educational outcomes. The multiple paradigms of education are not coincidentally multiple nor coincidentally related to

the multiplicity of psychological theories. But from these multiple paradigms come a system of multiple conceptions of education, its goals and processes. One fundamental problem of R & D will be to foster the use of uniparadigmatic procedures to advance a multiparadigmatic knowledge base. The other will be to further the reconciliation of the paradigms.

5. With respect to problem and program studies, this conception implies that a multiplicity of alternative viewpoints must be considered in analyzing the educational needs of children and in developing and evaluating various procedures to meet these needs. It implies that a number of problems concerning children's social, economic and political roles are legitimate as concerns for American education as seen by some participants as are needs in strictly "educational" domains. Further, it implies that an important function of educational R & D should be the identification and analysis of a wide range of procedures and programs for addressing educational needs -- perhaps leading to something of a "Consumer's Report" of alternative educational practices, a descriptive report of the variegated "menu of alternatives" in American education.

6. With respect to policy analyses, this conception again implies the importance of considering a plurality of goals and views if superordinate analyses are to address adequately the diverse concerns of their audiences. The conception of basic research set forth in this essay has additional implications for policy analyses. Basic research transmits into the policy sector -- not technology, not methods, not simply these things -- but educated guesses about how to proceed. Policy decisions involve guesses, bets on courses of action; responsible administrators will seek to make guesses as educated as possible. This is precisely the reason for

undertaking careful analyses which pull together pertinent information and outline options for decision-makers and practitioners dealing with educational problems. In commissioning and using such studies, educational decision-makers will be wise to insist on hearing from a number of different schools of thought, so that they may be sure that the options from which they choose do not represent the choices of some select advocates but rather represent the pro's and con's of the range of alternative courses open to American education.

Finally, what if there is multiparadigmatic knowledge at the everyday level of social and political debate? What does this imply? Among other things, that learning to tolerate diverse opinions, to build on and thrive on pluralistic conceptions, to be open to the values and value systems of others is essential to the functioning of American education. Do we have still another principle and goal for education in this country?

Comments on White-Bissell-Golenski Paper by David K. Cohen

My comments consist of several brief questions raised by this paper.

First, one attribute of the structure of knowledge in education as you describe it might be phrased as a weakness of system memory--deficiencies in the process of meaningfully storing information, retrieving it, and applying it to practice or new inquiry. One reason that weak knowledge systems exists is pluralistic theory. And one function an NIE might serve is precisely to strengthen system memory, to reduce the isolated production and dead storage of knowledge, and to introduce a greater degree of cumulation into the knowledge system in education. You didn't mention this idea, and I wonder if it was because it seemed hopeless or simply unimportant.

Second, I got the distinct impression that you thought that theoretical pluralism was not a good idea on the whole, and that simplification and greater coherence was. You seem to assume that education will become increasingly scientific but I wonder why you think it is true, especially since most of the evidence you present in the paper would support the argument that education is permanently pre-scientific or multi-paradigmatic. A related question is how one would know, as an empirical matter, that paradigms were becoming fewer and more general. Because we stopped talking about them? Another question is whether multi-paradigmatic knowledge progresses or merely changes.

Finally, if the knowledge base is multi-paradigmatic, and if it is likely to be unruly for some time to come, how could one decide what good R & D is?

what work gets funded, these two roles collapse into one. Since both N.I.E. personnel and researchers tend to subscribe, either implicitly or explicitly, to a particular paradigm or set of paradigms, who gets funds could ultimately resolve the issue of competing paradigms without confronting openly and empirically the conflicts inherent in these various paradigms. The choice then becomes how to stimulate competition among paradigms with a minimum of duplication while at the same time insuring the most appropriate synthesis of knowledge and the development of higher order replacement paradigms. To help resolve this dilemma and to minimize the likelihood of the N.I.E.'s adopting a particular point of view, it was suggested that a panel of senior researchers no longer ego-involved in their own work or seeking to establish their own reputation be established to evaluate research proposals.

At a more general level, one might question the need for and the desirability of the N.I.E. assuming a coordinating role in the education R & D effort. More specifically, should the N.I.E. only facilitate a dialectic with regard to the synthesis of education R & D efforts and results or should it establish research agendas? Here the consensus appeared to be that the role of the N.I.E. would often be determined by the state of the art in a particular area. In a relatively new field, more freedom with regard to agenda setting and research strategies might best be left to the discretion of the researcher, subject to appropriate project oversight at N.I.E.. In more advanced areas, where considerable work had already been undertaken and the process of knowledge synthesis had already been initiated, the coordination activity then becomes essential.

Two other areas deemed important to the development of a knowledge base in education were training researchers and the establishment of incentives which would reinforce the production of quality R & D and the attainment of N.I.E.'s goals. Researchers in education, given the diverse nature

of the knowledge base in the field, need an interdisciplinary set of skills and insights. However, in present academic institutions, interdisciplinary training and interests tend frequently to be interpreted as non-disciplinary. What is needed, therefore, is the development and continued longterm support of interdisciplinary research institutes. These institutes would tend to mitigate the discipline-oriented promotion, recognition and peer systems prevalent in most universities, thus freeing the researchers to answer the complex problems raised by synthesizing previously acquired knowledge in education. However, to prove effective in the long run, these research institutes need to have a sense of permanence with regard to the rewards, recognitions and incentives that they provide for interdiscipline-oriented individuals. Training is important but without the proper rewards and reinforcements, the standards and orientations of the academic disciplines will continue to prevail.

General Discussion of White-Bissell-Golenski Paper and Cohen Comments

By and large, the discussion centered on the prevalence of different paradigms in education and the advisability and means of reconciling them. To focus the discussion, it was agreed that there were two basic ways of viewing the multi-paradigmatic world of education R & D. First there were different ways of defining the questions to be studied (and there was a lack of agreement with regard to the questions to be asked). And second, there were different notions regarding the goals of schooling. Education has never had a unified paradigm due to multiple perceptions across a range of issues. Rather, the field of education has experienced a series of co-existing intellectual frameworks.

If one accepts the preceeding description as appropriate and applicable to education, one is led to ask what the role of the N.I.E. should be seeking to mediate between these co-existing and frequently competitive paradigms. Additionally, how and at what level should this reconciliation be sought? One obvious role the N.I.E. could and should play is that of maintaining a system memory bank. In the past, there had been no synthesis of knowledge in education R & D in the United States. The cumulation and synthesis of knowledge are essential precursors to the development of higher order unifying paradigms. Thus, the N.I.E. could, in this fashion, seek to initiate and stimulate the simplification and coordination of R & D efforts in the education field, thus minimizing the duplication, non-cumulativeness, and faddishness so prevalent at present time.

The N.I.E.'s potentially conflicting roles as the source of financial support for education R & D and the arbiter of competing paradigms raises more subtle and more serious issues. Indeed, at a very basic point, namely

THE CONDUCT OF DEVELOPMENT IN EDUCATION

Richard E. Schutz

SWRL Educational Research and Development
Los Alamitos, California

I. Antecedents and Analogs

Development in education is usually treated as a new phenomenon with unique characteristics, problems, and potentials. It should not be so viewed, for there are losses in disassociating it from relevant antecedent and concurrent endeavors. Current and projected development phenomena lose nothing and gain a good deal by admitting to phenomena that came before and that coexist. Here and throughout the paper the term "development" is used with the standard referents of the National Science Foundation definition (1965): "the systematic use of scientific knowledge directed toward the production of useful materials, devices, systems, or methods, including design and development of prototype and processes." Several alternative definitions of development in education have been offered to justify a wider range of phenomena or to incorporate characteristics regarded as unique to the educational context. These departures from standard usage appear neither necessary semantically nor desirable instrumentally, and again reflect the tendency to isolate education from the human mainstream.

Antecedents of Development in Education

The antecedents of educational development can be traced to antiquity, but the history may also be accounted in short order. Up until the last century, the history of educational development is totally accounted for by the history of educational practice. Materials, devices, and so on were being produced on the basis of the scientific information then available. But development was conducted as part of the general educational enterprise.

During the twentieth century, research in education came to have established referents as an enterprise apart from educational practice. Utilizing the methodology of the behavioral sciences, largely psychology, a large quantity of education research was amassed. Whether this literature constitutes knowledge is moot, but the research reports did occasionally spawn development activity. Practitioners also continued to use available knowledge to produce useful materials and devices. Thus during the last century up until the mid 1960's development in education was conducted either within the boundaries of educational research or within the boundaries of educational practice.

Within the last decade increasing attention has been given to development in education as an endeavor warranting consideration apart from educational research and practice. But the endeavor has been seen as

naturally building upon and extending historical antecedents in educational research and practice. At the risk of heresy charges from colleagues in educational research and practice with whom I proudly identify, I am compelled to a radical alternative perspective.

Development in education now and in the future can best be conducted by breaking sharply with its historical antecedents. Although the historical seeds of educational development lie in educational research and/or practice, the seeds do not have within themselves the wherewithal to provide the nurturance that will now permit development in education to thrive above ground. Energizing sustenance can, however, be provided by experiences in science and technology in fields other than education, and it is here that we shall look in this paper. By breaking through its historical shell of educational research and practice, and driving its roots into broader scientific and technological knowledge, educational development at this point has its best chances of flowering and propagating in the future.

Analogs for Development in Education

Since the botanical metaphor has serious implications for the conduct of development in education, I shall substantiate it by contrasting the manner in which educational development is conducted if regarded as operating within the boundaries of educational research and practice and the manner in which it is conducted if regarded as operating within

boundaries of scientific knowledge and technological capacity generally. The contrast will be made in terms of dominating paradigms, disciplines, institutions, and timeframes.

Paradigm Perspective

The contextual paradigm for educational research and practice has been Mark Hopkins on one end of a log and a student on the other. All manners of characteristics of students and teachers have been identified, compared, and correlated in an attempt to enhance the human characteristics and optimize the human enterprise of education. In focusing on the inherently human aspects of Mark Hopkins and the student, the inescapable log on which they sit has been largely ignored. As a result, the developed useful materials, devices, systems, and methods that characterize other facets of modern life are not now present in education.

Despite impressive edifices and embellished interiors, the typical school is an impoverished instructional environment. Compared with even the typical home the classroom is barren. The call here is not for clutter nor for what educators term realia. It is for a perspective that is orientated toward providing developed items with demonstrated functional utility in instruction rather than expecting especially creative teachers and naturally ingenious students to make do with items that have for the most part been developed to fulfill other functions. The "teaching machine" quest was in retrospect just as simplistic as a "household machine" quest would have been had development for the home been

similarly viewed. But had the conduct of development for the home all been restricted to manipulations of the occupants thereof, the household would not be the sophisticated functional environment it is today.

The product goals of development in education are in no sense inconsistent or incompatible with the intrinsically human characteristics of education. People are not products; people use products. Educational products, like other products people use, provide alternatives that human beings may choose to use to extend their capability. By scrupulously maintaining the distinction between the poeple involved in educational practice and the products derivable through educational development, humanistic problems are not completely solved, but they are addressable within a human mainstream that includes education rather than isolates it as an enterprise demanding unique attention.

Discipline Perspective

Educational research and practice have traditionally looked to the behavioral and social sciences for their knowledge base, since these disciplines are viewed as addressing the human characteristics of education. While development in education can certainly draw upon the knowledge base of these disciplines, it is unnecessary and undesirable to restrict its consideration to these disciplines. Fields including but not limited to aerospace, architecture, business, engineering, and

pharmacy have associated knowledge that is currently relevant to development in education and that can be expected to advance the state-of-the-art in the future.

The relevance of the fields just enumerated is by transfer of their structural and management strategies at a macro-level rather than transfer of their personnel or applications at a micro-level. These technically oriented fields have been less self-conscious concerning their methodologies than have the behavioral sciences, and thus have given only modest attention to packing their macro-strategies in a form readily cognizant to persons within the field or transferrable to other fields such as education. Thus the typical specialist within these fields is not now in good position to contribute to the conduct of development in education. He has adopted the view of education as a field apart and when confronting education either has no response or emulates his notion of how a teacher should respond. In either case he appears foolish.

It is the methodology not the men of aerospace, architecture, business, engineering, and pharmacy that renders these fields ripe for contributing to development in education at the present time. The prevailing methodology of the behavioral sciences seeks to isolate differences and to evaluate comparatively. The predispositions and techniques that follow from behavioral science methodology tend to be disfunctional in

development. A counter methodology seeks to identify commonalities and to evaluate cybernetically. The predispositions and techniques that follow from this alternative methodology tend to be highly functional in development and are shared in common by intellectual endeavors in the fields enumerated above. But the typical behavioral scientist must suppress strong competing orientations in order to acquire the "new" predispositions and techniques. His quest for comparative differences in phenomena blinds him to the syncretic characteristics of the same phenomena.

Institutional Perspective

Educational research and practice have looked to the higher education sector for supportive sustenance. The higher education sector has controlled the personnel and structure of schools through certification and accreditation complexes. It has also controlled the substance of schooling by providing the rhetoric, paradigms, texts, and research that determine this substance. One need not take a position on the quality with which these controls have been exercised to concede that there are disadvantages in lodging development in education within higher education.

Experience in development in other fields suggests that the industrial sector rather than the higher education sector is the best fulcrum for development in education. The reasons for this pertain to fundamental characteristics of the university and of the development enterprise that are not readily compatible. The university is the

undisputed home of science. Charles Eliot's turn-of-the-century definition of the university is still accurate: "a voluntary cooperative association of highly individualistic persons for teaching and advancing knowledge" (Wolfle, 1972, p. 94). The institution of higher education is admirably equipped to forward research, but the individual autonomy, conceptual elegance, and fragmented specialization characteristic of higher education are wasted in driving the development engine. This is in no way to suggest that educational development in a university context be proscribed or that research related to development in education should be relegated to the university. It is to suggest that justification for educational development activity in higher education be research-based and that justification for research activity in industry be educational development-based.

American industry has no current development capability in education; there has been no basis of support for this capability. The education market has been relegated to "publishers," a small and weak industry. American industry in general has been twice burned in the last decade in new venture areas in education other than publishing. Efforts to penetrate educational practice via teaching machines and via performance contracts did not prove profitable economically or effective professionally. Remnants of these two abortive movements remain alive today, but the scar tissue left in their wake is also present.

The industrial sector has no unique potential for conducting research in education or for operating schools. It has great unique potential

for conducting development in education. This potential includes matters of management capability and socio-technical strength that make it eminently reasonable and feasible to lodge educational development in the future in the industrial sector of the nation.

Temporal Perspective

Educational research and practice have been motivated either by blind faith and infinite patience or by alleged crises and instant "solutions." These patterns have been reflected in educational legislation and in public information pertinent to education, but they are irrelevant to development in education. While it is possible to motivate educational development by blind faith and infinite patience, it is not necessary to do so. The tangible referents for development outcomes and the finite estimates of development completion dates permit greater specificity of motivation than blind faith. Likewise, it is possible to cast development in a solution-to-crisis mold, but it is not necessary to invent a crisis to motivate development. Development can, indeed, make a contribution to the solution of real crises. But it can also contribute outcomes motivated by the anticipation of positive potential rather than by the avoidance of crises created for that purpose.

Apart from motivational rationale, both the "instant-infinite" and the "one-year" time frames in which educational research and practice have been fielded are altogether unrealistic for development in education.

In fields other than education a ten-year time frame is treated as the minimum possible time for fielding a development effort from commitment to completion, and a quarter of a century is not uncommon. This may seem like an inordinately long time, but it can be confirmed by adding up the years (Bright, 1969). Assume that a set of concepts can be demonstrated in application form so that a development effort may reasonably be initiated. "How long will it take to achieve a prototype for full scale or field trial? One to four years? Assume two years. Then how much longer until a commercially saleable product with necessary adjuncts in the form of maintenance, user training aids, promotional support, etc. is ready for sale? One to four years? Assume two years. Once first sale is made, how long will it be...until the innovation is in widespread use? Three to 10 years? Assume six years. Using these rough assumptions on the optimistic basis the total time is about 10 years! Now allow for the fact that we may be lucky in shortening some of these phases, but are more likely to have under-estimated at least one of them. Then a 15-year time span is a strong probability" (p.38). If you doubt the applicability of these parametric figures for education, check them against the current planning frames of educational publishers. You will find that their offerings for the late 1970's and early 1980's are now determined. Notice that the addition above started at the time that the scope of the development could be well-enough specified so that development effort could be started; the prior time for inquiry to achieve this level of specificity was not included and would still further increase the time span were it to be accounted for.

It has been a popular pasttime in both educational research and practice to despair over the time interval between the invention of a concept and its reflection in prevailing practice. But this despair has produced only hand waving or wringing. What is not recognized is that it is not the gap that is unique to education, but the effort that achieves successive waves to change the definition of the gap. General convention (Kuhn, 1962) recognizes these successive waves that change the nature of the gap as scientific and technological "progress." It is this absence of change in the nature of the gap between thought and practice that has uniquely characterized education, not the gap per se.

Coda

If development in education is viewed in terms of one person (looking like hundreds of university professors and their graduate students or like millions of individual teachers working in isolation), and in terms of thousands of dollars for a few years (looking like a research project or like an avocational pursuit), it can be dismissed as inconsequential. There is no way that the conduct of development in education can immediately match the sophistication of development in other areas where the development endeavor is currently established and valued. However, it is in these areas rather than in educational research or practice that relevant analogs are to be found.

The current state-of-the-art of development in education justifies a view of its conduct in terms of managed personnel units (looking like

industrial R&D organizations) and in terms of several million dollars for several years (looking like small resource commitments to development in areas other than education). Elaboration and defense of this contention will motivate the remainder of the paper. An elegant analysis of the precedents for the radical restructuring of the field that is being called for here has been presented by Platt (1970).

II. Structures and Systems

Radical restructuring of the conduct of development does not demand or imply either unfettered expansion of activity or dilatory appraisal of priorities. It does require a paradigm that provides a comprehensive structure for the endeavor. Paradigms growing out of educational research or practice that use categories of academic disciplines, human attributes, demographic divisions, and so on, do not fit the development endeavor and lead to gross distortion and misinterpretation when applied to the development context. Neither are macroparadigms that use categories of research diffusion, dissemination, evaluation, and so on, any more useful. Irrespective of their merits or deficiencies for other purposes, these catgories beg the issue of structure for the conduct of development. Finally, in eliminating structural contenders, we can discard miniparadigms of the development method that use catgories of design, engineer, test, iterate, and so on. Like paradigms of the scientific method, these miniparadigms prove useful if treated as background boilerplate, but are dangerous if believed as trustworthy templates for action.

What does this leave? Fortunately, the shelves are not bare. The goods have been delivered by structures growing out of the conduct of development in societal enterprises other than education. The best documented experience derives from defense and space development. The public information aspects of defense and space have subordinated the interpersonal and management aspects of these enterprises to hardware that can easily be

photographed and simply depicted in a form amenable to the popular media. However, after cutting below this press image, it becomes clear that all of the people problems that are found worrysome in development in education are also present in space and defense development. They have simply received less emphasis. Missiles are not men, but management structures are management structures and people are people in development wherever it is conducted. This undersimplification is offered not to support but to introduce the paradigms to be presented. The paradigms are offered as illustrative alternatives rather than as definitive imperatives. As further experience in development in education is gained, a more suitable paradigm will no doubt be produced. But future action need not be a bootstrap effort. Development in education may proceed by standing on the structures of giants who have come before.

DOD-Borrowed Paradigm

Table 1 is adapted from a survey of DOD categories presented by Glennan (1967). Glennan's categorizations are paraphased and freely adapted to reflect and incorporate distinctive characteristics of education. A brief description of each category is followed by suggested prerequisite criteria for effort initiation that define the boundaries of each category. This paradigm has several things going for it. First, it permits concurrent coordinate activity that successively reduces the uncertainty inherent in development. Second, it hedges the risks encountered in contracting for the total procurement at one time, promotes reasonable competition within each category, and guides the anticipatory

Table 1

CATEGORIES FOR AN R&D MANAGEMENT STRUCTURE IN EDUCATION

I. Research includes all efforts directed toward increased knowledge of natural pheonomena and environment and toward solutions to problems in the physical, behavioral, and social sciences. By definition, "research" includes all basic research in addition to applied research directed toward expanding knowledge in various scientific areas. It does not include time-oriented investigations and developments.

Effort initiation criteria:

1. The utility of the potential outcomes of the research is high.
2. The scientific or technological domain is judged to be ripe for exploration.
3. Talented scholars and scientists are available or recruitable.

II. Exploratory development includes all efforts to resolve specific problems short of major development projects. These efforts may vary from fundamental applied research to sophisticated experimental prototypes study, programming, and planning efforts. The dominant characteristic of this category of effort is that it is pointed toward specific problem areas, with a view toward developing and evaluating the feasibility and practicability of proposed solutions and determining their parameters.

Effort initiation criteria:

1. The technical feasibility of a promising model is uncertain and warrants further investigation, or
2. A requirement for a prototype or component can be specified with sufficient precision to permit further effort to refine the specifications, or
3. Experimentation is required to investigate the parameters or performance limit of a prototype or component of a subsystem, or
4. The effort involves the testing of a model preparatory to the development of a prototype or component of a subsystem and the technology for such effort is available.

III. Advanced development includes all efforts that have progressed to the development of systems for experimental or operational tests. Advanced development is characterized by line-item projects, normally involving systems designed for test or experimentation as opposed to those designed and constructed for operational educational use. The major distinction is in terms of readiness for use.

Table 1 - Categories for an R&D Management Structure in Education - Continued

Effort initiation criteria:

1. A promising exploitable technology is available and the priority or magnitude of the effort is too great to warrant consideration as exploratory development, or the nature of the effort is such that more extensive management is required to insure continuity or cost control than is reasonable under an exploratory development effort.
2. Primarily development rather than experimental effort is required, and the technology needed is sufficiently in hand.
3. The system and performance objectives have been defined.
4. The best technical approaches have been selected.
5. A trade-off analysis of alternative system configurations has been made.
6. The cost effectiveness of the proposed item has been determined to be favorable in relationship to the cost effectiveness of extant items.
7. Cost and schedule estimates are credible and acceptable.

IV. Operating program development includes efforts directed toward the full development, engineering, and testing of all of the essential systems, support programs, vehicles, materials, and procedures that have been demonstrated ready for installation and operational use.

Effort initiation criteria:

1. Primarily system articulation rather than system development effort is required, and the technology needed is sufficiently in hand.
2. The operating environment and performance envelopes are defined.
3. The best technical approaches have been selected.
4. A thorough trade-off analysis of alternative program configurations has been made.
5. The cost effectiveness of the proposed program has been determined to be favorable in relationship to the cost effectiveness of competing potential programs.
6. Cost and schedule estimates are credible and acceptable.

V. Installation/operation. The category subsumes operating cost evaluation, production-marketing, installation, and operation. Operation is relevant to R&D only to the extent that it reflects such post-installation activity as the setting of standards over time.

Table 1 - Categories for an R&D Management Structure in Education - Continued

Effort initiation criteria:

1. All systems involved in a new operating program are available or a firm availability date can be projected.
2. The cost effectiveness of installing the new program has been determined to be favorable when compared with that of current operating programs.

expectations of all parties involved. Third, it permits a healthy range of contractor specializations with sufficient differentiation and stability, and with reasonable redundancy to span the full development time frame. Other advantages and implications could be cited, but I shall resist further embellishment, since the purpose is only to set forth the paradigm.

NASA-Borrowed Paradigm

Table 2 is adapted from an impressive analysis of management style and organization stucture presented by Sayles and Chandler (1971) that relies heavily on NASA experience.

The information in Table 2 is given only incidental treatment by Sayles and Chandler and is not the basis for my praise of their book. I offer the paradigm to show an alternative formulation that happens to be compatible with the structure in Table 1 and that presents additional facets of the endeavor, such as the appropriate differentiation of agency-contractor responsibility and type of contract. If the paradigm were to serve no function other than to provide a contrast with the erstwhile "procurement practices" of USOE to write RFPs for R&D based on the pet notions of agency staff at expenditure levels of whatever they could scrounge out of annual appropriations residuals, it would serve a useful purpose. But the Table implies more than this. In elegant simplicity it quietly structures the resolution of complex socio-technical issues that in education have generated nothing more than heat in skirmishes surrounding "free competition," "federal control,"

Table 2

PHASED DECISION POINTS FOR THE CONDUCT OF DEVELOPMENT IN EDUCATION

Phase	Objectives	Agency	Contractor
A. Preliminary Analysis	Analysis of alternate overall approaches and concepts	Primarily an in-house effort	Support role for study contractors (FP or CPFF contract); need not be capable of Phase B, C, or D.
B. Definition	Selection of one of several approaches for further definition and eventual development of this seems advisable; effort may be cut off here	An analysis role	Study contractors develop information (FP or CPFF contract); not a competition for Phase D contract.
C. Design	Definition of detail of the approach selected in Phase B	Integration and validation of contractor data	Major portions of work are contractor conducted (CPFF or incentive contract); generally 2 or more prime contractors selected; only firms capable of performing through Phase D are eligible since Phase C provides competition for Phase D.
D. Development	Final design, development and testing	Monitoring and review functions	Major portion of work is contractor conducted; restricted to Phase C contractors except in unusual cases, one prime contractor (incentive contract).

"autonomy and independence," "public and private," and so on. Again, I shall resist further embellishment of these desiderata.

Both of the foregoing structures are characterized by (1) incremental acquisition, based on a sequence of decision points and a succession of development phases and (2) pronounced austerity in the early phases of development (Perry, 1972). It may be professionally impolitic to advocate austerity of any sort in this present season of financial adversity for development in education. I have no aspiration to outslash the budget slashers. The austerity being advocated is structural, not financial. It is offered as an alternative for the present anarchy that imposes austerity controls late rather than early in development. This anarchy is exemplied in the erstwhile USOE practice of laissez faire development followed by a "county fair" competition among "product" entries vying for NCEC-sponsored "dissemination" prizes, with the judges supplied by the higher education and school communities and the fairground operated by ETS under contract to NCEC. The "county fair" strategem would be inconsistent with the largest national resources imaginable for development in education. The decreasing austerity strategem is operable with the national resources presently available for development in education.

"Austere initial development is an important element in any incremental acquisition strategy....During development, the desired product is information, and only information. Hardware is merely a means of acquiring

the information needed to proceed to another phase....Irrelevant information is inevitably expensive and frequently worthless. It is unlikely, for instance, that...consumption rates and maintenance requirements can be accurately calculated before test articles are in hand and test experience has been accumulated. Making such calculations is costly. Acting on them before they can be validated is very costly" (Perry, 1972, p. 358).

Industry-borrowed Paradigm

The notions of purposive phasing in R&D are not restricted to public R&D enterprises. Table 3 shows classifications and definitions recommended by the Committee on Research Definitions of the Industrial Research Institute (Brown, 1972). The Committee "did not find it helpful to classify R&D by how it is done (fundamental, basic, applied), or by where it is done (central labs, divisional outposts, semi-works, on the bench) or by whether the research is product or process oriented. They found it most helpful to classify the research on the basis of why it was done" (p. 56). By substituting educational for business referents, the tabled categories appear readily generalizable to development in education.

System Modification Alternatives

The paradigm in Table 4 dimensionalizes alternative routes for modifying an educational system. Several implications may be drawn from the array.

First, the array indicates that it is possible to structure the management

Table 3

OBJECTIVES-BASED R&D CATEGORIES

Support of Existing Business Research - This research is that conducted in direct support of the given company's existing business to maintain or improve its profitability, and to improve its social acceptance. It is conducted to retain or increase market share by introducing new products, by improving the quality of existing products, by decreasing the cost of manufacture, or by preventing excessive increases in cost of manufacture, by extending the market of existing products into new applications, by enhancing safety, reducing pollution, or in other ways improving product or market acceptance

Exploratory Research - Exploratory research is that research performed for the purpose of advancing knowledge of phenomena of general company interest and also for finding major new high risk business projects. It is usually long range in nature but may include literature searches, laboratory scouting experiments, preliminary application and engineering studies, and preliminary economic evaluation. A new product, process, or service is in view, but the work, by definition, remains "exploratory research" until a product or process objective is established.

New High Risk Business Project Research - New high risk business project research is that conducted with the intention of developing a product, process, or market in which the sponsoring company has no direct manufacturing or market experience, or both. It includes those projects which involve a diversification or a totally new way of accomplishing an important function. It is high risk in nature. This research may result from the successful accomplishment of exploratory research or may be a new program related to otherwise acquired technology. It can include all the technical categories of work associated with research and development.

Table 4

ALTERNATIVE ROUTES FOR MODIFYING AN EDUCATIONAL SYSTEM

Route	Examples	Prerequisite	Vehicle	Effect Dependent Upon
Administrative Organization	Experimental schools Storefront schools Day care centers	Defined workable structure	Institutional organization	Institutional leadership
Authoritative Analysis	Kerr Commission reports Levien report California Master Plan	Analysts and commission	Report and publicity	Public acceptance and executive interpretatior
Development	Sesame Street SWRL Kindergarten Program PLATO	Development technology	Devloped system	Quality of development
Evaluation	EEO survey National Assessment State Testing Programs	Evaluation technology	Report of findings	Evidence and recommendations
Judicial Ruling	Supreme Court ruling Circuit Court ruling State Court ruling	Statute and claim	Court order	Legislative and executive reaction
Legislation	Federal law State law Local law	Legislative majority	Statutes	Executive enforcement
Media	Cable TV Computer systems Audio cassettes	Medium and message	Hardware-software	System quality
Personal Service	Education extension agent Management consultant Psychotherapist	Knowledge-embedded-in-individual and receptive client	Interpersonal relations	Individual expertise

Route	Examples	Prerequisite	Vehicle	Effect Dependent Upon
Technique	Micro-teaching Busing Finance formulas	Reliable defined procedures	Media for transmission	Technique and medium
Training	Teacher education Graduate fellowships AERA presessions	Training system	Instruction	Instructor and materials
Trend	Career education Preschool education Open schooling	Multiple convergent determinants	Disjointed incrementalism	Continuity of determinants

of educational change independent of references to categories of students, teachers, disciplines, and methods. It is quite feasible to use multiple descriptors to characterize any modification effort, but excluding descriptors appropriate to a given effort in deference to descriptors appropriate for the management of another enterprise is indefensible.

Second, the array encourages an open and pluralistic approach to educational change. It recognizes development as one of several feasible routes. Each route has unique strengths and potentials, and none is without its constraints and weaknesses. It is inappropriate to subsume all the routes under the rubric "development," as USOE has done in the past, since differential treatment is in order for each. The array recognizes that there are several ways to skin a cat, and that each way has its distinctive features. Grabbing the sharp end of the knife with both hands creates difficulty because it fails to respect the distinctive features of that route. I apologize for the crude metaphor, but it aptly expresses the clumsy efforts at educational system modification that dot the past.

Third, the alphabetical ordering of the alternatives indicates that they are complementary rather than competitive, coordinate rather than hierarchical. A popular gamesmenship ploy in recent years has been to disparage all routes except the one being promoted as "minor tinkering" in contrast to the great "cost benefit" promised by the route being advocated. As a matter of fact, it is unnecessary to select one route as "best" since

it is possible to pursue concurrent routes at no increase in cost to education. Each route has unique but complementary strengths and potentials, and none is wothout its constraints and limitations.

Fourth, the array provides a basis for parametric estimates of the resources required to pursue efforts along each of the routes. Present experience permits the assignment of time and cost estimates to efforts within several of these categories. Table 5 addresses this point. I do not vouch for the precise accuracy of these estimates since my own estimates differ, give or take a few million and a few years. It is my understanding that more refined study of parametric effort boundaries were prepared in connection with NIE planning, but so far as I know this study has not been made public. My point is that the state-of-the-art does now permit such parametric estimates and that development in education in the future need not be planned and fielded in ignorance of such estimates. Parametric estimates may also be applied to the costs of creating the institutional capability required to provide the prerequisites for pursuing each route.

Finally, the array suggests differential regulatory criteria appropriate to guide efforts along the various routes. The establishment of regulatory criteria is a traditional arena for healthy focused cooperation between a governmental agency and allied professional associations. Had NCERD

Table 5*

ESTIMATED[1] COST FOR
EDUCATIONAL INNOVATIONS

Type of Innovation	Cost	Delivery Time
Major Curriculum Projects (such as BSCS, new math, etc.)	10-15M	5 - 7 yrs.
New Innovations in Media and Technology (such as Sesame Street, Computer Assisted Instruction)	10-15M	3 - 5 yrs.
New Efforts at Assessment - Accountability (National Assessment Program, Belmont Project)	15-20M	8 - 10 yrs.
An Experimental School	5M	5 yrs.
Major Studies in Financial Reform	5M	3 - 5 yrs.
Training One Hundred Senior Researchers	6M	4 - 5 yrs.

[1]Estimates provided from National Center for Educational Research and Development, Office of Education.

*Excerpt from testimony before House Select Subcommittee on Education by James J. Gallagher, February 18, 1971.

officials and AERA members, for example, conjointly devoted half the attention to defining and legitimizing the criteria for development outcomes that they directed toward evaluating and legitimizing the evaluation of development efforts in the total absence of such criteria, both the state-of-the-art of development in education and the public would have been better served. In their zeal to "win one for old OE," NCERD consistently vacillated between a "We'll solve all your problems," and "Don't bring any of your problems to us," position with its development contractors. This orientation totally abrograted the regulatory mechanisms that have traditionally proved successful in other areas of federal government. A federal agency runs a high risk in meddling directly in contractor affairs. It smooths out problems by establishing firm regulatory boundaries within which contractors may operate in the interests of all parties concerned.

III. Management And Manpower

Management Considerations

"Management" is a nasty word both in many quarters of academia where its connotations of regimentation are abhorred and in many quarters of government where its connotations of control are feared. However, management is inherent to the conduct of development, and its pejorative connotations can be avoided by explicit action toward this end. Such action involves addressing and accommodating dilemma that inherently affect all persons involved in a development effort. These relate to such matters as:

Precision - Ambiguity

Independence - Dependence

Competition - Cooperation

Refinement - Completion

Effecting - Marginalizing

Classically, the terms on the left have been considered to be the preferred choices. However, the terms on the right in reality come closer to a reasonable resolution. There are no magic rules for handling these management matters, any more than there are for handling other matters of development.

Literature on the above topics is sparse. Derek Price (1970) has pointed out that technology in general tends to be papyrophobic in contrast to the papyrocentric concerns of science. These tendencies appear to result from determinants of personal property rather than intellect.

"If you want to make capital out of technological discovery (whether the discovery be individually or institutionally referenced), the last thing you want is that open publication that determines... private property for the sciences" (p.8). It is possible, however, by borrowing heavily from Sayles and Chandler (1971) to at least sketch the boundaries of the apparent paradoxes. (Page number citations without name references in the text below are to their book.) Resolving the paradoxes is one of the intellectually exciting aspects of development remaining for the future. The pursuit of development in education should deliberately contribute to this resolution.

Precision-Ambiguity

"There is a sharp contrast between the precision of specification and recordkeeping in high-technology projects and the managerial process associated with their effective pursuit. The latter is characterized by a highly fluid, iterative, and seemingly imprecise series of activities that require a high degree of personal interaction (p.225)." The classical ideal of management that has for a decade been recognized as naive (Braybrooke and Lindblom, 1963) includes a set of milestones converging on the attainment of a precise solution to a problem that was specified with pristine clarity prior to beginning the effort. Never 'twas so, and never 'twill be in development in education or in any other field.

"In traditional management theory administrators are expected to collect and weigh facts and probabilities, make an optional decision and see that it is carried out. In large-scale development projects, a clear sequence of action is not possible because of their extended duration, the many technical unknowns, the changing balance of power among interest groups, the continual discovery of new 'facts,' and constantly changing constraints and pressures....It is assumed that the problems are simply and directly solved by rational analysis when, in fact, a great deal of interplay and negotiation may be necessary" (pp. 7-8).

A development effort that is conducted as a mechanical completion of milestones will either trivially advance the state-of-the art or intellectually misrepresent the complexity of its operations.

"Modern development programs have life histories filled with unanticipated crises, unpredicted barriers and impediments. What appear to be reasonable designs, given prior knowledge and experience, turn out to have neglected some small, crucial factor, and some subpart... fails to work. This in turn means that the subsystem may have to be redesigned to 'work around' the problem, which in turn affects other subsystems and the larger system.... These complex technical endeavors...require not less but more human ingenuity, improvisation, and negotiation than old-style business and government organizations" (pp. 10 and 16).

Independence-Dependence

"A major paradox...is that effectiveness in development programs requires a high order of responsible autonomy and the opportunity to innovate and even to change plans. But large scale projects...also require unbelievably precise integration and coordination among the parts...Thus a wide array of intellectual and economic commitments must be simultaneously focused on a very explicit task without destroying the motivations that release energy and commitment" (pp. 5-6).

The American society historically and now has placed a high value on independence. This tradition demands that both individuals and institutions be officially recognized as "independent." However, a mission orientation introduces constraints on all parties contributing to the mission.

"A mission orientation...clearly is not consistent with a literal interpretation of the 'independent contractor' concept. [All parties] must be able to act in concert to be immediately responsive to a program's needs. A certain degree of separation from external pressures that might prematurely abort potentially significant advances is also required. Thus, the development group needs a working arrangement that will insulate it from its environment, and a monopoly or near-monopoly of certain relationships is one way of achieving this goal. To get on with the job, the sponsoring agency is almost forced to make itself the central figure in a closely knit group of organizations, insulated

from external pressures--from the environment--and therefore dependent upon the sponsor. To secure this relationship, the sponsor is obliged to provide unusual guarantees as protection against risk, such as compensation for losses that may be incurred [with] approaches that show promise but eventually prove unfeasible.... Management of this style of relationship is of special significance because it is by no means limited to advanced technologies. In the future we can expect a much greater use of mission-oriented aggregations composed of a sponsoring agency and a diverse group of satellites who have banded together to achieve a major social or economic goal" (p.71).

Such interdependence includes profit making as well as public organizations and individuals as well as institutions. Interdependence has never been strongly pursued in development in education since it has been viewed as a reflection of weakness rather than strength. An opposite view appears to best forward both a mission orientation and a development enterprise. A poignant anecdotal illustration of this point is NIE's curt referral of its contractors to the Small Business Administration for loan assistance in contrast to the elaborate agency efforts that led to arranging federal guarantee for the $250 million Lockheed loan.

Matters of independency-dependency are often viewed as unilaterally involving individuals or institutions other than the sponsor, but the effects are reciprocal. A sponsor that disdains all interdependency relationships is itself totally dependent upon others in forwarding its interests externally. This dependency typified NCERD and its predecessors in USOE.

Competition-Cooperation

This dilemma is closely allied to that of independence-dependence. Both competition and cooperation are each useful mechanisms for enhancing excellence. The accommodation of the incompatibilities of the two mechanisms appears to lie in a self-forcing, self-enforcing system. "To achieve this goal, a pressure system must be devised that will function to correct significant errors and prevent major distortions from arising. Relying heavily on indirect means, management provides pressure in the right direction so that most of the time the system will be brought back to its original course. Management of large-scale endeavors essentially involves the skillful creation of such a pressure system" (p.104).

The system envisioned is still an aspiration rather than a reality in any field of development. Techniques toward this end in educational development that have proved useful in practice are described in Schutz (1972).

Refinement-Completion

Research work, like woman's work, is never done. Development work must be treated as complete at the earliest setting sun, although it is clear that it could be extended and refined to good effect for a much longer period. "Letting go" of a piece of work is one of the most difficult things for a novice in development to learn how to do gracefully. The tricks of the trade known to me are described elsewhere (Schutz 1970 a and b), but there is no "single best" resolution.

"The complex intermeshing of scientific needs, engineering requirements, budgetary limitations, organizational constraints, and personal goals and values almost ensures that project decisions will involve a complex of trade-offs among many different gains and a variety of losses. Experienced and knowlegeable participants cannot eliminate the need for trade-offs, but they can approach the bargaining with a realistic evaluation of the possible outcomes" (p.64).

Effecting-Marginalizing

Even economists tend to prefer direct effects over marginal accomplishments when there is a choice between the two modes (Charlesworth, 1972). In development, marginalizing is often more efficient as well as more effective than direct manipulating.

Development efforts in education are simply too complex to be handled by one individual in a hands-on, do-it-yourself fashion. The development specialist "acts the role of a marginalist. He widens or narrows limits, adds or subtracts weights where trade-offs are to be made, speeds up or slows down actions, increases the emphasis on some activities and decreases the emphasis on others" (p.209). He finds that "there is often not a precise, rational solution to most questions; rather the answer is a product of flexible give-and-take" (p.215). He "strives constantly to keep an appropriate balance in relative effort for what are always somewhat conflicting objectives and to avoid the usual degredations by which high hopes are dashed on the rocks of 'realistic

solutions'.... The process thus becomes a kind of continuous test of the perspicacity, alertness, and omniscience of those involved in the project. As such it provides very useful feedback to the manager, perhaps much more useful than the data provided by traditional appraisal mechanisms" (pp. 216-217).

Manpower Considerations

Considerations of manpower have been postponed to this point not because human resources are incidental to development in education, nor because qualified persons are available in good supply. Neither could be much further from the case (Levien, 1971). However, unlike the conduct of research, the integrated group rather than the isolated individual is the reasonable unit for considering the conduct of development. This does not reduce the importance of the individual in any development enterprise. It simply requires greater attention to insuring an environment that will make it possible for each individual involved in the conduct of development to be professionally productive and personally satisfied.

It is thus inappropriate to impute the attributes of a researcher to create a "developer" role. Some person's have assigned the term "educational developer" to themselves or their students. However, a one-man "developer" will be superficially trained and will operate superficially. The conduct of development in education requires highly competent specialists, not prima donna generalists. Now, and likely forever, personnel qualified as journeymen contributors to development in education are likely to be trained and to identify themselves as discipline specialists rather than as "developers."

The temptation to anthropomorphize abstractions such as "development" and "evaluation" into "developer" and "evaluator" has been, however,

compelling popular. University training programs for "educational developers" have been established, and training materials for such programs have been solicited and contracted for by USOE. These programs and materials can do little at present than to communicate irrelevant dogma and obsolete technology regarding development in education.

The technical sophistication of development in education is still so primitive, but is advancing so rapidly, that it appears both premature and inadvisable to attempt to pack it into degree programs. Books (e.g. Baker and Schutz, 1971, 1972) are feasible, and courses are, perhaps, reasonable, but any additional academic trappings are empty pretense.

The methodological and substantive competence of graduates of prevailing bachelors, masters, and doctoral level programs is quite adequate for development efforts. The deficiencies in the training received by these individuals relevant to development in education are not in methodology or substance. The deficiencies pertain to personal and discipline attitudes inculcated by academic training that forward research contributions, but impede development contributions.

Fry (1972) has conveniently summarized the attitudinal dispositions that must be adjusted in moving from the university laboratory to the development laboratory.

1. "The technical sophistication of a concept is no guarantee of its commercial success." In the university laboratory, the cleverer the idea, the more attention it is likely to receive. In the development laboratory, the goal is the quickest, most direct, most simple, least expensive means of reliably accomplishing desired functions; technical sophistication is at best a means toward this end. Conceptual complexity usually leads away from the goal.

2. "The work of a development laboratory is creative and synthetic." It is the creative synthesizer rather than the critical kibbitzer who forwards the work of a development laboratory. "Anyone with reasonable intelligence can do a good job of choosing between alternatives for objectives. The valuable man is the one who defines the need for a new activity, or who realizes that a certain characteristic of a product, which was brought to the present level only with apparently great difficulty, is in fact rather low in terms of what is ultimately possible."

3. "Work in the development laboratory is frequently empirical manipulation of highly complex and poorly understood systems." The plea, "We don't yet know enough," begs the development question, as do simplified model systems abstracted from the real or natural system of interest. In the development laboratory, personnel are perforce dealing with a whole process in its full complexity, whose mechanism is not fully understood and whose variables are incompletely defined. Moreover, they are asked to affect changes in a relatively short time.

4. "Work in the development laboratory is largely a group activity." This point has been elaborated earlier. "In the development laboratory, the final product or process is rarely associated with one individual."

5. "Development projects take much longer to complete than research projects." Again the point has been made above. "Commercialization of an idea in the development laboratory may take up to ten years."

6. "Because of the effects of process changes on efficiency, production functions may be reluctant to adopt such changes." It is modest comfort to persons in education to recognize that production managers, like school administrators, have concerns beyond technical soundness and financial benefits. "To the production operation in the short run, change, in the short run, means lower efficiency and higher unit costs...Manufacturing will never be as willing to adopt process changes as the [researcher] might anticipate."

For the forseeable future, the development-desirable predispositions enumerated above can best be produced in quick on-the-job orientation or in an internship in a good development laboratory in conjunction with an academic training program. Such internships would also be useful for established researchers. However, for an established researcher to pass through the doors of a development laboratory is as difficult as for an established camel to pass through the eye of a needle. The reason has nothing to do with matters of heaven, but it has a lot to do with matters of earth.

REFERENCES

Baker, R. L. and Schutz, R. E. Instructional product development. New York: Van Nostrand Reinhold, 1971.

Baker, R. L. and Schutz, R. E. Instructional product research. New York: Van Nostrand Reinhold, 1972.

Braybrooke, D. and Lindblom, C. E. A strategy for decision. New York: Free Press, 1963.

Bright, J. R. Some management lessons from innovation research. Long Range Planning, 1969, 2, 36-41.

Brown, A. E. New definitions for industrial R&D. Research Management, 1972, 15, 55-57.

Charlesworth, J. C. Integration of the social sciences through policy analysis. American Academy of Political and Social Science, Philadelphia, Monograph 14, 1972.

Price, D. J. de S. Citation measures of hard science, soft science, technology, and nonscience. In C. E. Nelson and D. K. Pollock (Eds.) Communication among scientists and engineers. Lexington, Mass.: Heath, 1970. Pp. 3-22:

Fry, F. H. Six differences between long-range and development labs. Research Management, 1972, 15, 64-71.

Glennan, T. K. Jr. Research and development. In J. Enke (Ed.) Defense management. Englewood Cliffs: Prentice Hall, 1967. Chap. 20.

Kuhn, T. The structure of scientific revolutions. Chicago: University of Chicago Press, 1962.

Levien, R. E. National Institute of Education: Preliminary plan for the proposed Institute. RAND Corporation, Santa Monica, CA, 1971.

National Science Foundation. Federal funds for research, development, and other scientific activities. NSF 65-19, 1965. Pp. 56-57.

Orlans, H. The nonprofit research institute. New York: McGraw-Hill, 1972.

Perry, R. European and U.S. aircraft development strategies. Policy Sciences, 1972, 3, 349-359.

Platt, J. Hierarchical restructuring. General Systems, 1970, 15, 49-53.

Sayles, L. R. and Chandler, M. K. Managing large systems. New York: Harper & Row, 1971.

Schutz, R. E. Programmatic instructional development. SWRL Educational Research and Development, Los Alamitos, CA, Professional Paper 11, 1970. (a)

Schutz, R. E. The nature of educational development. Journal of Research and Development in Education, 1970, 3, 39-63. (b)

Schutz, R. E. Educational R&D program management considerations. SWRL Educational Research and Development, Los Alamitos, CA, Technical Report 40, 1972.

Wolfle, D. The home of science. New York: McGraw-Hill, 1972.

Comments on Schutz Paper by David R. Krathwohl

The perspective which Dr. Schutz brings to this topic is both useful and important. As one of the more successful educational laboratory directors, he has first hand, in-depth knowledge of the development process. Further, he has long been a student of the topic. In his own laboratory he has commissioned persons to draw generalizations about the development process as they saw it operating. Further, as is particularly apparent in this paper, he is a student of the work on development wherever it appears in or outside the field of education.[1] The paper is a rich mix of both of these perspectives covering a wide range of aspects of development.

The importance of the paper may not be immediately apparent to all readers. A restatement of what appears to be conventional wisdom rarely is; but the winnowing of wheat from chaff among the wide range of "conventional wisdom" statements that could be made is its main contribution. In many respects, therefore, Schutz' paper will be most appreciated by those who have already tried to generalize about the development process.

These same individuals will also least be bothered by the large number of overgeneralizations which are sprinkled through the paper. They will appreciate the zest with which Schutz breezes to convictions about the necessity of development breaking with past practice, about the kind of start which development has made in business today, about the difference between

---

[1] It is worth noting in Schutz' writing prior to NIE's authorization, that Glennon's work on the development process as cited is especially worthy of attention. This lends extra credence to Schutz' current use of the material and bodes well for NIE that its first director has been an analyst of the development process.

the research and the development points of view, and many other topics. For me, the overgeneralizations serve as signals to pause and try to appreciate the author's point of view; others may find them distracting to the point of reducing the value of the paper.

In addition to describing the development models, the paper makes important points about the labor intensive nature of the educational process; the poverty stricken nature of current classrooms contrasts starkly with the thousands of dollars per worker which industry invests in order to achieve its goals. The fact that much of what is now done by the teacher could, with adequate development work, be achieved even better by technological means is becoming apparent as the products of that technology become increasingly available. Schutz' contrast of development with the research model is useful but overdrawn. In this context, the contrast of experimental and control groups as a base for research with the "cybernetic approach" as a base for development mainly puts a fancy name on the use of a group as its own control, a well known experimental technique. Nonetheless, there probably are some differences in attitude, interest in detailed tasks and problem solving orientation between researchers and developers, which make the point worth considering.

Before raising the major question about the paper to be discussed at length, two notes should be made. First, basic to the whole development process is better knowledge about how to bring about change in institutions. This was commented upon in almost every one of the papers prepared for the symposium. The products of research and the most complex development process are of no value if they are not used. Part of the difficulty of resolving this problem is that it quickly gets into the realm of value conflicts of the kind noted by White and Bissell in their paper. What is the proper role of the teacher? What does "teacher proof" curricula do to the teacher and the teaching process?

Why don't teachers make use of the carefully prepared alternatives which are set before them in some development projects? These and many other questions relating to the process of educational change must be examined in order for the development process to approach its potential.

This knowledge is particularly critical to the final phase, the dissemination-operation aspect of the model Schutz puts forth. In this connection it is worth noting that the NASA and DOD models form the basis for Schutz' education model, and both of these organizations have firm control of the dissemination-adoption process. They can thus afford to allow it to be held to a final step. Educational development in no way controls this process, and, as will be brought out later in this discussion, there may well be advantages in merging the dissemination-adoption process with the middle developmental steps. Thus, further knowledge of the change process is essential even to understanding and describing the steps in the educational development process itself.

Second, Schutz suggests that industry is the proper locus for development work. Certainly with respect to hardware, it already is the locus. Problems have existed, however, where there have been inadequate profit margins, such as with the new complex modularized curriculum materials. This is also true where the dissemination of an idea is more a process than a tangible product. Schutz particularly finds the university campus a poor place for development. But it may well be that only on the university campus does the crazy kind of idealism flourish that is willing to try to disseminate processes as well as products, or to deal with products which have a real potential for improving education, but which are not likely to make a commercial profit.

There are many good points about the paper, but there is also a special problem. This is the adherence to a model of development which is built

around a sequence of development steps, each preceding the other. In this approach one first does the research; the development in several successively more extensive steps follows, with installation and operation the culmination of the process. This and similar approaches which follow a strictly rational framework of problem solving, like PPBS and management by objectives, have a great deal of appeal, and are not without their successes. The question to be examined, however, is whether and how this model fits the education development process, as the paper advocates.

Although developers may feel that the example is not the best one, the beginning of the programed instruction movement does provide a useful example of this process which is within the memory of most readers, and which graphically illustrates some of the problems. Programed instruction began with Skinner's pointing out the research base. These findings indicated that small learning steps plus immediate reinforcement in the form of the right answer were a model for learning materials which should work. Many such programs were created and field tested, some going through a number of revisions. Indeed, they showed that the students did achieve. Though numerous programs were of the homegrown variety. Some came into being when publishing firms set up development centers, and established teams to coordinate and evaluate field testing. But, as these materials flowed into the dissemination market, they failed to be installed, and never achieved the potential that was proclaimed for them. For example, of the over 352 programs listed in the catalog of programed materials issued in 1963, very few are still in extensive use.[2] In part this was due to the fact that there is an artistry and skill in the development of this material, which was not encompassed in the rules which were developed from the research to guide its creation.

---

[2] Programs 63, A Guide to Programed Instructional Material Availability to E' cators by September 1963. Catalog No. FS5.234:34015-63. Washington: erintendent of Documents. 1963

But it was more than that. One could argue that the failure was due to inadequate field testing; one never got past the Hawthorne Effect of the experimental situation. It might have been that the side effects of student boredom and lack of motivation were not adequately observed by the developers. Perhaps the installation procedures were faulty and did not involve the teacher, who felt displaced and managed by the materials rather than vice versa. Teachers were not trained in ways of using these materials as complements to, not a replacements for their skills. Teachers are too oriented to getting their success feelings from standing in front of class lecturing, and they could not play and did not like the role of waiting to be involved only when a student didn't learn from the program. Probably all of these criticisms have some validity. But at the same time they also point up some of the common problems of this approach: inadequate research to guide the development process, so that unforeseen problems occur; difficulty ingetting field tests under realistic conditions; inadequate attention paid to side effects, since the major evaluation is aimed at whether the product achieved its goal; inadequate attention to the human aspects of engineering the development into the classroom in terms of current mores, attitudes, logistics, administrative structures and the like, and in particular inadequate knowledge about how to install such products in ways which are self-reinforcing and grow on their own, "take off" so to speak.

This is an example of the linear model, though not an example of its best application. It was chosen in part because it does point up a number of the flaws. Though not all the flaws are inherent in the model, we just don't know enough yet to overcome all these and others that occur in any given application.

In order to gain a perspective on this model, it is helpful to contrast it with the so-called multi-dimensional, or as I prefer to name it, the

multi-stage model of development. This is the model advocated by the team of foreign examiners from the Office of Economic Cooperation and Development[3] in their examination of the U.S.'s educational R and D efforts, and in the discussions with the member nations of OECD in Paris which followed. They criticized the apparent wide-spread adherence by our own laboratories and R and D Centers to what they termed the "linear" approach or production line. They felt we were missing many of the values of the multi-stage approach.

An example of this approach is the success of the British schools in developing and implementing the open classroom type of school with no formal development organization, only teachers, schools, teacher centers, very good headmasters and headmistresses, and Her Majesty's Inspectors. In this approach each teacher, each school, and a variety of teaching centers are all working on the same general problem, coming up with solutions to parts of it, exchanging information on the best of what they have developed, adopting the best of the new material, and then developing further on a wide front. Installation is almost simultaneous with development, and is not a process which is engaged in solely at the end, as in the linear model. Like the linear model, it uses the cybernetic approach to development, successively modifying products to achieve the desired effect. Evaluation in the linear model is more likely to involve sophisticated evaluation devices and technicians to develop assessment and evaluation process. The multi-stage model is more likely to rely upon informal devices or the teacher's own observation and judgment of what does and does not work. The linear process is the more likely to have some kind of summative evaluation which permits one to have objective evidence that the goals are indeed achieved. The multi-stage process may have no such evidence; indeed it has no real

---

eviews of National Policies for Education: United States. Paris: nization for Economic Co-operation and Development, 1971, p.67.

summation--it is an ongoing, continually evolving process. Reliance is placed on the fact that others thought it looked good, have tried it, and judged it to be successful. It assumes that teachers can adequately sense when they are and are not achieving their goals.

The multi-stage method makes use of the common sense and wisdom of those on the firing line. It assures that the economics and logistics are feasible and compatible with the system; that this is something students and teachers can and will want to do. It serves to stimulate those in teaching positions to grow and to become more competent. Nobody more than teachers is aware of unwanted side effects, and so these are immediately picked up. Alternation between research and development is continuous as new thrusts are made. As already noted, installation is almost simultaneous with development, and does not wait until the end as a separate step.

There is another kind of contrast of the linear to the multi-stage model which needs to be made, a comparison on their political viability. Certainly, any model which is recommended as a base for political funding must meet that test. From the Federal point of view, a problem with any developmental model is to maintain sufficiently clear direction and to have sufficient momentum at the right times that the project can be carried to completion. Many projects of the U. S. Office of Education, among which Project English and Project Social Studies are but two, have clearly failed in this regard.

Schutz' text indicates that the development process, including research, take between ten and fifteen years. In one of the tables he notes lower delivery times where the research has already been done of three to five years for media and technology, and five to seven years for major curriculum projects. Even these lower intervals are ong periods, considering that congressional elections occur every two years, that presidential elections occur every four, and that at each of these points there is a need to "show and tell."

Maintaining direction and momentum over delivery times of ten to fifteen years seems extraordinarily difficult to sustain. Further, school teachers, parent, students and administrators are clamoring for solutions now! Their support is extremely helpful, perhaps even vital to getting the kind of funds needed to maintain development at any stage, but certainly at its very expensive final stages. Thus, factors like interim products, a widespread base of support from users, and some interim evidence of success are very important factors in political viability of a development model.

Both the linear and the multi-stage model can produce interim products, though this is a more natural part of the multi-stage process, and may have to be contrived with certain product lines in the linear model. A broad base of support is built into the multi-stage model, as is belief in the success of the product among those who make up that base. Such support in the linear model comes only from the field tests. If laboratory personnel descend from the laboratories only to use the field to gather their data, and return to the laboratory without adequate involvement of the school personnel, students and parents in their study, that base of support simply will not be there. Unfortunately, this is a common pattern; teachers and those on the firing line tend to be involved only at the field test stage, and often only peripherally then. They have no long-term stake in the success of the project.

With the multi-stage model, through the spread of the communication network as additional teachers and schools seek to become involved, there is a natural build-up of momentum as the project proceeds. This continues so long as the approach which forms the basis for the development effort continues to pay dividends. The build-up is not quite as natural with the linear model, although it can be managed with the availability of interim products, with the involvement and the gaining of commitment of personnel

in the field tests, and with the massing of evidence of success. Unfortunately, the latter is more difficult for Education in terms of convincing Congress and the public than in other fields. We have nothing as persuasive as NASA's landing on the moon or the DOD's building a successful missile. We especially do not have adequate measures of success outside the realm of cognitive objectives.

Thus, both of these methods can meet the criteria of political viability, but it is a natural part of one, and must be taken into account in terms of appropriate management in the other.

Finally, the multi-stage model, as the name implies, blurs the distinctions among stages, distinctions which even the most hardened linear model convert will admit are hard to maintain in practice. With out present limited knowledge of the educational process, for example, it is almost impossible to be sure that one will not have to return to research once one begins to engage in development. We just can't be sure that all the needed research has been done. The desirability of blurring the distinctions between development and dissemination has already been noted.

How well then does the linear model elaborated in the paper fit a desirable development process for education? Clearly the contrast with the multi-stage model suggests that the latter has real strengths to be reckoned with. But it must be clear that the multi-stage model is much like the method that we have depended upon for years to bring about change in education. While that by no means entirely discounts its value, it does suggest that except as it were engaged in by persons outside the present educational system, it is unlikely to result in anything which is a rapid, radical departure from what we presently have. Incremental change is desirable, but many cry out and are impatient that education changes so slowly.

Neither is the multi-stage method likely to result in the development of

complicated and sophisticated technological products. These are more likely the result of a more nearly linear type development process. In general, it appears that products which are labor intensive may develop well in the multi-stage pattern.

As a final comment on the linear model, some quotations from the OECD review cited earlier seem especially appropriate. In summarizing the discussion with regard to American conceptual and management models of R & D, the OECD examiners posed three questions: "First, is the R and D process necessarily a neat and logical sequence of research, development, dissemination and application?" The American team's reply was, "Obviously not." "Secondly, is there any necessary preference for theoretical-deductive as against empirical-inductive models of R and D?" The American answer was again, "no." "Thirdly, can and ought the 'consumers' be involved in goal identification? The American Delegation thought that 'the researcher ought to be one of many seeking to arrive at national goals of importance to education.'" The OECD examiners pointedly stated, "the 'many' must include teachers."[4]

They continued: "Thus far there are no disagreements in principle. The disagreements concern the extent to which it is possible to create management procedures for ensuring that, in the end, theories are established in practice, and new techniques and materials are tested and disseminated, without destroying the initiative and participation of the classroom practitioner."

Further on, they operationalized their concern in terms of the products ("packages") of development: "If the packages are good, teacher determination of them can become weaker. The more they meet consumer needs or desires, the more they reinforce the teacher's own role as consumer rather than as a producer." Then summarizing at another point: "American schools are not the

---

[4] Organization for Economic Co-operation and Development, op.cit., pp.25-26 ff.

best examples of education development arising from a creative teacher force. The young, but strong, federal initiative in R and D could either consolidate the role of teachers as 'consumers' or reinforce them as active participants in educational development."

The above analysis suggests that Schutz' paper is remiss in considering only one model of educational development. There are other models, the multi-stage one discussed here, and the many positions between the linear and multi-stage which could be taken. Each has its advantages and disadvantages. Except where the U. S. Office of Education has forced them to do otherwise, many of the laboratories and centers appear to be using models which lie between the two extremes analyzed here. Anyone considering development in education would do well to consider making a conscious choice of a model to fit the kind of product he hopes to develop, the political situation that he faces, and the many other factors, such as staff, field facilities, resources, and the like, which must enter the decision. A further broadening of the Schutz discussion on development, extensive as it already is, therefore, is in order.

General Discussion of Schutz Paper and Krathwohl Comments

The discussion on development activities in education was concerned with the definition of the products being produced and the market being served; the best locus in which to conduct these activities; and the role of the government in sponsoring and overseeing development work. Development in education differs from development efforts in other fields because of the diffuseness, the decentralization of activities , the difficulty in defining and evaluating products and the problem of determining the users of these products. Other fields, such as DOD or NASA, tend to have a more tangible product, to operate through centralized office and staffs, and to be able more easily to measure the impact of their development efforts. In addition, there tends to be more consensus on what their goals should be and, until recently, the levels and types of commitment to be made in the development of new products and processes to facilitate the attainment of these goals.

In education, however, the products can be as intangible as ideas or concepts or as tangible as new textbooks or teaching machines. Further, in determining what should be the substantive and time frame for development activities in education, a number of potential user groups are involved, though few are directly consulted about or involved in development efforts. This is especially true of teachers and administrators, who would logically be expected to know what is needed to improve learning in the classroom and who can ultimately affect the manner in which innovations are adopted and implemented. Their involvment in the development process might increase their committment to the innovation at later stages, and could serve to minimize subsequent implementation problems due to lack of practical ledge and insight on the part of the R & D personnel.

The apparent failure of education R & D to have an impact on activities in the field has forced many in the education development area to look to other fields for models or means of improving the implementation process. In addition to the government's space and defense efforts cited earlier, developers tend to offer the pharmaceutical and textbook publishing industries as examples of profit-oriented corporations which appear to be successful in developing efforts. However, although the drug industry engages in long term basic research, the majority of development activities are very much like those of the textbook publishing companies and for largely the same reason. In both cases, the innovations tend to be incremental in nature and the R & D efforts to have a relatively short time frame and a relatively high probability of success. The profit motive, while stimulating invention in some ways, tends to dampen it in others. Those projects which have too low a probability of success or which require too long a payback period frequently are not undertaken.

This raises an important question: does education, with its essentially conservative environmental and organizational pressures, need to introduce additional potentially conservative elements of the profit-oriented market model for the development and dissemination of educational innovations? Is industry, after all, a better locus for educational development activities than those presently being used? If so, what role should the government play in sponsoring and overseeing development work? These, and other questions, need to be raised and answered before education adapts development processes and procedures from other areas. Any changes in the development process (and changes do seem called for), must be made with a better understanding of the existing organizational and environmental pressures on school systems and researchers if they are to be effective.

To achieve this kind of success in the development area, it will be

important to define the various development strategies currently operating in the education area. For example, are the current development strategies consumer, interest group, or school based, or do they fall along some such continuum? Should they have such a focus? If so, how do we identify these different strategies and choose among them? How does one obtain and retain support for development activities generally and for specific innovations? Are these bases of support different? If so, how do we seek their help? As is obvious, these questions do not lend themselves to quick answers or patent solutions. Rather, it will be up to the NIE over the next few years to examine these issues and to experiment with one or several of these development strategies.

THE DISSEMINATION OF EDUCATIONAL R&D PRODUCTS: RESEARCH AND POLICY ISSUES FOR THE FEDERAL GOVERNMENT

Daniel Weiler*

The Rand Corporation, Santa Monica, California

I. INTRODUCTION

This paper discusses issues related to federal government programs for the dissemination of educational R&D products, and suggests general policy and research guidelines in this area for the National Institute of Education.

Federal dissemination activities are spread throughout numerous agencies and offices that play some role in educational research and development, principally in Washington, but at the regional and local level as well, including many federally sponsored experimental or "impact" programs such as Experimental Schools or Title I. The formal institutions of dissemination are largely centered in the National Institute of Education, in particular with the National Center for Educational Communication (NCEC), which manages the decentralized ERIC clearinghouse network, as well as other dissemination programs. The "product" of most education research is a document which presents the conclusions of a study or an evaluation. The product of actual development can be (a) a way of organizing and structuring some set of behaviors that are designed to help make education more efficient or productive (e.g., PBBS, individualized instruction), (b) a physical product designed to accomplish the same objectives (e.g., a new science kit, a new building design), or (c) some combination of these (e.g., the Sullivan reading program). Most "physical" products are developed by commercial publishers or other segments of private industry; many behavioral products are developed by these companies, but many are also

*Helpful comments on an earlier draft of this paper were made by David Clark (University of Indiana), Arthur Cohen (ERIC Clearinghouse for Junior Colleges), and John Pincus (The Rand Corporation).

developed by private and nonprofit research corporations under federal or local contract. The products of educational R&D are diverse, but most research does not lead to development--it leads to a written study or report that is rarely carried further.

In addition to the products of research and development there exists something called educational information--which is neither research nor developed product, but simply, as the name implies, some kind of information about education--about research in progress, about practice, about debate and argument among professionals on various topics, and so on.

Federal dissemination efforts embrace all of these possibilities--research products, development products (which imply some preceding research), and information.

This paper first summarizes key problems now faced by education practitioners who might wish to use R&D products or information, and the factors that contribute to these problems. The paper then goes on to consider a general research and policy posture for NIE in this area, relevant research topics of interest, and the question of appropriate NIE policy for the short run.

II. THE PRESENT SITUATION*

Practitioners who try to use educational R&D products and information face three major problems.**

First, practitioners often find it difficult to identify, locate and acquire potentially useful materials. District-level specialists in large school districts with well supported research libraries, and some practitioners who live in the vicinity of special information centers, have somewhat less difficulty than others. However, even for these somewhat special cases, and certainly for most practitioners, these difficulties are very real.

Second, practitioners find it hard to get the help they need in order to oversome these difficulties. Many educators do not at first have a clear understanding of the relationship between their problems

---

*In reviewing current problems of access to educational R&D products, we argue as if it can be taken for granted that most education practitioners are motivated to find and make active use of the best available information. As we shall see, this assumption (and dissemination activities that rely on it) is open to question.

**Education practitioners can be specialists at either the district or building level, including administrators as well as specialists in testing and guidance, curriculum, or personnel training; or generalists, meaning largely classroom teachers. However, the distinction between specialists and generalists does not appear to have an important bearing on dissemination requirements, for the kinds of dissemination activities required to meet the needs of the specialist and those required to meet the needs of the generalist are sufficiently similar to make the distinction uninteresting from the perspective of federal policy [Greenwood and Weiler, 1972].

The research community faces a number of related problems in attempting to use educational R&D products or information. However, since the needs, resources and training of researchers (both basic and applied) are quite different from those of practitioners, this paper does not discuss these problems in detail. In practice it seems likely that federal policies designed to improve dissemination for practitioner users will lead to improvements for the research community also. [Ibid.]

and their requirements for R&D products or information. They may need assistance in defining problems clearly, or in understanding the character of the difficulties they have identified. Even with adequate problem definition, most practitioners are not well trained in the appropriate search techniques, and will require the assistance of a professional specialist in locating information that might be useful.

Finally, even with problems adequately identified, relevant search areas pinned down, and information in hand, most practitioners will need some assistance with the practical interpretation of the products they have found. They are usually being asked to use either printed material or microfiche, in the case of a research report, or a printed manual of procedure in the case of developed materials. Printed materials, however, are necessarily somewhat abstract, and often fail to capture the specific character of the implementation problems faced by practicing educators. Moreover, these materials are not interactive. If upon reading a report or research study a practitioner formulates additional questions for which answers seem required before he can take action, he cannot query the printed page, and will usually look for additional assistance before acting. At present he finds this assistance difficult to obtain.*

Finally, the practitioner who has successfully identified candidate R&D products related to his interests finds that he often has little guidance as to their probable utility, reliability or validity, and must essentially make these judgments for himself on the basis of his own instincts.

---

*In California recently, the state legislature mandated the introduction of program budgeting procedures in every school district, and hired a firm to develop detailed implementation guidelines. The guidelines were duly circulated; local districts immediately began searching for experts who could help them interpret and implement the new procedures.

III. CONTRIBUTING FACTORS

Some thought has been devoted to these and related problems in recent years, and it is now possible to summarize a number of reasonable explanations for this situation. What follows is an attempt to provide such a list as the basis for further discussion of federal research and policy in this area.

1. Many R&D products and much of what passes for educational information are of little or no practical utility. Thus, while it is appropriate to consider possible changes in current dissemination activities, it is important as well that we understand the very real constraints on the system posed by the material it has to work with. Even the best dissemination system in the world will be of little use to anyone if it has little of consequence to disseminate.

2. At the same time, research or development products of potential value to educators do exist, from evaluations of specific programs, to studies of the reliability and utility of achievement tests, detailed descriptions of new curricula, reports of practical solutions to administrative problems, and dozens of others. Thus, while it is a truism that there is great need for better research of direct utility to the practicing educator, there are doubtless useful products in the system as well. However, there is at present virtually no organized effort to distinguish these products from those which are less useful. For practical purposes, the dissemination system now treats all information as having equal value. Hence, though the average practitioner may be poorly equipped for this task by training and temperament, and has little time to spare, he must invariably perform the screening and assessment function for himself.

3. An enormous amount of educational research has been accomplished during the last several decades, leading to the publication of a great many books, reports, professional articles and occasional papers. However, only a comparatively small effort is made today to review and synthesize research outcomes across the various areas that might be of

interest and practical utility to professional educators.* In part perhaps because of this low level of effort in research synthesis, there has also been little attempt to translate the results of research into practical program advice for professional educators.

4. We have not yet developed adequate procedures for the collection and validation of information on the best of current practice--although it is this kind of information that educators rely upon most heavily--nor for the translation and extrapolation of successful local practices into general policy guidance for the implementation of educational programs designed to meet the varied needs of different student populations.**

5. The practitioner who tries to identify and acquire a helpful research or development product faces multiple, partly redundant sources of information with widely varying visibility and accessibility--federal, state, local and private--with virtually no coordination of disparate efforts. He can try district resources; if those do not yield results, he can look in his local public or university library. But he may also need to go directly, by mail, or by phone to an ERIC clearinghouse, a local Information Resource Center, a county-run curriculum lab, or to any one of dozens of other possible sources [see, e.g., Wanger, 1971].

6. If the practitioner overcomes the problem of multiple and physically remote access points, he must still cope with confusing arrangements for system entry from any given locale. He faces a plethora of different indices, card files, microfiche storage systems, technological aids and guides to information. These various sources of entry to the information system have little procedural or structural uniformity, for they were not created through system-wide agreements

---

*NCEC attempts to do some of this job through its Targeted Communications Program and support of various ERIC Clearinghouse Information Analysis products. While these efforts appear to be fairly popular with practitioners who are aware of them, their quality and scope are uneven, and they have not received a high level of federal funding support (System Development Corporation, 1972; Greenwood and Weiler, 1972).

**Objections to the idea that this kind of work might prove valuable are discussed briefly on pp. 18-19.

on the ground rules for decisions pertaining to indexing strategies, selection of key descriptors, or subject matter partitioning.

7. The practitioner who seeks the advice and intercession of professional experts finds a system that is often passive, that is not structured to respond in depth to direct inquiry from the user, and makes little attempt to anticipate information demand and consumption patterns. Marketing mechanisms--the identification of client needs, and an active attempt to meet those needs--are rare. To some extent, this appears to reflect a system bias toward the research community as the client, and away from an operational mode of dealing directly with the practicing educator as the principal user. It could be argued that this is in some respects a sensible way of dealing with multiple and partly redundant resources, for researchers are better able to wend their way through this complex system than are practitioner users. At the same time, this could be said to represent acquiescence in the creation of what might be described as a "closed loop" for research information, wherein the research community uses the system for assistance in the creation of new research results, which in turn go back into the system and again to the research community. There are, to be sure, many breaks in the loop--applied researchers do communicate directly with education practitioners and practitioners do have access of sorts. Nevertheless, the apparent system-wide bias against direct response to the practitioner user effectively "freezes out" many potential practitioner clients from timely access to the information they seek.*

Federal Government Policies and Assumptions

We have discussed problems faced today by practitioners who seek educational R&D products or information, and factors that contribute

*The new education extension agent program is designed to remedy some of these problems, but it is not yet clear what its impact will be, since there have not been many accompanying changes to related elements of the dissemination system [Sieber, Louis and Metzger, 1972].

to these problems. We turn now to a discussion of federal government policies and assumptions that appear to have played an important role in the creation and maintenance of present dissemination arrangements. We then go on to suggest some research and policy guidelines that might be considered in order to resolve these problems.

Much of the present information dissemination system was inherited by the federal government either as it now exists or in some incipient version, when the government first expanded its education activities significantly in the middle 60's. The multiplicity of sources, for example, is a natural consequence of the decentralization of education in the United States. When the federal government did initiate dissemination activities of its own, it identified and moved to ameliorate an important problem--the absence of a central collection and indexing agency for information that did not find its way into professional journals, and was therefore lost forever to most potential users. The limited sponsorship of selected information analysis products, such as bibliographies, research reviews and state-of-the-art papers was also begun. These initiatives could not have been adequate to the task of overcoming the problems we have discussed above [Burchinal, 1968].

Federal policies to date have been characterized by a relatively low level of total effort, and by reluctance to pursue objectives that go much beyond the limited goal of providing a passive archival system. At the same time, practitioner demands on the system (requests for direct assistance) have led on the one hand to a variety of uncoordinated efforts to respond ad hoc, and, more recently, to a more formal response through the creation of the education extension agent program. The picture that emerges on balance is one of some confusion about the appropriate clients for federal programs and the appropriate objectives of federal policies. For example, while regional laboratories have been supported by the federal government in their efforts to develop improved curricula for the public schools, federal efforts to disseminate information about the results of these development programs have been relatively low key, and it is difficult for a practitioner

today to acquire pertinent information about this work by querying the federal dissemination system.

While many decisions must of course await better information about the nature of user needs and the best ways in which to meet them, the federal government has made relatively poor use of available knowledge about possibilities for improvement--from existing research studies, from the field operations of federal and other (state, county, private) components of the information dissemination system, and from informed judgments of professionals who work in the system. Regular system evaluation procedures have not been implemented, and there has been no consistent program of research designed to lead to system improvement. Federal policymakers may have been either unaware of many of the problems discussed above, or unable to point to ways in which key problems could be resolved.

The federal government appears to have been making some key assumptions that may be unwarranted. First, the government appears to have assumed that there is an abundance of useful educational information and many good R&D products. In fact, this does not seem to be the case.

Second, the government appears to have assumed that the passive, archival mode for the dissemination of R&D products can provide adequate access to the information that is needed, for both the research and practitioner communities. The assumption also seems unwarranted.* Multiple and physically remote access points and nonuniform, poorly designed search tools have in fact made it quite difficult for both researchers and practitioners to have convenient access to the information they seek.

Finally, the government appears to have assumed that adequate access will in turn be sufficient to insure that education decisionmakers at the local level will make use of existing products and information in

---

*It now appears to be assumed that the provision of system-user intermediaries such as education extension agents will further insure adequate access for the practitioner community (see, however, footnote on p. 7, above).

order to improve education. This assumption would appear to ignore the realities of the actual incentives of educators to make decisions that could result in changes from current and accepted practices.

If the argument presented so far is valid, it is not unreasonable to suppose that the institutions and policies for dissemination that have emerged at the federal level, while useful in many ways, do not yet meet the real needs of educators today. However, they may provide the essential basis for moving to an improved system in the years ahead.

IV. NIE'S RESEARCH AND POLICY POSTURE

Two things are clearly required--decisions and information. Decisions are needed to elaborate a clear federal policy in this area, information is needed to help validate, refute, or amend the hypotheses and assertions posited here and elsewhere about dissemination activities, and to provide fresh insights into ways in which improvements can be effected. The NIE might think in terms of four broad steps:

1. Develop a clearer understanding of the range of activities, products, and services that now exist. We now have, or can obtain in short order, most of the information we need in order to lay out the spectrum of existing practice in the dissemination of educational R&D products and information. The proper frame of reference is not just federal activities, but the broader perspective of all existing educational information resources--federal, state, local and private. Before any subsequent decisions can be made, and before it will be possible to know where to direct research efforts most efficiently, we shoud try to have as clear an idea as possible of the range of services, resources, and institutional arrangements now in existence.

2. Make an initial assessment of ways in which we would like to see these activities, products and services improved. In effect, this requires at least a tentative decision about what a national educational information system might best look like. At one end of the spectrum, for example, something like the present arrangement might be viewed as most appropriate, in that the decentralized nature of educational practice may require decentralized, multiple, redundant sources of educational information. In this case, the cost of eliminating redundancy and confusion in the system would be viewed as exceeding the benefits that could be expected from such an effort. Alternatively, it might be argued that present arrangements are wasteful and inefficient, and that they hold little hope for providing practitioners with the information they will need in the years to come, particularly with anticipated growth in both the amount and quality of educational R&D products over the next decade. These judgments are certain to occasion

a good deal of heat, and should indeed be the subject of widespread debate. Nevertheless, they are a key step, for judgments about desirable directions for a national educational information system, while they will not be immutable, must surely be made, and made early, if federal programs that will inevitably be part of a larger set of activities throughout the country are to be rationally cast within the framework of explicit objectives for the nation as a whole.

3. Make tentative decisions about the proper role of the federal government in leading, sponsoring or implementing programs designed to effect desired changes. This implies a requirement for analyses of the extent to which desired practices and services cannot or should not be provided elsewhere, a review of present and anticipated resources available to the federal government, and difficult judgments about what it would be appropriate for federal ambitions to embrace. On the one hand, for example, it could be argued that the federal government should not be interested in any further expansion of its presence to the local level, and that its activities should be restricted essentially to funding state, county and local efforts to disseminate information, together perhaps with some modest federal effort roughly equivalent to that which is being made today. Alternatively, it might be concluded that it would be wrong to count on local initiatives to provide the necessary services in all cases, and that only ambitious federal programs will ensure uniform and adequate access to educational information in every school district.

4. Assess the utility of various institutional arrangements at the federal level for accomplishing desired objectives. A decision about the institutional form for federal efforts follows logically--but must follow, not precede--the first three steps described above. Institutional structure should be dictated by antecedent decisions about federal objectives, and these decisions must in turn be preceded by a clear understanding of the kind of national information dissemination system that would best serve the requirements of both practitioners and researchers.

Each of the broad steps described above implies a requirement for various kinds of information about the best way to proceed. Much of this information is already available and can be pulled together in order to assist policymakers; much remains to be gathered. We take up this topic below.

V. MAJOR RESEARCH QUESTIONS

We have suggested that a logical first step would be the development of a clearer understanding of the range of activities, products and services that now exist. Existing research studies and professional judgments could also be mined for insight and advice regarding federal programs in this area. In addition, a carefully thought out and coordinated program of research, institutional design, and experimentation could be initiated over the next several years in order to better inform and continue to test tentative decisions that may be made in the interim. Such a research program should address at least the following major questions:

1. How do the incentives of education practitioners to seek, acquire and use educational R&D products and information vary with variations in

- the substance of the information,
- information format and style,
- product availability,
- strategies for marketing information products,
- practitioner roles in the education system,
- practitioner objectives (e.g., maintenance, improvement, reform), and
- opportunities to put products and information to practical use?

We can be reasonably certain that incentives to seek and use educational information and R&D products do not exist independent of incentives and opportunities to make and implement policy decisions about educational programs in the classroom, school, or district. These may be incentives and opportunities to maintain educational programs at an existing level of quality, to implement new and improved programs, or to reform (or retain) inefficient practices. What matters is that these incentives and opportunities often precede incentives to acquire and use educational information, and can determine the extent of the information acquired, the kind of information sought, the speed with

which it is desired, the format considered most appropriate, the style of its acquisition, and the uses to which it will be put. While we can say this much with reasonable certainty, we do not know the ways in which these variables interact. It is probable that incentives and opportunities are related to perceptions held by different actors of:

- o opportunities for professional advancement (status, esteem, income);
- o the risks that may accompany a decision to act or withhold action;
- o opportunities to pursue deeply held beliefs.

Thus the need for and uses made of educational information and R&D products are related in some way to the capacity and willingness to act in education, and the nature of this capacity--the incentive structure, the opportunities, the origins of policy decisions--is something we must know more about if information dissemination is to have a direct bearing on the process of educational change.

2. What institutional arrangements would be most efficient for meeting the range of product/service objectives selected by the federal government? We have argued that form should follow function--but how can different functions best be performed? Research in this area would attempt to determine the most appropriate source of support for information system components, the function and client focus for each component, the research and data collection responsibilities of the system, and the location of various components and sub-components (e.g., centralized or decentralized locations for different functions). The nature of the services to be provided by the system will influence the number, size and location of system access points for various clients as well as the responsibilities of system management and professional staff. Procedures for the management and coordination of system activities would have to be considered, as would system capacity for growth and renewal. System design would reflect not only decisions with respect to overall objectives, but technical judgments concerning the best way to maximize information processing and communication efficiencies, attract the most

skilled and dedicated staff, and develop the most effective relationships with other institutions, both federal and non-federal. We are some distance from being able to judge what shape such a system should take, what its component parts should be, where they should be located, what management procedures should be considered or what kinds of staff to look for.

3. How should specialists of various kinds be used to assist practitioners to make optimum use of available R&D products and educational information? Some practitioners today have access to information specialists who mediate or negotiate their requests for information in various topic areas. In addition, NCEC is now planning to put a number of education extension agents in the field, in part to fulfill this function. We have not yet collected and analyzed the bulk of available information pertaining to existing services of this kind, and little attempt has been made to study the impact on educational decisionmaking of the provision of varieties of services with personnel who have different kinds of training. We are therefore not yet clear about what kinds of people such specialists should be--what kinds of training they should have, what organizational affiliations or institutional character they should adopt, and what kinds of services they should provide. We do not have a clear understanding of the criteria that should be applied to their selection, nor of the arrangements that should be made for their continued training. A university reference librarian is an information specialist; so is an Information Resource Center specialist who has been to a special training program at a regional laboratory. What is the difference between the kinds of services these two people can provide; how effective are they in their different roles; how much has their training cost; who can they best service? We should try to look at these and at other existing "personnel" models in order to get a better idea of the most desirable characteristics for this critical component of the dissemination system.

4. Is it possible to translate research results into practical policy guidance for educators? Educational R&D, like that in many

social policy fields, is characterized by a division of functions and specialties which largely removes the research community from the responsibility for drawing operational implications from the results of its work. While some applied research does make a serious effort to take advantage of the results of basic research into human behavior and social organization, there remain as yet few interdisciplinary attempts to design specific educational programs (together with detailed specifications for implementation) that are based directly upon a spectrum of education and social science research results, and could be tested in the light of those results. It appears, for example, that research has not yet been able to identify any particular combination of education resources that is consistently and unambiguously related to educational success [Averch, et al., 1972]. At the same time, because contradictory research results do exist, and because research tools are not adequately refined in many areas, more information of a practical nature--testing the application of research hypotheses as they relate to operational programs--might help to eliminate some of this ambiguity. But it is not clear that this goal, though it may be desirable, can be attained. It is difficult to imagine what the practical program implications might be for many research findings. We need more work in this area in order to see whether or not guidelines for the translation of research into practical programs can be established, and to refine ways in which program outcomes can be tested in terms of original research conclusions. One way in which to begin such work might be to devise institutional efforts to bring basic and applied researchers together with the designers of educational programs, in such a way that program designers are forced to test the theories and assumptions underlying their designs against the substantive knowledge and hypotheses of the research community. Since there is so much diversity of opinion and belief in the research community with respect to fundamental issues of human behavior, and since this lack of consensus is in part responsible for the vacuum into which program designers have moved, this would be an extremely difficult task, but one that could be rewarding not only for program design, but for the research community as well.

5. Can we establish reliable procedures for the wide-spread collection of information on the best of current practice, and the translation of this information into usable models of exemplary practice? This kind of research is extremely difficult. At the least, such efforts might include an attempt to provide descriptive syntheses of the systemic effects of different combinations of resource inputs (including student characteristics), teaching processes, and organizational structures. Information thus acquired might then be inspected for our ability to extrapolate "general rules" for program success under various circumstances, and such rules would in turn have to be translated into practical implementation advice for working educators. Even so, the settings for program replication will be largely unique, implying a requirement for the kind of expert assistance with program implementation techniques that is not readily available today (and where research under 3 above might be of some assistance).

Eventually, a broad information collection effort might have to be initiated, based on decisions about relevant performance criteria, program characteristics of interest and related matters. The machinery for such an information collection effort is not available and would have to be designed. Careful attention would have to be given to the mechanisms through which this information was to be aggregated, analyzed, synthesized, and translated into a product of some utility.

When this research issue is discussed it is commonly objected that we cannot reach agreement on what "good" practice looks like (multiplicity of objectives); that we could not identify it even if we agreed on what it was (crudeness of current measures); and that, in any case, current practice is mostly bad and not to be encouraged, so that this kind of work is not worth the effort. While these objections deserve to be taken seriously, they would appear on balance to lend added support to the suggestion that much serious work of the kind described above remains to be done before it will be possible to judge

the desirability of mounting expensive efforts to collect information on current practice.*

6. How can technology be used more effectively to give both practitioners and researchers better and faster access to the existing knowledge base? One way in which an extensive set of dissemination services might be decentralized would be to create a decentralized technology, tying districts into the services of a regional information net that could respond to requests from remote users. We do not yet know how--or if--this could be accomplished, and this issue deserves further exploration before final decisions are made about the location of federally managed or sponsored information system access points. In addition, little has been done to experiment with ways in which the school building (meaning building administrators and teachers) could be tied directly to information resources. It is widely understood that neither teachers nor principals have adequate time to use standard library resources, even when such resources are physically convenient and accessible. (This is quite apart from the time required at such a location to sort out the confusion of multiple points of entry to the information system.) It has been argued that an attempt to tie schools directly to information would not be worth the cost and effort. This may well be the case, but we know of no experimental efforts to verify such assertions. Indeed, what may be the case today may not be the case tomorrow, for if the NIE is successful in many of its objectives, and if better information for practitioners can be developed by the research community, the products that may be available for teachers and principals may improve in quality substantially over the next decade. Should that be the case, we might want to begin to experiment now with ways in which the school building could be given more direct access to those products in the years to come.

---

*It is hard to resist making the observation that many of the strongest and most eloquent attacks on current practice come from professionally successful, highly educated critics who are products of the public school system.

In addition, education researchers are not well served by present system technology. In particular, we cannot now collect and make available to researchers around the country the growing mass of longitudinal educational and related social science data presently stored on computer tapes in a variety of federal, state, university and private data banks around the United States. Every new research project that attempts to scan available data for new insights, or use data already collected as the context for comparison with fresh information, must start essentially from scratch to search the country seeking access to various existing data files. It is not yet possible--though it may be technologically feasible--for the researcher to have access to a wide variety of data without moving from a local computer terminal. Technological options for providing these services also need to be explored.

7. How can we establish widely acceptable screening, rating and classification procedures for educational R&D products and information? One serious problem now facing the potential user of educational R&D products is that little attempt is made on his behalf to differentiate high-quality products from those which are less useful. It seems obvious that rating procedures cannot simply assign one-dimensional classifications to R&D products; at the same time it is not clear what these procedures should be, nor whether procedures can be agreed upon that will be acceptable to the broad mass of practitioner users as well as to researchers. Fairly complex procedures may have to be devised--procedures that differentiate among various levels of information utility for different purposes and for different users. A first-level screening might be fairly crude, seeking only to establish whether the material is of sufficient professional quality overall to be accepted into the system. These might be comparatively easy judgments to make, as long as the bases for the judgments were made explicit and the professional credentials of the judges were widely accepted. In addition, rejected material would probably have to be placed in a separate information pool for inspection by users who did not trust these judgments. After this initial screening, more refined and complex rating procedures might have to be instituted, so that material could

be rated, not along a scale of excellent to poor, but in a manner that indicated the level of sophistication required of the potential user, the objectives of the research, the scope of the research conclusions, the professional interests of users who would stand to benefit most from the product, uses to which the product has been put, user response to date, and so on. These categories remain to be devised and tested.

The research topics described above can be approached in a variety of ways· it is not our purpose in this paper to present an extended discussion of research strategies or designs. One aspect of this research that does deserve special mention, however--and special consideration when designing a research strategy--is the probable systemic and interactive nature of various activities and outcomes of interest. In particular, key interactive effects may exist between institutional arrangements for the dissemination of information, the use of various kinds of specialists to assist practitioner and researcher clientele, procedures for the collection and validation of information on current practice, and the uses of technology. Ambitious and comprehensive research designs aimed at increasing our understanding of these interactive effects may eventually be required in order to support informed policy decisions.

VI. POLICIES FOR THE SHORT RUN

We turn finally to the question of policies that may be most appropriate in the short run for the support or modification of existing programs and institutions, for though NIE may initiate new research and experimentation in this area, the world is not likely to stand still. Three criteria suggest themselves for weighing policies to be undertaken in the immediate future:

1. The policies should maximize the probability that the government will be in a good position to implement policy recommendations that may eventually flow from a concerted program of research, design, and experimentation.

2. In the event that these research, design, and experimentation efforts prove disappointing, or in the event that research points to organizational, practical, political or economic considerations that militate against the implementation of significant system change, short-run policies should at the same time serve the objective of improving the existing system.

3. The policies should capitalize on existing resources.

The application of these criteria to current options need not yield conflicting policy recommendations. On the contrary, by attempting to meet all three criteria the NIE may be able to place itself in a good position both for obtaining desired system changes in the long run, and for achieving needed short-run improvements.

The first criterion implies a requirement to (a) keep long-run policy options open and, (b) attempt to build substantive, psychological and political momentum towards the possibility of eventual comprehensive system change. The best way to achieve these objectives may also be the optimum way to meet the second criterion of achieving needed improvement in the short run:

First, in order to keep long-run policy options open, the NIE would want to prevent existing programs from hardening into permanent institutions, and existing momentums from growing out of control.

This suggests a policy, for example, of deliberately withholding long-term commitments to existing components of the federal information dissemination system, and an emphasis on experimental variations with new components of that system in order to collect as much incidental information as possible without making a commitment to any one variant. Clearly, this approach would also support (and in some respects be a prerequisite for) a policy of seeking short-run improvements to the existing system.

Second, one way in which to build substantive, psychological, and political momentum for eventual comprehensive system change would be to begin now to make changes that will almost certainly be necessary in any case. Thus, whatever long-run policy recommendations emerge from further research, it now seems reasonably likely that a number of steps, among others, will almost certainly be required, even assuming quite modest federal ambitions for the future:

- More coordination, through federal leadership, of diffuse state, local, federal and private efforts.
- Further consolidation of existing federal programs in order to eliminate redundancy and management inefficiencies.
- Introduction of improved screening mechanisms in order to cut down on the amoung of low-quality material that is collected and disseminated by the system.
- More intensive efforts to collect detailed information on local practice, together with efforts to screen and classify that information for different varieties of users.
- More emphasis on the review and synthesis of existing research.
- Some effort to translate existing research findings into practical program implementation advice for practitioners.

- Continuing attention to the improvement of system management and organization.
- Work on the improvement of access and entry to the system--better search tools, more aggressive product marketing, and more visible entry points.

These are also the kinds of changes one would probably want to make in order to improve the existing system. Here, policies that seem likely to be included in recommendations for the achievement of most long-run objectives probably support short-run improvement goals as well.

Finally, one efficient way in which to approach the task of improving the existing system is to capitalize on existing resources. The most important existing resources are people and technology. There appear to be many bright, inventive people working in the field of educational information dissemination. These people have valuable experience, and they have been responsible for some of the most cogent and constructive criticism of the dissemination system. By involving the best people at an early stage in efforts at improvement--bringing them into the debate and planning stages of the work, insuring open lines of communication as work proceeds--the NIE would surely benefit from their knowledge and experience. At the same time, it would be creating important psychological and political momentum for potential long-run change among a constituency--the working professionals in the field--whose support for federal policies will be of critical importance to the government.

By exploring ways in which technology can be utilized to improve the efficiency and responsiveness of the system, the NIE might well be able to make important system improvements in the short run while laying the groundwork for more ambitious and sophisticated technological changes in the future. For example, it is almost certainly important to begin at an early stage to acquire more information about certain kinds of man-machine interactions, for if long-run policy recommendations should include proposals for substantial new federal investments in technology to serve both the

researcher and practitioner user, policymakers will need high-confidence estimates of the impact and utility of various technological aids. To make these estimates they will require longitudinal data, and should not wait until after long-run objectives and policies have been thrashed out to begin collecting that information.

To recapitulate briefly, it is the thesis of this paper that because of present serious difficulties with the effective dissemination of educational R&D products and information to education practitioners, the NIE should:

1. Undertake a carefully thought out program of research, design, and experimentation that will enable it to:
   (a) Step back and look at the national picture taken as a whole.
   (b) Decide how and why this picture differs from what policymakers would prefer to see, and what the federal government should and can do about it.
   (c) Acquire the necessary information for federal policy decisions in the years to come.

2. Implement parallel efforts to improve the existing system without making any long-term commitments to present or currently planned institutions or programs.

3. Involve the best professionals now working in the system in designing and implementing both long-run research and short-run improvement efforts.

If properly planned and executed, these initiatives should be mutually supportive, and the chances for obtaining beneficial long-run system change should be enchanced without sacrificing attention to immediate problems.

SELECTED BIBLIOGRAPHY

Many of the arguments and propositions in this paper are based on findings or hypotheses presented in an extensive literature on research dissemination, knowledge utilization, and incentives for innovation, as well as on the author's own research and thinking. For the most part, this literature has been used here as the basis for composite summaries of research findings, so that reference to one or a few specific studies would be misleading. Key studies relied on for this purpose, as well as other selected works of interest in this area, are included in this bibliography, which is not intended as a comprehensive listing of the relevant literature in the field.

Averch, H., et al., *How Effective is Schooling?* The Rand Corporation, Santa Monica, California, R-956-PCSF/RC, March 1972.

Baldridge, J. Victor and Rudolph Johnson, *The Impact of Educational R&D Centers and Laboratories: An Analysis of Effective Organizational Strategies*, Stanford University, May 15, 1972.

Banathy, B., et al., *The Educational Information Consultant: Skills in Disseminating Educational Information*, Far West Laboratory for Educational Research and Development, Berkeley, California, July 1971.

Bennis, W. G., et al., *The Planning of Change*, Holt, Rinehart and Winston, New York, 1969.

Burchinal, Lee G., *ERIC--Development Through June 1968*, Department of Health, Education and Welfare, Office of Education, Washington, D. C., October 1968.

Carlson, R. D., *Adoption of Educational Innovations*, The Center for the Advanced Study of Educational Administration, Eugene, Oregon, 1965.

Carlson, R. D., et al., *Change Processes in the Public Schools*, University of Oregon, Eugene, Oregon, 1965.

Chorness, M. H., et al., *Decision Processes and Information Needs in Education: A Field Survey*, Far West Laboratory for Educational Research and Development, Berkeley, California, December 1968.

Chorness, M. H., et al., *Use of Resource Material and Decision Processes Associated with Educational Innovation*, Far West Laboratory of Educational Research and Development, Berkeley, California, February 1969.

Coney, Robert, Vernon Plaskett, Robert Roggenbuck, and Paul Hood, *Educational R&D Information System Requirements: A Task Force Report*, Far West Laboratory for Educational Research and Development, Berkeley, California, March 1968.

Crawford, J. J., et al., *Evaluation of the Impact of Educational Research and Development Products*, American Institute for Research in the Behavioral Sciences, Pittsburgh, 1972.

Dershimer, Richard A. (ed.), "The Educational Research Community: Its Communication and Social Structure," *A.E.R.A.*, April 1970.

Donohue, Joseph C., Jules Mersel, and William A. Morris, *Information Transfer in Educational Research*, and *ERIC and the Future*, Informatics Inc., Sherman Oaks, California, April 8, 1966.

Eidell, T. L. and J. M. Kitchel (eds.), *Knowledge Production and Utilization in Educational Administration*, Center for the Advanced Study of Educational Administration, University of Oregon, Eugene, Oregon, 1968.

Fairweather, G. W., *Methods for Experimental Social Innovation*, John Wiley and Sons, Inc., New York, 1967.

Fleigal, F. C., et al., "Attributes of Innovations as Factors in Diffusion," *American Journal of Sociology*, Vol. 72, No. 3, 1966, pp. 235-248.

Gideonese, H. D., "Research Development and the Improvement of Education," *Science*, 1 November 1968.

Gideonese, H. D., *An Output-Oriented Model of Research and Development and Its Relationship to Educational Improvement*, U. S. Office of Education, Washington, D.C.

Gideonese, H. D., *Research and Development for Education: A Market Model*, National Center for Educational Research and Development, U.S. Office of Education, Washington, D. C., September 1970.

Glaser, E. M., and H. L. Ross, *Increasing the Utilization of Applied Research Results*, Human Interaction Research Institute, Los Angeles, 1971.

Glaser, E. M., and S. Taylor, *Factors Influencing the Success of Applied Research*, Institute of Mental Health, Department of Health, Education, and Welfare, Washington, D. C., final report on Contract No. 43-67-1365, January 1969.

Greenwood, Peter and Daniel Weiler, *Alternative Models for The ERIC Clearinghouse Network*, The Rand Corporation, Santa Monica, California, R-951-HEW, January 1972.

Havelock, R. G., *A Guide to Innovation in Education*, Center for Research on the Utilization of Scientific Knowledge, Institute for Social Research, University of Michigan, Ann Arbor, Michigan, 1970.

Havelock, R. G., *Planning for Innovation Through Dissemination and Utilization of Knowledge*, Institute for Social Research, Center for Research on Utilization of Scientific Knowledge, University of Michigan, Ann Arbor, Michigan, July 1969.

Human Factors Research Laboratory, *Evaluation of the Information Retrieval Center of the Northern Colorado Board of Cooperative Services, Boulder, Colorado*, Colorado State University, Fort Collins, Colorado, June 30, 1971.

Klein, B. (ed.), *Guide to American Educational Directories*, B. Klein Publishing Co., New York, New York, 1969.

Machlup, Fritz, *The Production and Distribution of Knowledge in the United States*, Princeton University Press, Princeton, New Jersey, 1962.

Maguire, L. M., *Observations and Analysis of the Literature on Change*, Research for Better Schools, Inc., June 1970.

National Academy of Sciences, *Scientific and Technical Communication: A Pressing National Problem and Recommendations for its Solution*, a Report by Committee on Scientific and Technical Communication of the National Academy of Sciences, 1969.

National Center for Educational Research and Development, *Educational Research and Development in the United States*, U. S. Office of Education, Department of Health, Education, and Welfare, Washington, D.C., Government Printing Office, 1969.

New York Research Coordinatirg Unit, *ERIC and the New York RCU: Guidelines to the Use of the Educational Resources Information Center Service in New York State*, The University of the State of New York, State Education Department, Albany, New York, January 1971.

Persell, Caroline, *The Quality of Research on Education: An Empirical Study of Researchers and Their Work*, Columbia University, Bureau of Applied Social Research, 1971.

Rittenhouse, Carl H., *Innovation Problems and Information Needs of Education Practitioners*, Stanford Research Institute, Menlo Park, California, May 1970.

Rogers, E. M., and F. Scoemaker, *Communication of Innovations*, The Free Press, New York, 1971.

Rogers, E. M., *Diffusion of Innovations*, The Free Press, New York, 1962.

Rosenau, L. H., and J. Hemphill, *Utilization of NIE Output*, Far West Laboratory for Educational Research and Development, Berkeley, California.

Sieber, Sam D., Karen S. Louis and Loya Metzger, *The Use of Educational Knowledge*, ERIC Documents Reproduction Service, 1972.

Stufflebeam, Daniel L., *Proposal to Design New Patterns for Training Research, Development Demonstration/Dissemination and Evaluation Personnel in Education*, U. S. Department of Health, Education, and Welfare, Washington, D. C., December 1970.

System Development Corporation, *Evaluation Study of NCEC Information Analysis Products: Final Report*, Santa Monica, California, June 1972.

Wanger, Judy, *Directory of Educational Information Resources*, CCM Information Corporation, 1971.

York, L. J., *Arrangements and Training for Effective Use of Educational R&D Information*, Far West Laboratory for Educational Research and Development, Berkeley, California, December 1968.

Comments on Weiler Paper by David L. Clark

In attempting to organize this brief critique of a far-ranging paper, I will begin with a three-fold conceptual challenge to the orientation chosen by the author. This may seem to lead logically to a position which disregards most of what follows in the paper but this is, in fact, not true for two reasons. First, the author does not lean as heavily as the first sections of the paper itself suggest on his own logical structure, and, second, he simply has many sensible and provocative observations to make within his frame of reference. Consequently, I will move, as a second step, to the identification of the material I found most useful; will then proceed to more specific disagreements with the author on points which may be of some interest to the reader; and will conclude with a list of "implications" for N.I.E. which were suggested to me by my review of the paper.

A Conceptual Challenge

In my opinion, Mr. Weiler was seduced by an attractive and popular alternative when he chose to define the problem to be addressed by the paper in terms of difficulties faced by the practitioner in using educational R and D products and information. That particular focus leads to an overemphasis on second level technical problems, e.g., difficulty in locating information and products, redundancy in the system, or lack of quality control in admission of items to the system. These problems though technical are not trivial in a long range sense but if one is defining, as Mr. Weiler is, "the situation at present," they are actually trivial considerations. For example, the concern that "practitioners find it difficult to identify, locate and acquire useful information on R and D products" and that this difficulty is attributable in any significant measure to a lack of "district-level specialists" or "well supported research libraries" distracts attention from more telling problems, e.g., the fact that no development community on education exists (Shutz), or that the structure of the educational bureaucracy is such that it has a predisposition to resist certain types of change (Pincus).

Throughout the paper, in my opinion, there is too much emphasis on the individual practitioner in contrast with "the institution of education."

The central question in dissemination and adoption tactics and strategies in education should focus much less on the individual practitioner, who in this paper seems to assume the status of a private entrepreneur, and much more on the bureaucracy of education, its bureaucratic sub-system, and individuals in their defined gatekeeper roles within that bureaucracy.

This emphasis on the educational practitioner leads to yet another conceptual problem of some importance. Mr. Weiler treats the problem of dissemination as a problem of serving the client, i.e., in his terms the educational practitioner, efficiently and effectively. This strikes me as an a priori goal displacement. The objective in being concerned with dissemination in education is to effect positive modifications in educational practice where the practitioner's satisfaction is incidental (except insofar as it influences the modification) and the client is viewed as the user of the system, i.e., the student. The practitioner although sometimes an active participant in and stimulator of the change process may, in other instances, be a major constraint to needed change. An appropriate view of his role would consider him to be an intervening variable affecting change not the end product to which the change process is directed.

Highlights in the Paper

The single most interesting and useful of the introductory sections in the paper was the sub-section "Resistance to Change." I agree with Mr. Weiler on two counts; first that the assumptions noted have been accepted at least in an implicit sense for operational decisions by the Federal government; and second that they are unwarranted and have, consequently, been misleading in decisions made about dissemination. I would suggest that a provocative exercise which might well be undertaken would be the explicit statement of the counter-assumptions and an effort to design a hypothetical system based on those statements.

I am at least in partial agreement with Mr. Weiler's suggestions under the heading "N.I.E.'s Research and Policy Posture Over The Next Few Years." Surely he has identified an appropriate (in terms of the state of our knowledge) and rational starting point in suggesting the status study of what now exists. I would add to this only the caveat that this effort should be more broadly construed at the federal level than NCEC. I think it is not true, as Mr. Weiler notes, that "Federal programs for the dissemination of educational R and D products are largely centered in the National Center for Educational Communication (NCEC)." Steps 2 and 3 under this heading are ones with which I would agree in a very general way, but as my difficulties with the conceptual rationale for the paper indicated and as I will point out in my next section, there would probably be a number of operational disagreements between Mr. Weiler and me in formulating programs to accomplish these agreed-upon ends.

Oddly enough, I will now suggest to the reader that he turn to Mr. Weiler's "punch line" and take very seriously the three summary suggestions offered to N.I.E. My first reaction to agreeing with Mr. Weiler's recapitulation of his recommendations after disagreeing with his central conceptual orientation was to charge this off to the fact that he was simply proposing a universal, rational solution to a planning problem at a sufficiently high level of abstraction to avoid disagreement. To some extent that may be true. However, beyond that I think he has established a tone of introspection and systematic movement which is important in an area where N.I.E. will be pressed to deliver "something rather than nothing" but where something is likely to turn out to be much worse than nothing by obfuscating the lack of real problem solutions through apparent activity.

Disagreements and Questions

Following seriatim as they appear in the text of the paper are a number of points of disagreement or questions that concerned me about the paper:

• Federal Level Dissemination Activities - As was noted earlier, the definition of the federal government's interest and investment in dissemination of educational R and D products as being "largely centered" in NCEC is far too narrow. Even a cursory examination of such activities as Title III of ESEA, EPDA, a variety of training and re-training programs sponsored in U.S.O.E., the activities of USOE in vocational education, the education of the handicapped, etc., lead quickly to the conclusion that a significant portion of all these activities are directed quite specifically to dissemination objectives. The problem with the narrower definition, of course, is that it overlooks a myriad of dissemination programs and investments which overshadow NCEC quantitatively and, at the very least, affect dramatically how NCEC might modify its investments, policies, and procedures.

• Efficacy of "Best of Current Practice" Reports - This objection is more in the form of a cautionary note than a basic disagreement. Best of current practice syntheses have had a desultory history in education whether the efforts took the form of the printed word, e.g., "What Research Says to the Teacher," or an organizational entity, e.g., school study councils. This is not to say that no dissemination occurred but that precious little significant change which modified practice in major ways was effected. That may be because the efforts were poorly implemented not meeting such criteria as Mr. Weiler would pose, e.g., "generalized ground rules for the implementation of reasonably successful educational programs." A competing hypothesis, however, is that this genre of information being, as Weiler notes, "precisely the kind of information that educational practitioners rely upon most heavily," fits the classification of innovations which Mr. Pincus, in his paper, notes

are among those the schools are already willing to adopt. This is not an unreasonable hypothesis since the substantive input is based upon innovations that some system within the bureaucracy has found congenial. The sharing of best current practice may simply be a method for reinforcing behavior the bureaucracy will find a way of manifesting without government intervention.

• Major Research Questions - There is considerable doubt as to whether the questions noted as major research questions by Mr. Weiler are either "major" or "research." The majority of these questions seem to me to be, in fact, questions of system evaluation - evaluation questions of both a process and product type. The significance of this distinction, which may appear to be only semantic quibbling, is in the strategy for dealing with them. My guess is that N.I.E. need not mount a research program for this purpose but can, instead, build an emphasis on such questions into evaluative designs for operating programs saving time, money, and energy and reserving research support in this area for more basic inquiry. To re-emphasize this latter point, I found missing in this section an identification of basic questions on the change process in education in contrast with the narrower function of dissemination within that process. I would venture the guess that an investment in research on change would have higher payoff than the more delimited emphasis suggested by Mr. Weiler.

There is serious doubt in my mind about the significance of the questions raised. It is admittedly difficult to argue or assess their significance at their present atheoretical level. This is not a very devastating criticism if they are to be taken simply as examples, but if they are pushed beyond this point as Mr. Weiler does when he notes that, "This research program should address at least the following questions---," I am inclined to argue. I would argue, for example, that question #1 is a relatively low priority research area in contrast with an identification of the characteristics of the educational bureaucracy which impede or facilitate change; or an empirical verification of Pincus' classification of innovations likely to be adopted by schools. This preference for research directed to the organization called education rather than the individual practitioner is a logical extension of my earlier noted dissatisfaction with Mr. Weiler's conceptual orientation to the topic. It is re-introduced at this point only to temper the reader's enthusiasm for Mr. Weiler's interesting set of questions as a possible conceptual map for an N.I.E. research program in this area.

• The "Experimental Region" - This, I think, is a bad idea. Earlier in the paper Mr. Weiler has built a good case for the need for substantial descriptive research in the arena of dissemination. I concur. The initiation of an experimental region at this time would be premature and the experiments encompassed in the region's program would be crude and wasteful. The hypotheses cited are not really hypotheses at all, i.e., they are not derived from a theoretical or even a logical structure.

The idea looks very much like a raw empirical fishing expedition. The low payoff in terms of new knowledge, in my opinion, would be complemented by very high cost. Most of the money devoted to the enterprise would go not for research activities but for operational costs. Rather than establishing such a region, I would attempt a field based research program in natural settings supplemented, as needed, by targeted interventions and experiments.

In Summary

Although this paper has concentrated on disagreements with Mr. Weiler, it must be obvious to the reader that I found Mr. Weiler's suggestions provocative, helpful, and interesting. From N.I.E.'s point of view in planning for this area I would be concerned with:

1. Concentration on the ecology of the change process in education with dissemination viewed as a part of that process, rather than an emphasis on the technical problems of an immediate sort associated, for example, with the operation of NCEC.

2. Coordination of the total investment of the Federal government in the educational change process if not in an operational sense at least in a conceptual sense.

3. Focus on the institution of education as an institution rather than the individual in an entreprenurial role.

4. Initiation and maintenance of a rational planning attack on a difficult area rather than attempting to achieve fast results which simply will not be forthcoming in this arena.

5. Utilization of multiple tactics, e.g., evaluation designs in operational programs, rather than a single overall approach, e.g., an experimental region or a master research plan.

INCENTIVES FOR INNOVATION IN THE PUBLIC SCHOOLS

John Pincus

January 1973

P-4946

INCENTIVES FOR INNOVATION IN THE PUBLIC SCHOOLS

John Pincus*

The Rand Corporation, Santa Monica, California

I. INTRODUCTION: THE INSTITUTIONAL SETTING

This paper sets out some propositions about the structure and incentive systems of public schools as they relate (1) to the *adoption* of innovations and (2) to their *implementation* in the schools. These propositions may have certain sytematic implications for education R&D policy as well as for such broader questions as how to implement planned change in bureaucracies. The propositions are not based on careful testing of hypotheses, but on a blend of evidence and speculation and are aimed at influencing how we might think about educational R&D policy. Furthermore, implicit herein is the notion that society will be better off if schools could offer a more diverse menu of alternatives in respect to both organization of schools and curricular emphasis. Section II discusses incentives to adopt innovation. Section III discusses the problems of implementation. Section IV suggests some implications for R&D policy. The general thesis of this paper is that the *market structure* of the public school "industry" has a major effect on schools' decisions to adopt innovations; while the *bureaucratic structure and incentives* of schools shape in specific ways the transition from adopting innovations to implementing them. This distinction is somewhat artificial. The ultimate objective should be a testable theory which integrates the incentive effects of both market structure and bureaucratic structure. This paper is therefore a halfway house toward that goal, and not a fully specified model of how school systems behave in response to opportunities for change.

*I have benefited in writing this paper from discussions with George R. Hall. Several people, including David K. Cohen, T. K. Glennan, Jr., Gordon Hoke, Herbert Kiesling, Robert Klitgaard, Milbrey McLaughlin, David Mundel, and Daniel Weiler offered useful suggestions and comments on an earlier draft. None of them bears responsibility for any shortcomings herein.

Public elementary and secondary school systems in the United States are, like many governmental functions, a form of public utility. The public schools are given by state government action a virtual local monopoly of schooling services. The monopoly is not complete because there are four alternatives open to parents who do not choose to send their children to the local public schools: (1) private schools; (2) another public school in the district; (3) another public school in another district; (4) religious schools. The first and third options, private schools and moving to a "better" school district, are open primarily to the wealthy or to those who are both upwardly mobile and also attach primary importance to schools as a determinant of residential location. The second option, another school in the same district, is generally limited by such factors as transportation constraints, school capacity limits, school district regulations limiting transfer, and the relative homogeneousness of neighboring schools. The fourth option, parochial schools, is open primarily to Catholics, and is the most important single alternative to public schools. However, with the progressive reduction in numbers of Catholic parish schools in recent years, most Catholics' alternatives are also being narrowed, a trend which could be reversed only by significant and unlikely changes in church policies toward racial integration or by government subsidy.

For the great majority of clients, the public schools are a de facto local monopoly, which is different from many other local public utilities in several respects. First, unlike most telephone, gas, and electric service, and so on, the quality of school service can vary substantially within a district, which often creates serious perceived issues of equity along income, race, and neighborhood lines.*

Second, in contradistinction to many other local public utilities, the aims of schooling are unclear, or at least there is no consensus about what priority should be given to the various aims.

---

*Similar perceptions of unequal service *between* districts have led to recent court decisions which hold that it is a violation of the Fourteenth Amendment to base school spending differences on differences in the local property tax base. Title I "comparability" requirements are designed to offset some of the intradistrict variations.

Third, the technology of schooling is unclear, unlike that of most public utilities. In economists' language, we don't know what the educational production function is, or even if there is one, except perhaps in some non-operational sense.* More generally, we are often unsure whether one method of providing school services is consistently better in terms of output effects, however defined, than any other method.

Fourth, school districts may have very little incentive to be economically "competitive"--to increase their registration at the expense of other districts. The perceived financial gain or loss from interdistrict shifts in public school registration depends on several factors, including the ratio of state and federal aid to local taxes, effects on the local tax base, possibilities of adding or dropping staff or facilities, socio-economic characteristics of shifting students, and so on.

Fifth, although the schools are educational institutions, they apparently provide only a small part of the student's educational resources. Other influences--heredity, family and peer-group environment, communications media, etc.--appear to be the prime determinants of how much people learn, how they learn, and how they respond to contacts with other people and social institutions. This situation makes it very difficult to gauge the effects of schooling on people's lives and learning.

The public schools, of course, do share a number of common characteristics with other non-market oriented public utilities. They are self-perpetuating bureaucracies, thanks to tax-supported status, certification practices for teachers and administrators, and custom of promotion from within. In these respects, the schools resemble many civil service agencies, notably public health, welfare, and criminal justice systems. They also share with these systems a

---

*It is non-operational because we now have no satisfactory way of measuring many of the multiple outcomes of schooling, nor of adjusting for differences in teacher and student quality, nor for taking account of the interaction among students, teachers, and curricula, which introduces systematic bias into empirical estimates of educational production functions.

characteristic which profoundly affects their institutional response to innovation: they cannot select their clients and the client must, as a practical matter, accept the service, whether or not he is satisfied with its quality [Carlson, 1965b].

Like the systems cited above, school districts operate under a highly decentralized system of governance, but a highly complex structure of influences. There are nearly 18,000 school districts in the United States, each subject to a range of local community influences, as well as to the influence and legal authority of state and federal governments.

Finally, like certain other self-perpetuating bureaucracies (police, fire, public health services), the schools are a labor-intensive craft industry whose managers often present to the outside world the impression that their craft is highly specialized, that its functions cannot be carried out by replacements whether in the form of uncertified labor or machines.*

*The recent widespread introduction of teacher aides may represent some potential competition for the teacher guild. This threat is presumably offset in teachers' minds by the value of having assistance in routine and menial chores. It is predictable that teacher aides will themselves "professionalize" before long.

II. CONSEQUENCES FOR DISSEMINATION OF INNOVATIONS IN THE PUBLIC SCHOOLS

How would we expect a self-perpetuating bureaucracy to respond to R&D findings if (1) it is not market-oriented; (2) is widely considered to be socially necessary and therefore deserving of public protection--is in fact the captive servant of a captive clientele; (3) is open to a good deal of public scrutiny on issues having to do with perceived equity, quality, and goals; (4) cannot unambiguously define its aims or clearly identify technologies that are dominant in light of aims that might be specified; (5) its contribution to its clientele's life and learning is uncertain and also modest as compared to other societal influences; (6) its governance is highly decentralized, yet subject to a wide variety of influences, so that each unit perceives itself as facing an unique configuration of clients and masters.

An obvious response is that organizations facing these influences might have fewer incentives to innovate than in situations where market forces or the clarity of institutional goals dictate invention or the adoption of innovations. Yet, as has often been pointed out [Rogers and Scoemaker, 1971; Carlson, 1965b; Miles, 1964; Havelock, 1969], the schools have tried out and adopted a large number of innovations. Certain innovations (the new mathematics, PSSC curriculum, language laboratories) have spread very rapidly; others (junior high school, kindergarten, driver training) more slowly; still others (ungraded classes, open schools, decentralization of decisionmaking from district level to school level) very slowly. Then there are some educational innovations (voucher systems, abolition of teacher tenure, abolition of formal schooling, parent evaluation of school staff as a basis for retention and promotion) which have not yet spread at all. Finally, there have been many innovations that have been adopted but are often not successfully implemented. (A great many applications of new technologies, such as audiovisual equipment and CAI, appear to fall into this category, as well as, in all probability, such new management techniques as PPBS, accountability, administrative decentralization of large districts, etc.) The impedimenta of these innovations--in the form of equipment, or a new set of management structures, or the vestiges of "bold, new" curricula--remain beached by

the wake of ephemeral educational revolutions, while the system continues to operate as before.

The responses of schools to opportunities for innovation appear therefore to be complex; and between the adoption and the implementation, innovations routinely disappear or suffer sea-changes [Gross et al., 1971; Havelock, 1969; Goodlad, in Committee for Economic Development, 1969].

The Schools' Response to Innovation: A Market Analogy

In order to understand under what circumstances schools will or will not be likely to adopt and implement innovations, it is instructive to compare the responses one might expect from the public schools with those one might expect from a competitive private firm (say a private vocational school, such as a computer programming school or a secretarial school).

Educational innovations can affect the schools' operations in diverse ways: increasing the level of resource use only ("more of the same"--e.g., a smaller class size); changing the resource mix (a higher proportion of teacher aides, relative increase in capital equipment); changing instructional processes or methods without significantly changing resource level or mix (new math, new reading curriculum); affecting administrative management, without significant effects on organizational power structures (computerizing data management, new accounting systems); changing either the organizational structure of the schools or their relation to external authority (community control, open schools, voucher systems).

Compared to a competitive firm, we would expect the public schools to:

1. Be *more likely than the competitive firm to adopt cost-raising innovations,* since there is no marketplace to test the value of the innovation (e.g., smaller class size) in relation to its cost. Therefore, any cost-raising innovation that is congenial to the public school authorities and acceptable to local taxpayers or state and federal funding sources will be adopted.

2. Be *less likely than the competitive firm to adopt cost-reducing innovations*, unless the funds so saved become available for other purposes within the district.

3. Be *less likely than the competitive firm to adopt innovations that significantly change the resource mix* (e.g., a higher ratio of teacher aides to teachers, sharply increased use of capital-intensive technologies), because any consequent productivity increases are not necessarily matched by greater "profits" to the district, and because any replacement of labor by capital may threaten the guild structure of the schools.

4. Be *more likely than the competitive firm to adopt new instructional processes or new wrinkles in administrative management* that do not significantly change institutional structure, because such innovations help to satisfy the demands of the public, of state and federal governments, and of teachers and principals themselves for change and progress without exacting heavy costs to the district in the form of organizational stress.

5. Be *less likely than the competitive firm to adopt innovations that change the accustomed authority roles* and established ways of doing business, because changes in these relations represent the heaviest kind of real cost to bureaucracies.

6. Be *equally unwilling as competitive firms to face large-scale encroachments* on protected markets (voucher systems, metropolitan-areawide open enrollment), although for somewhat different reasons.

From this perspective, the public schools can be seen as more likely than private firms to adopt innovations that do not require complex changes in management structure or organizational relations. Such innovations help to satisfy staff and client demands for change, without requiring from the organization the difficult task of self-renewal, which many of the organization's clients, as well as the

organization itself, might resist. Such innovations are also safe, in that it is nearly impossible, given the present state of educational information systems, to document whether a new curriculum, or new physical plant, or an audiovisual system helps or hurts children's learning or attitudes. Therefore, the innovating district, if it uses reasonable sense, is unlikely to get in trouble as a consequence of adopting or abandoning such innovations. Private vocational schools, whose policies are closely tied to student success in job placement, are quite conservative about adopting the latest thing in curriculum, because the risks are excessive in the absence of evidence [Belitsky, 1969].

Although there are probably significant differences in the kinds of innovations that schools and competitive firms are likely to adopt, it is impossible to generalize about whether public schools will adopt more or fewer innovations than competitive firms. It is often pointed out [e.g., Mansfield, 1963] that competitive industries characterized by relatively small firms (e.g., farming,* apparel, hardware) are likely to innovate less than large firms in less competitive industries (e.g., pharmaceuticals, electric equipment, computers), but in this context that is somewhat besides the point, as is the presumably correct argument that no firm, public or private, likes to make uncongenial changes. The point is that differences in market structure tend to lead to different patterns of innovation, through differences in the nature of incentives, whether positive (higher profits, larger Federal grants) or negative (impending bankruptcy, threatened teacher strikes).

Bureaucratic Factors Supporting Innovation

These market considerations are quite general, of course. What are some of the more specific bureaucratic conditioning factors that

*Farming is actually an innovative industry in the United States, but, as in the case of education, the research and development bill is almost entirely paid by the Federal government and oligopolistic farm supply industries. Competitive industries that are less powerful politically than agriculture receive little or no subsidized research and development, while individual firms are too small to finance ally significant levels of R&D.

lead school districts to adopt innovations? For years a dominant view was that the primary determinant of willingness to innovate was the level of per capita school spending [Mort and Cornell, 1941; Carlson, 1965b]. This view was based on extensive studies of school district administration conducted by the Institute of Administrative Research at Columbia Teachers College. These findings, which were widely disseminated, buttressed the school superintendent's natural desire to maximize his per-pupil budget, providing thereby a happy coincidence of organizational self-interest and socially endorsed "progressive" behavior.

More recent research [Carlson, 1965a; Havelock, 1969; Mansfield, 1963; Gross et al., 1971] casts doubt on this contention with respect to schools as well as industrial firms, without denying the usefulness of command over resources. A more complex view of the determinants of innovation in the school emerges. Three factors seem favorable to innovation in the schools:

1. *Bureaucratic Safety*--When the innovation is perceived as favorable with respect to the current status and organization of the bureaucracy (because in a self-perpetuating non-market system, these bureaucratic values become socialized and tend to dominate other criteria; or in other words, the bureaucratic costs are the real costs of the system).

2. *Response to External Pressure*--When external pressures for innovation are perceived as irresistible (because school systems cannot be entirely unresponsive to external pressures and financial constraints).

3. *Approval of Peer Elites*--When key figures in the bureaucracy and their colleagues in other educational bureaucracies can reach consensual agreement about the acceptability of the innovation (because in the absence of clearly defined output criteria, consensus among the elite is often the primary decisionmaking criterion).

These elements are interrelated. For example, external pressures can lower the school district's perception of bureaucratic safety, thereby providing negative incentives to innovate; or if a particular innovation is neutral with respect to bureaucratic safety, then peer approval may act as a positive incentive. They are also complex. For example, approval of peer elite can be used by individual administrators as a justification for pursuing deeply held beliefs, while it can be ignored when it endorses policies that the same administrators are opposed to. Finally, they are relative. In each organization, depending on the circumstances, a constraint may be more or less elastic, and one object of R&D policy may be to make these constraints more elastic, thereby creating greater willingness to change (see pp. 14-15 below).

In addition to these factors which apply particularly to the public school setting, there may be elements present in any organization, whether or not educational, that encourage innovation. These have been discussed widely in the literature on innovation [Bennis et al., 1969; Lippitt et al., 1966; Havelock, 1969; Lippitt, Watson, and Westley, 1958; Marcum, 1968; Rogers, 1971]. These elements, although clearly important in many instances, will not be discussed in detail here. The kinds of factors that students of planned change have identified as generally supporting innovation in organizations include, after the outline of Glaser (1971): organizational attitudes that support change (such as free communication, support from administration and colleagues, high staff morale); clarity of goal structures; organizational structures that favor innovation (sufficient decentralization of authority, existence of a large number of occupational specializations, existence of structures for self-renewal); professionalism of staff; organizational autonomy (not excessively dependent on public opinion or tests of political feasibility to validate planned change); and few strong vested interests in preserving status quo methods of operation.

Some of these elements (e.g., lack of clear goal structures or organizational autonomy) are implicit in the three factors described above (Bureaucratic Safety, Response to External Pressure, Approval of

Peer Elite). Those that are not implicit obviously can affect the propensity to adopt innovations in schools, as elsewhere, and we would expect different school systems to exhibit these qualities in varying degrees.

But if we accept the proposition that the unique elements in the schools' response to opportunities for innovation stem from their special institutional role, market structure, and the systematic set of economic and bureaucratic incentives so created, then there emerges a somewhat different perspective from that usually set forth in the literature on innovation. The three factors described above can, in this perspective, be considered as reflections of institutions, markets, and consequent behavioral incentives facing the public schools.

Therefore, if we can identify the kinds of innovations that are likely to be adopted by school districts that follow such behavioral styles, we may be able to identify ways that R&D products can be oriented in order to gain acceptance. As soon becomes apparent, the three conditions favorable to adoption of innovations in the present setting are themselves rather restrictive. Advocates of substantial innovation in the public schools aren't likely to be very satisfied by a R&D dissemination strategy which takes these conditions as operative constraints. Therefore, it is useful to examine the ways in which R&D dissemination policy could take advantage of the existing structure of market and bureaucratic incentives and also to examine the ways in which these incentives could be modified by conscious R&D policy.

The *bureaucratic safety constraint* means that schools are unlikely to accept radical changes in educational institutions, such as taking instruction out of the classroom, introducing capital-intensive instructional technologies, or new forms of educational market organizations, because such changes might be expected to affect the organization of the system substantially.

The fear of *external social and political pressures on the school system* means that schools will be reluctant to enter into genuine collaboration with other social groups at the policymaking level, such

as community or student participation in control of school district policy, or providing the public with educational information systems that could be used as a step towards "accountability." Extra-system knowledge of school affairs is perceived as leading to greater extra-system pressures for reform, thereby creating unwanted problems for the school system.

The *elite consensus constraint* tends to prevent any but marginal changes from current practice. School districts are of necessity unclear about educational goals, and educational research and development has failed to enlighten them substantially about the relationship between various educational technologies and any specified instructional aim. Therefore, faced by such enormous uncertainties, a rational bureaucratic elite would be unlikely to experiment voluntarily with major changes in structure or method. Social and political consequences would be incalculable (e.g., busing, sex education) while benefits would necessarily be uncertain.

Given these constraints, and the market structure of the public school "industry," schools tend voluntarily to adopt innovations which promote the schools' self-image by demonstrating that the schools are:

- *"Up-to-date"*--introducing modern physical plant, new curricula not requiring changes in bureaucratic organization or staff rules, reduction of class size, use of teacher aides, team teaching.
- *"Efficient"*--adoption of electronic data processing, new budgeting and accounting systems, portable classrooms.
- *"Professional"*--adoption of curricula that are espoused by the educational leadership, hiring well-trained teachers, subsidizing in-service training and workshops, consulting with faculty of leading schools of education.
- *"Responsive"*--establishing formalized links to parents, using blue-ribbon advisory committees to submit reports on policy issues, establishing counseling and guidance functions, establishing special programs for handicapped, gifted, slow learners, etc., providing vocational programs

that respond to needs of local industry, offering a variety of adult extension courses.

Because the 18,000 school districts have a great deal of autonomy in deciding whether and how to innovate, we would expect adoption of innovations often to be a selective and idiosyncratic process varying according to administrators' tastes and their perceptions of school and community needs. The empirical evidence shows that small districts adopt fewer innovations than large ones [Lindeman et al., 1968], presumably because large districts are more able to keep informed of new methods, and face a wider variety of both external and system-generated pressures for change.

Those innovations that are widely adopted generally share common characteristics of substantial consensus in their favor among the elite and presenting no major bureaucratic or social problems. The most widely adopted instructional innovations, as of 1969, (adopted by half or more of the largest school districts) were: teacher aides, ability groupings, team teaching, elementary resource teachers, movable partitions, TV instruction, and non-graded sequencing. Curriculum innovations were widely introduced by large districts over the period 1965-69 in science, math, and reading [Lindeman et al., 1968]. The curriculum innovations were influenced by the NSF science and math curricula and by the sales efforts of new commercially marketed curricula (e.g., the EDL reading labs and SRA reading program).

These findings indicate that large-scale carefully planned R&D efforts are, in curriculum change, likely to be more effective in gaining adoption than more modest efforts (the current success of the SWRL prereading program is another case in point). It should also be noted that the NSF and SWRL prgorams were worked out in close collaboration with practitioners, which helped encourage adoption. Finally, these programs were widely publicized and praised by professional education groups, so that there were social pressures for adoption.

This last characteristic has significant general implications for acceptance of R&D products. Since the incentives for a school superintendent or principal to adopt or reject an innovation are so diffuse and so closely related to administrators' preferences, and their perceptions of internal and external constraints, R&D organizations

should clearly do their best to work closely with school administrators and those who influence them (county and state school officials). In part, this is a question of co-optation. But considering the characteristic remoteness from the client of educational R&D organizations, which have traditionally been university-centered, it also can serve as a form of reality therapy for the researcher. Of course, the opposite problem also can arise, as witness the rather poor record in innovation of school districts' internal research staffs, probably because they share too closely the bureaucracy's perspectives and priorities.

Adoption can also be catalyzed by pressure, subsidy, or other incentives from external jurisdictions or interest groups (e.g., federally mandated or subsidized innovations, such as Head Start or Follow Through; state mandating of kindergarten programs or programs for education of the handicapped; court decisions on desegregation or finance; influence of industry or interest groups on obtaining special programs, such as vocational education or driver education; minority community influences in achieving black or Chicano study programs, or varying degrees of decentralization).

School districts face a certain set of incentives which systematically affect their preferences for different kinds of innovations. State and federal R&D policy, to the extent that it aims at encouraging innovations that schools would normally be reluctant to adopt, should devote most of its funds to innovations that are uncongenial to the schools, with payments based on evidence that reforms actually are undertaken. If reformers seek to affect the ways that districts respond to internal and external institutional pressures, they will have to pay for it. Therefore, most federal and state subsidy, both for R&D and for innovation (both directly to schools and indirectly to R&D agencies) should go not for things that schools want to do, but rather for things that they would otherwise be reluctant to do.

Large-scale, well-planned support for innovative efforts aimed in part at rectifying the existing institutional biases, and a conscious policy of collaboration with school administrators (and

increasingly with leaders of teachers' organizations) are therefore two potential catalysts for adopting policies which rank high in reformers' preferences--the first aims at reducing existing barriers to innovation, while the second aims at achieving more innovation within the existing constraints.

A school district, whatever its critics may aver, is a going concern, one whose "survival" is under existing laws, threatened only to the extent that school boards can replace superintendents and that the public can replace school boards. Other employees are relatively invulnerable to these possibilities. Therefore, given the risk-avoiding mentality that we might normally expect in such a bureaucracy, real costs of innovations that affect internal or external relations of the system are likely to be magnified. At the same time, gains from such innovations are likely to be discounted, because institutions' operational focus is necessarily short-range which tends to stress immediate costs to the system, while the benefits of such changes are typically uncertain and remote. This creates a built-in conflict between practitioners and those researchers who seek innovation through methods that require reform of structures (which incidentally provides further argument in support of external subsidy for uncongenial innovations, such as vouchers, ungraded schools, or alternative schools). The research community typically complains that practitioners and politicians are unrealistic in their desires for immediate results. One method that researchers can use to appear responsive to this desire is to promise more performance from an innovation than the evidence warrants. This response, which is the more understandable if the R&D organization stands to benefit in prestige or money from the adoption, tends to heighten the district bureaucracy's skepticism about the merits of any R&D initiative which engenders significant organizational stress.

Because so many factors, not the least of which are the uncertainty of benefits and the certainty of resistance, tend to operate against any substantial voluntary change in the structure of the schools, desires for progress and reform therefore tend to be channeled into "safer" areas--those that involve spending more money on the existing

resource mix (more teachers, more administrators), or those that involve the kinds of changes in curriculum or administration that don't seem to threaten organized groups in or out of the bureaucracy. This is a collateral reason for the oft-noted prevalence of faddism in school reform. If structural changes are prohibitively costly in real (institutional) terms, then the attractiveness of less costly reforms, or even of chasing after will-of-the-wisp, is heightened.

III. FROM INNOVATION TO IMPLEMENTATION

Frequently change may be made still less costly in terms of the system's values by not implementing innovations along the lines proposed by their developers. Perhaps the most common complaint of the R&D community about adoptions is that the innovations are not actually implemented as prescribed, so that they never get a fair trial. This has clearly been the fate of most audiovisual developments, for example. Goodlad (1970) has pointed out that despite years of discussion and professional support leading to widespread adoption of such innovations as ungraded classes and team teaching, these innovations are rarely implemented. A school district will adopt ungraded classes, then implement it by teaching essentially a graded curriculum in the "ungraded" class.

These are several reasons for this failure to implement innovations effectively:

- R&D organizations frequently do not provide sufficient implementation guidance, in light of the variety of school situations where adoption is tried.
- Teachers, administrators, and students may not accept the obligation to change their behavior patterns in ways consistent with implementation.
- The schools may simply not know how to implement the innovation. As Smith and Keith (1971) have said, describing one such effort:

  > In spite of prodigious effort, common guidelines that guided did not exist; the language of school organization, teaching and goals for pupils remains metaphorical and literary but neither practical nor scientific (pp. 52-53).

- As a sort of corollary, if the language of the schools is "neither practical nor scientific" but metaphorical and literary, it may often be the case that school personnel will be more interested in the language of innovation than in the complexities of translating that language into innovative practice. This style of operations

referred to as the "alternative of grandeur" [Smith and Keith] may well be entirely reasonable in the absence of evidence that conclusively supports the advantages of innovation. For the schools' purposes, verbal adoption of innovations may be entirely sufficient, and a preference for the verbiage of magnificent vistas has been noted by various observers.

The problems associated with implementing major innovations in public schools are too substantial to discuss in detail here. Suffice to say that even when motivations to implement are strong, innovations that are perceived as radical by the schools and their clients are exceedingly difficult to implement [Smith and Keith, 1971; Gross et al., 1971]. Frequently cited barriers in cases where there is widespread support for the innovation include unclear objectives or methods, and too little time allotted for planning change and informing people of what is planned and why. In instances where there is not widespread support, then such factors may come in to play as the need for stability, personal or institutional perceptions of threat or vulnerability, inertia, perceptions of client response, etc.

The Institutional Setting for Innovation

The principal funding sources that support innovations in the schools are federal and state governments, either directly through grants to school districts or indirectly through funding educational research and development. Cases in point are Head Start, Titles I and III of ESEA, the Emergency School Aid Program, Career Education funds, Miller-Unruh reading programs in California, urban aid in New York State, federal support of regional labs and R&D centers, and so on.

In practice the only real control that the Federal government has over district use of funds is the relatively unlikely option of withdrawing support. As the history of efforts to obtain Title I "comparability" indicates, use of this weapon is largely symbolic, as an adjunct to moral suasion [Wirt and Kirst, 1972].

Local school authorities know that once they receive a grant, they have much more freedom to use it in accord with their own priorities than the granter might wish.

School districts are characteristically hard-pressed, squeezed between voter reluctance to raise property taxes and employee wage demands. This squeeze tends to buttress whatever preference the school authorities have for system maintenance over innovation, and the actual flow of funds is likely to reflect those preferences [Coleman, 1972].

Some attributes of federal aid enhance these tendencies, and act to discourage incorporation of innovations into school systems.

- There is a tendency to subsidize educational research and development without particular reference to the effects of the developments on various outcomes of schooling.
- There is a tendency to ignore in setting policy the evidence of evaluation reports on innovative programs, allowing districts to introduce or perpetuate pet projects without regard to the alleged aims of innovation. (This does not imply that all evaluations are worth heeding, a fact which buttresses the policymaker's natural tendency to support whichever innovations his personal or bureaucratic preferences may dictate.)
- There are too frequent changes in program priorities and too short a life for educational experiments. Many federally funded innovative programs are based on trials of one to three years, with two major consequent disadvantages: (1) not enough time is allowed to separate effects of the innovation from effects of the frictions arising from the effort to implement; (2) because the districts know that the programs cannot get a fair test in such a short time, they are unlikely to take much interest in the programs as exemplars.
- A related difficulty is the tendency of federal and state agencies to view their contributions as seed money to be replaced by district funds if the program is a success. But school districts know that the typical cost of such programs ($100—$500 or more per student per year) is beyond their ability to finance for the student body at

large, while using district funds for applying the innovation to only a small number of students raises serious ethical questions for a regulated public utility.

- The school districts do not perceive the federal government as demonstrating clear or consistent policies toward innovation. There is no clear long-term benefit or penalty to a district if it adopts or fails to adopt one set of of innovations in preference to another. This tends to reduce the schools' respect for federal policies toward innovation, and to breed a certain cynicism as to the merits of serious efforts at innovation. Furthermore, since federal aid fails to systematically support hard alternatives and scamp easy ones, it in effect encourages a strategy of "grantsmanship," as witness a favorite saying among school administrators in response to federal agencies' description of new programs--"Yes, yes, just give us the money."

- The schools interpret these peculiarities of federal aid policies as meaning that federal aid is unreliable--"soft money" that will disappear as suddenly as it arrived. Therefore, school districts characteristically refuse to use federal money as the basis for any substantial long-run changes in ways of doing business.

- Furthermore, the federal government's support of innovation is relatively small scale compared to other programs such as impact aid and compensatory education. Therefore, funds for innovation, while helpful to a school district on the hunt for federal largesse, are a second order quest. This is rather ambiguous, though. If federal support of innovations were larger than it is, the institutional pressures to call almost anything by the name of innovation would be irresistible. Apparently, under the existing set of institutional relations, no federal investment in innovation is optimal--low levels of funding are insufficient to call forth substantial efforts of organizational reform, while

large-scale funding would clearly lead down the path already blazed by impact aid and Title I--compensating the schools for following the path of least resistance.

IV. IMPLICATIONS FOR EDUCATIONAL R&D

The schools are a unique social institution, molding the clients who, in ways reflecting reverence and resentment, also control it. From a certain point of view, the schools are primarily the agents of socialization, teaching successive generations how to accommodate to social and economic institutions, in the interests of the existing social order [Bowles and Gintis, 1972; Dreeben, 1968]. To accept such a view is, in effect, to deny the possibility of significant innovations in schooling, except as dictated by changing interests of ruling social and economic forces. But this view is far more optimistic about the merits and possibilities of innovation than some current empirically-based analyses [Averch et al., 1972; Jencks et al., 1972]. It posits the inevitability of effective innovation under the influences of changing social regimes, while Averch and Jencks seem to cast strong doubts about the ability of the schools to affect learning and attitudes in any systematic and significant way.

If we are to belive Bowles and Gintis, the only way to change the schools is to reform or revolutionize society, because the schools today are in effect performing the role that the majority influences in society want them to, and not the way that reformers want them to. If we are to draw reasonable inferences from Jencks and Averch, it makes more sense to invest in innovations in the non-school environment than in the schools themselves, because environmental factors account for far more of the variation in achievement tests, college attendance rates, lifetime earnings, etc., than school factors do or can.

If these researchers are correct in their conclusions, much of the money spent on educational innovation, however carefully allocated it may be, is wasted by social policy criteria. Whether or not they are right is debatable. After all society does not choose to systematically affect heredity; the outcomes of its attempts to influence the broad social environemnt are characteristically far from the mark; in practice this leaves the educational system as the principal vehicle for policy reform. The fact that the schools are not omnipotent in shaping educational outcomes is partly irrelevant. In current

circumstances, the market structure of the schools, the uncertainty about their goals and technology, and the particular set of institutional incentives that school districts face lead to systematic preference for certain kinds of innovations over others. This paper aims at suggesting ways for the schools to become more open to a variety of innovations, particularly those that the schools have not yet adopted. Measurement of the consequences can appropriately come later. As long as the schools can become more open to certain kinds of innovations, their prospects for performing better will be enhanced in the long run.

The discussion of Sections I and II above leads to an important conclusion for R&D policy. *If goals are in some sense undefinable, it is inappropriate to adopt the standard rationalist approach of first defining goals, then seeking means appropriate to achieve them efficiently. Instead, R&D strategy should be based at least in part on the converse approach.* If the present situation is unsatisfactory, then it is wiser to try out systematic innovations and assess their consequences than to continue to pursue uncertain goals with unclear technologies. (For a similar view, see March, 1972.) Adopting this strategy means finding ways to do three things:

- Adopt R&D policies that appeal more effectively to the existing set of bureaucratic incentives; and also policies that attempt to modify those incentives.
- Adopt policies that permit the public, acting through a variety of institutional means, to assess more accurately what the schools are doing and how well they are performing.
- Introduce changes in the market structure of the schools that will facilitate innovation.

This suggests five broad emphases for R&D policy in encouraging adoption of innovation: (1) large-scale experimentation; (2) collaboration between R&D agencies and educational leadership networks; (3) case studies of successful and unsuccessful innovation; (4) research that will improve the R&D community's understanding of the existing pattern of incentives in the schools; (5) trying out methods

of restructuring system incentives. Most of these approaches have been tried to greater or lesser degrees. The following discussion attempts to link them to the discussion of incentives.

1. Large-Scale Experimentation

Most educational innovations are tried out on a small scale in one school or one district. They tend, whether considered successful or not, to disappear from view. The National Center for Educational Communications, through ERIC and other devices, tries to disseminate information about innovations, but the results to date in terms of adoptions so generated have not been impressive. Large-scale experiments, either planned or emerging as offshoots of other programs, include Head Start, Follow Through, Titles I, III, VII, and VIII of ESEA, the NIE experimental schools program, Higher Horizons, More Effective Schools, Sesame Street, The Electric Company, NIE career education models, and performance contracting. There has been wide variation in the perceived success of these efforts, both between and within programs [Averch et al., 1972]. But the experiments have in general not been designed or evaluated in ways that would allow anyone to assess the reasons for their success and failure in the real-life setting of the schools. This kind of assessment is difficult, both because education is a complex phenomenon and because innovations that impinge on bureaucratic values make headway slowly. There is not only the obvious point--experimenting with a major educational innovation for one year or a few years is unlikely to reveal much about its merits even in its own terms--but also a less obvious and far more general one: any substantial intervention in an existing social system is very likely to have important unintended effects, reflecting the system's effort to respond and accommodate to the new stimulus. For example, one of the unintended effects of New Deal agricultural price support programs was to subsidize large commercial farmers heavily without significantly halting the decline of family farming. This effect reflected both changes in agricultural technology and the strong influence of commercial farmers in the structure of agricultural p[illegible]s, which in turn was able to exert its influence on the broader

structure of national politics. By the time these unintended effects became apparent, it was too late to rectify them--had experimentation been possible, the eventual outcome might have been avoided through different policies.

In education, suppose that an unintended effect of ESEA Title III were that Title III schools or districts behaved no differently than others three years after federal support expired. This result would provide strong evidence that system behavior is extremely stable with respect to perturbations introduced by temporary funding in support of innovation. This might in turn argue either for longer term support of effective innovations or for abandonment of the present Title III program.

As noted above, some experimentation has already been undertaken and offers a substantial opportunity for seeing how R&D initiatives have actually affected the schools as institutions, offering thereby guidelines for future R&D policy. But two kinds of new, large-scale natural or planned experiments are also needed. The first kind of experiment involves finding out more about the effects of new methods on educational outcomes, given the current institutional structure. Examples include: (1) long-term analysis of cohorts that, through chance or design, receive different educational treatments (the Progressive Education Association's Eight-Year Study, 1934-1941, is the only extant example of such an approach); (2) highly capital-intensive forms of education; (3) curricula that make sharp changes in existing pupil-teacher, pupil-method and teacher-method interactions.

The second kind of experiment is more deliberately aimed at modifying the current structure of institutional or market incentives. Examples, discussed in more detail below, include: (1) educational vouchers; (2) youth endowment plans; (3) alternative schools within a district; (4) decentralized governance; (5) merit pay; (6) compensating R&D agencies and school personnel for both the development and the implementation of innovations.

But all such research and experimentation should focus not simply on the effectiveness of meeting stated goals, but also on the systemwide

effects of the experiment, in particular the institutional response of the schools to the new stimulus. This approach will help create a corpus of knowledge about the response mechanisms of schools to innovation in different fields, as advanced in different ways--in particular it will show which innovations, if any, are most effective under current incentives, and which ones effectively modify those incentives.

Large-scale, appropriately publicized experiments are important to demonstrate to schools and the public that a particular innovation can succeed in a variety of settings. They are also important in some cases to provide convincing demonstrations of failure. Educational research and development organizations espouse a wide variety of innovations. Some large-scale experiments, even if carefully chosen, are likely, after a resonable test, to fail of their objectives. It is appropriate that knowledge of unpromising innovations be as widely disseminated as promising ones. For example, if the one-year OEO performance contracting demonstration had been continued for long enough to demonstrate that contractors could, given a reasonable time span, neither improve cognitive skills nor encourage schools to adopt new methods faster or more cheaply than otherwise, the conclusion would have been of value for policymaking and well worth disseminating widely.

2. Collaboration with Educational Leadership

There is strong evidence that school district administrators rely primarily for research and development information on personal contact with researchers and with other administrators, through informal channels, workshops, and professional meetings [Havelock et al., 1969; Carlson, 1965a; Greenwood and Weiler, 1972]. It also seems clear that most educational research and development has been oriented to academic peer approval rather than to adopting innovations in the schools [Havelock et al., 1969; Glaser and Taylor, 1969]. It has been said, for example:

> Many academic scientists value the prestige that their tributions to basic research and theory give them in the

eyes of their peers more than whatever rewards might be obtained from clients who would find their work useful .... Much of the applied work in disciplinary departments is done by those who for one reason or another do not compete for the highest prizes of their disciplines [National Academy of Sciences, 1969, p. 93].

Policymakers who come to social scientists for advice often go away empty-handed. A local school superintendent in California addressed exactly this charge to the staff of one research and development center. 'They're always chasing theoretical rainbows, and frankly I doubt that there's a pot of educational gold at the end [Baldridge and Johnson, 1972, p. 33].'

It is clear that these misunderstandings (between researchers and practitioners) develop because there has not been a meeting of the minds between the research and the organization. The atmosphere, during early stages, of cordiality, implicit mutual assurances, and reciprocally unrealistic expectations compounds an already precarious balance .... The 'loser' is not just the agency or a disappointed researcher; it is the field, the clients, and all participants as well as future research endeavors [Glaser and Taylor, 1969, p. 91].

We have here a vicious circle: (a) many educators do not conceive of the scientific method and research as being of primary significance to their work; (b) this state of mind creates an atmosphere in which low priority is given to the conduct or utilization of research; (c) because of low evaluation or neglect, research continues to be a dubious enterprise; and (d) because condition (c) exists, condition (a) is perpetuated [Pellegrin, in Carlson, 1965b, pp. 71-72].

The present situation tends to combine several disadvantages:

- Researchers are interested in disciplinary prestige more than in problem-solving in the schools.
- Even when, as in the case of regional labs, there is considerable incentive to produce R&D results that can be applied in the schools, the gulf between innovation and implementation remains all too often unbridged.
- Researchers disseminate results through journal articles and reports; practitioners learn through briefings, meetings, and informal discussion.

- Research and development agencies follow an R&D change model that views the schools as passive adopters of new products, but the schools themselves decide to adopt and implement innovations in light of a host of organizational considerations which are not considered in the R&D model of change.
- Researchers and practitioners often don't talk the same language because their operating styles, perceptions of issues, and professional priorities are so different.

The policy implications appear evident, although the remedy is likely to be slow.

First, educational R&D organizations should be interdisciplinary and problem-oriented rather than disciplinary and methodology-oriented. This is not a criticism of either basic research or focusing on discipline or methodology. But in the context of this paper--how to increase the adoption of R&D products in the schools--they are evidently of little proximate value.

Second, R&D organizations should work more closely with principals, district administrators, and teacher representatives during the development period. Several such organizations regularly employ school administrators, on leave from their districts, in R&D planning. This practice should be extended. There is a delicate balance, of course, between systematically improving researchers' and school staffs' mutual understanding; and allowing research and development to be dominated by the institutional perceptions of the schools.

Third, R&D organizations should conduct regular seminars, workshops, and institutes for school district and state education agency staff, designed to communicate both R&D results and schoolmen's perceptions of appropriate priorities, implementation problems, and technical assistance requirements.

Fourth, it is important for R&D agencies to understand the nature of regional and national influence networks, and to identify

potential innovators. In the public schools, as elsewhere, there are organizations and individuals who are more disposed to innovate, and who feel less threatened than others by the prospects of change. For example, USOE in its experimental schools program and OEO in its voucher and performance contracting demonstrations have in effect identified a few such districts. R&D organizations can work with such innovators to demonstrate the new methods and find out how they work in practice, meanwhile working with broader leadership networks to disseminate the findings.

Fifth, and most important in the transition from innovation to implementation, is the need for R&D personnel to work closely with school staff during the implementation period. Otherwise, it is clear from the evidence [Goodlad, 1970; Gross et al., 1971] that the R&D task is cut off before its fulfillment. The view taken here is that incentives to adopt and incentives to implement are largely different from each other. Innovation and implementation work through different agents in the institutional setting. The federal or state agencies propose; school superintendents or principals dispose; the teachers and students transform.

Therefore, the R&D job does not end at the school district line or the schoolhouse door, and close collaboration with the schools is probably a necessary condition of implementing any innovations that depart from the established pattern of innovations that, as we have seen, the schools customarily accept. This approach means that R&D agencies will have to assure the training and recruitment of people who work well with both researchers and people in the schools. This form of technical assistance for implementing innovations will be expensive.

For the major innovations that proponents of school reform are seeking, it may often be a matter of years, not months, to build up the kind of orientations and mutual understanding that will be required and through a process of successive approximation, to create new institutional structures and values. It will in effect require R&D institutions to turn much more to a clinical model of change (one

which adapts general findings or processes to the specific circumstances of the client) and away from the engineering model, which offers a standardized product to the clients at large [Weiler, 1972; House, Kerins, and Steele, 1972].

3. Case Studies of Innovation

There is a sizable literature on educational innovation (see bibliography), including some interesting analyses of the success and failure of particular innovations [Smith and Keith, 1971; Gross et al., 1971; Carlson, 1965a]. However, the literature, with a few exceptions, does not describe the implementation process. As Goodlad (1970) and Gross et al. have pointed out, it is impossible to judge the merits of an innovation unless we have substantial information about how, and even *if*, it was implemented. If some innovations are, as Goodlad claims, implemented in name only, then the innovation remains untested. At the same time, such evidence clearly indicates a failure in the R&D process. Innovations that consistently remain unimplemented can hardly be regarded as arguments in favor of perpetuating existing R&D styles. Either the innovations are inappropriate, or the implementation arrangements, or both.

The discussion of the previous sections indicates a number of reasons why innovations might not be adopted or implemented. These arguments, based on unsystematic observation supplemented by a few case studies, need to be rejected or confirmed by more systematic case studies. Such studies can point the way to more effective strategies for development and implementation. Some R&D agencies have shown an ability to work with schools to implement innovations, while others have not; yet there is surprisingly little documentation of the record.

4. Analysis of Incentive Patterns

What are the institutional incentives that motivate school districts, administrators, school boards, teachers, state and federal educational agencies? Are the respective sets of incentives consistent with each other? If not, how are inconsistencies typically resolved?

In general, we would expect school districts' values to dominate in the resolution of interjurisdictional differences, since they are closer to operations than other jurisdictions, and exercise de facto control over funds, no matter how they are nominally earmarked.

But the relationships are complex. The Federal government has clearly forced state school agencies and local districts to pay more attention to disadvantaged students and to innovation than they would have otherwise. Changes in state education codes and in financial support regulations systematically affect local school districts' incentives and responses. The emergence of strong teacher unions has reduced school boards' and administrators' freedom of action, as has the emergence of a number of vocal and conflicting community interest groups.

In general, groups and institutions involved in the multi-bureaucratic structure of educational governance do not appear to gauge each other's motivations and responses well. Evidence for this lies in: (1) the frequently voiced disappointment of federal and state agencies in local districts' failure to do a good job in carrying out mandated programs; (2) local community groups' perception of school authorities as unresponsive; (3) district administrators' frequent impatience or contempt for state and federal agencies' inability to understand the local perspective; (4) R&D agencies' frequent ignorance of or disrespect for district administration; (5) teacher groups' increasing militance, reflecting impatience with the perspective of school boards and administrators; (6) the public's increasing unwillingness to vote more funds for schooling.

There is no accepted theory of interbureaucratic organizational behavior. Organization theory has concerned itself mostly with the internal structure and incentives of individual or representative bureaucracies and first steps toward a more realistic description of how bureaucracies interact are barely under way [see Levine, 1972]. Even this has not been worked out in enough detail at the school district or state and federal education agency level.

There have been formal treatments of interbureaucratic financial behavior [Barro, 1972] and descriptive treatments of individual

bureaucratic levels--federal, state, local, R&D agencies (see bibliography). But it seems safe to say that most external efforts to promote innovation in the schools have foundered in part through their ignorance of the tunes to which school districts must dance. Federal programs, for example, often seem to assume that because schools want to prevent high school dropouts, therefore Federal funding of dropout prevention programs will result in a coincidence of Federal and local interest. The reality is far more complex. School districts have a number of priorities, and dropout prevention ranks much higher in the verbal agenda than in the hidden one. For good reasons, the schools feel that some people should be encouraged to drop out and others discouraged. But dominating those perceptions is the need to prevent any important client groups from creating crises--to keep them at least relatively satisfied. Therefore, dropout prevention funds--like compensatory education funds, driver education funds, or any other largesse--will be spent as much as possible to keep parents, teachers, students, school boards, and "external" bureaucracies in some kind of equilibrium. The nominal purposes of the funds are regarded at the district level as constraints on the objective function, and one measure of an administrator's success is his ability to make the constraints non-operative, to allocate external funds so that they do double duty.

Therefore, if externally-encouraged innovative efforts are to avoid a great deal of waste motion, they must be based on a far more detailed appraisal of the reality of the schools as institutions than is now the case. For this reason, studies of the operative behavior of school districts in their relation to their own clients and to the state and federal bureaucracies they must deal with should be of high priority in R&D funding.

5. Restructuring System Incentives

A theme of this paper is that the schools, as a peculiar form of regulated public utility, have a different set of incentives to innovate than do competitive firms. It is undesirable to take the private market model as a general exemplar for school district

behavior (most people would be reluctant to allow only those who can afford schooling to obtain it), but it seems well worthwhile to experiment with changing the incentive system of the schools in a variety of ways. There is no guarantee that new system incentives will result in performance that satisfies society more than the present systems. But, given the great expense of schooling and widespread dissatisfaction with current performance, the social costs of experimental restructuring of incentives cannot be very great.

Experimentation with restructuring incentives should take four forms:

A. Changes in Market Structure

These experiments would cover:

- A range of voucher alternatives from the public school open enrollment version currently under way in Alum Rock to those that would include establishing new schools and allowing participation of existing private and public schools.
- Youth endowment plans under which each young person would have a lifetime entitlement of money to be spent on supplemental schooling or other beneficial use at the recipient's discretion--for example, supplemental educational or extra-curricular experiences during the elementary and secondary school years; college expenses, cost of private vocational schools as a substitute for high school; costs of going into business, etc. One version of this proposal, the educational bank, has been described by Killingsworth (1967).
- Permitting open enrollment across district lines among the public schools of a metropolitan area, with public funds following the student.

B. Changes in Locus of Control

Both greater centralization of control and greater decentralization of control are likely to lead to their own sets of systematic biases in incentives to adopt innovations and incentives to implement them. The object of experimentation and analysis should be to discern the nature of these effects. Obvious candidates for initial analysis are responses to innovation in New York City schools, as an example of decentralization to the neighborhood level; private schools, free schools, and alternative schools, as an example of decentralization to the school level; and jurisdictions such as the French and Swedish schools or Los Angeles and Chicago districts, as examples of centralized decisionmaking.

It should also be possible to mount new experiments, with assurances of long-term funding, such as paying school districts to decentralize decisionmaking to principals or to community boards or to teacher-student governance; or subsidizing a state government to centralize and implement innovative policies.

C. Changes in Individual Incentives

The schools have long resisted any moves to "deprofessionalize" the system, whether by paying people on the basis of performance or by allowing the schools to hire anyone they want to as teachers or administrators. Certification and the unified salary schedule are the shibboleths of professional educators. Some of the reasons are obvious: (1) certification offers the advantages of a sort of tariff barrier; (2) it also offers status--certification enhances the esteem of lawyers, doctors of philosophy or medicine, licensed plumbers and morticians--why should schoolmen not garner the same psychic benefits?

Some of the reasons are less obvious. Many teachers and administrators believe that both ends and means of their work are uncertain. Others believe that ends and means are certain, but unrevealed to those in position of authority. In either case, where does merit lie, and who should decide it? If salary differences are desirable as incentives and as recognition of increased social responsibilities

associated with aging, then why not condition salary rewards on objective measurable stigmata--years of experience in teaching and in learning--rather than on unverifiable judgments about individual merit? Stated differently, productivity criteria are one thing where some form of market appraisal or a generally approved surrogate exists (batting averages, journal articles, or shorthand speed)--the criteria may be resisted, but they are hard to gainsay entirely. Matters are quite different when each observer is free to assert his own criteria, or when centrally imposed criteria are widely regarded as arbitrary.

This leads to some conclusions for experimentation. Dispensing with certification requirements for recruitment and promotion should be tried out, under subsidy, in school districts. Merit pay experiments should preferably be implemented in association with reasonable evidence that certain kinds of teacher characteristics or behavior lead to better student outcomes than others. There has been a good deal of research on teaching, but rather little of it has been associated with student outcomes [Averch et al., 1972; Hanushek, 1970]. Beyond that, research is fragmentary and not conclusive.

In light of uncertainties about what merits should be compensated, it seems advisable to study through natural experiments (longitudinal studies of teachers and students) and planned experiments (assigning teachers with certain characteristics randomly to students) whether the objective correlates of merit can be determined. For both teachers and administrators, one dimension of merit to be compensated might be the successful implementation of specified innovations. Compensation could take the form of salary, or perhaps more acceptably, some agreed level of "free" funding for innovating schools or districts.

D. Clearer Standards for Accountability and Better Information Systems

Accountability and information systems for the schools have been carefully discussed [Barro, 1970; Coleman and Karweit, 1972; Dyer, 1970], but they remain in public discussion largely catchwords, two more footnotes for the historian of educational rhetoric.

Nonetheless, such catchwords, in this case as in others, represent a recognition of issues which, though dimly perceived, are fundamental to social choice. Why do such vague concepts as "accountability" and "information systems" represent something fundamental, and what can planned experiments do about them? The phrases are probably important because they recognize implicitly a search for consensus; and one task of experimentation can be to give that search some content. If the objectives of schooling are multiple and unclear, if there are no market tests of efficiency, if there are generally only weak performance criteria for R&D product adoption, and if, at the same time, the public is dissatisfied with its youth, and therefore with its schools, where should reforms begin? Why should the public endorse or the schools adopt, at considerable travail, new methods that will create political and institutional problems, when the resulting prospects for school improvement are so uncertain? In some sense, then, the call for accountability and for information is more than a blending of old nostrums--searching for scapegoats--and new ones--appealing to the balm of technology. It is a recognition of a disturbing situation. Unless the things the schools do can be tested in light of well-established and widely disseminated criteria, there is not much rational basis for preferring one policy over another. Even the obvious ones, like spending less money to put the same number of children through school at some average achievement levels, are unreliable. Saving money may be less important not only to the schools but also to the public than continuing to do things as before.

This paper contends that the search for accountability cannot be based on agreed objectives starting from first principles, because

there will never be agreement about the nature or priority of social objectives. Who is to decide ex ante what is the right combination of basic knowledge, vocational skills, child care, socialization, or motor development for the schools to produce? Instead, as suggested above, the present uncertainties should lead R&D planners to a strategy in which the process of experimentation is consciously used as the mechanism that helps define social values.

Market-oriented innovations, such as vouchers or the educational bank, are more or less consciously aimed in that direction. Many of the experiments suggested in this section are directed toward the same general goal, within the present public utility framework, primarily by trying to compensate for the innovative biases created by the local monopoly status; and by trying to assure that educational R&D is carried through to the implementation stage--the analogue in a competitive market would be production engineering--a function which does not automatically emerge from the dynamics of the educational marketplace, because there is no necessary payoff for implementing planned change in ways that mirror the developers' intent.

Once planned experimentation and analysis of existing natural experiments offer some idea of what different people in and out of the schools value, and what costs in money, in bureaucratic upheaval, and in alternative outcomes forsaken they are prepared to pay, we are at the threshold of genuine accountability, systems that could allow assessment of the progress of a teacher, an administrator, a school district toward specified goals. But it is only a threshold in the absence of widely disseminated information about the outcomes of schooling--in achievement, attitudes, career paths, in social integration, and perhaps ultimately in people's conceptions of education. Without comprehensive information flow to policymakers and the public, any new era of experimentation is likely to end up where past experiments have, in the research libraries. The widespread dissemination of information will give the public, as individuals and in various institutional roles, opportunities to campaign or be inert, on the basis of some more realistic appraisal of cause and effect than has yet been possible. If this is no guarantee of more effective

schooling, it at least comes closer to an uncertain ideal--public participation in an informal decisionmaking process. The devising of such comprehensive information systems in support of accountability has been discussed elsewhere [Coleman and Karweit, 1972; Farquhar and Boehm, 1971].

In conclusion, it should be pointed out that accountability in the sense of standards of accomplishment for school staff would require constant revisions. This would not reflect an effort by society to speed up the assembly line once initial norms were achieved; but society's tastes change and therefore the ordering of its preferences as expressed through experimental results would perennially impose new standards on the schools.

But in a diverse society, perennial change in standards implies that at any one time there will be a variety of standards--including, as pointed out above, the coexistence of incompatible standards. In private markets when consumers want different things, the response is to provide a variety of alternatives, allowing each consumer to choose the particular kind of housing, insurance, or toothbrush that comes closest to meeting his preferences in light of his means. Given a somewhat analogous set of competing demands in the public utility market of the schools, diversity will have to take place primarily in a public marketplace, which, as we have seen, behaves differently from a private one. In light of the standards implicit in this paper, a major focus of R&D policy should be--through experimentation and through incentives that encourage new patterns of institutional behavior--to encourage a long-overdue diversity of approaches to schooling. Even with more sophisticated approaches to R&D management and to the realities of implementation, the task will be long, costly, difficult. In the current state of knowledge, this process must be justified primarily on the grounds that an educational system which develops effective mechanisms for innovation is more likely to respond to changing social needs than one which is primarily centered on preserving the existing institutional order. This viewpoint implies that diversity in organizational response itself should be a prime target of policy.

BIBLIOGRAPHY

A Model for Innovation Adoption in Public School Districts: Research on the Characteristics of Selected School Systems as They Relate the Need for Appraisal, Acceptance and Use of Innovations, Arthur D. Little, Inc., Boston, Massachusetts, March 1968.

Agger, R. E., et al., *Educational Innovations in the Community*, U.S. Department of Health, Education and Welfare, University of Oregon, Eugene, Oregon, 1965.

Alternative Futures in American Education, Committee on Education and Labor, House of Representatives, January 1972.

Averch, H., et al., *How Effective is Schooling?* The Rand Corporation, Santa Monica, California, March 1972.

Baldridge, J. Victor and Rudolph Johnson, *The Impact of Educational R&D Centers and Laboratories: An Analysis of Effective Organizational Strategies*, Stanford University, May 15, 1972.

Barro, S. M., *An Approach to Developing Accountability Measures for the Public Schools*, The Rand Corporation, Santa Monica, California, P-4464, September 1970.

Barro, S. M., *Theoretical Models of School District Expenditure Determination and the Impact of Grants In Aid*, The Rand Corporation, Santa Monica, California, February 1972.

Becker, M. H., "Factors Affecting Diffusion of Innovations Among Health Professionals," *American Journal of Public Health*, Vol. 60, No. 2, 1970, pp. 294-304.

Beckhard, R., "Helping a Group with Planned Change: A Case Study," *Journal of Social Issues*, Vol. 15, No. 2, 1959, pp. 13-19.

Belitsky, A. Harvey, *Private Vocational Schools and Their Students*, Schenkman Publishing Co., Cambridge, Mass., 1969.

Bennis, W. G., et al., *The Planning of Change*, Holt, Rinehart and Winston, New York, 1969.

Blum, R. H., et al., "Staff Response to Innovation in a Mental Health Service," *American Journal of Public Health*, Vol. 54, 1964, pp. 1230-1240.

Boulding, K. E., "The Schooling Industry as a Possibly Pathological Section of the American Economy," *Review of Educational Research*, Vol. 42, No. 1, 1972, pp. 129-143.

Bowles, Samuel and Herbert Gintis, *IQ in the U.S. Class Structure*, Harvard University, Cambridge, Mass., July 1972.

Brickell, H. M., *Organizing New York State for Educational Change*, New York State Department of Education, December 1961.

Buchanan, P.C., *Change in School Systems*, National Training Laboratories, Washington, D.C., 1967.

Bushnell, Don D., et al., *Proceedings of the Conference on the Implementation of Educational Innovation*, SDC, 1964, pp. 1-318.

Bushnell D. S., *Planned Change in Education: A Systems Approach*, Harcourt Brace Jovanovich, Inc., New York, 1971.

Caffery, J. G., *The Innovational Matrix*, paper prepared for Institute of Government and Public Affairs Conference on Educational Innovations, UCLA Lake Arrowhead Conference Center, 17-20 December 1965.

Carlson, R. D., *Adoption of Educational Innovations*, The Center for the Advanced Study of Educational Administration, Eugene, Oregon, 1965.

Carlson, R. D., et al., *Change Processes in the Public Schools*, University of Oregon, Eugene, Oregon, 1965.

Cartwright, D., "Achieving Change in People," in W. G. Bennis, et al. (eds.), *The Planning of Change: Readings in the Applied Behavioral Sciences*, Holt Rinehart, and Winston, New York, 1962, pp. 698-710.

Chesler, M., et al., "Resistance to Research and Research Utilization the Death and Life of a Feedback Attempt," *Journal of Applied Behavioral Science*, Vol. 3, 1967, pp. 387-469.

Chorness, M. H., et al., *Decision Processes and Information Needs in Education: A Field Survey*, Far West Laboratory for Educational Research and Development, Berkeley, California, December 1968.

Chorness, M. H., et al., *Use of Resource Material and Decision Processes Associated with Educational Innovation*, Far West Laboratory of Educational Research and Development, Berkeley, California, February 1969.

Coleman, J. S., et al., *Medical Innovation: A Diffusion Study*, Bobbs-Merrill, New York, 1966.

Coleman, J. S., and Nancy L. Karweit, *Information Systems and Performance Measures in Schools*, Educational Technology Press, Englewood Cliffs, N. J., 1972.

Coleman, J. S., "Incentives in Education, Existing and Proposed," unpublished manuscript, 1972.

Committee for Economic Development, *Innovation in Education: New Directions for the American School*, McGraw-Hill Book Company, New York, 1969.

Committee for Economic Development, *The Schools and the Challenge of Innovation*, McGraw-Hill Book Company, New York, 1969.

Cooper, C. R. and B. Archambault (eds.), *Communication, Dissemination and Utilization of Research Information in Rehabilitation Counseling*, proceedings of a regional conference sponsored by the Department of Guidance and Psychological Services, Springfield, Massachusetts, 1968.

Croker, G. W., "Some Principles Regarding the Utilization of Social Science Research Within the Military," in *Case Studies in Bringing Behavioral Science Into Use*, Vol. 1, Institute for Communication Research, Stanford University, 1961, pp. 112-125

Dreeben, Robert, *On What Is Learned in School*, Addison-Wesley Publishing Co., Reading, Mass., 1968.

Dyer, Henry S., "Toward Objective Criteria of Professional Accountability in the Schools of New York City," *Phi Delta Kappan*, Vol. LII, No. 4 [Dec. 1970], pp. 212-216.

Educational Research: Prospects and Priorities, Committee on Education and Labor, House of Representatives, January 1972.

Eidell, T. L. and J. M. Kitchel (eds.), *Knowledge Production and Utilization in Educational Administration*, Center for the Advanced Study of Educational Administration, University of Oregon, Eugene, Oregon, 1968.

Evans, R. I., et al., *Resistance to Innovation in Higher Education: A Social Psychological Exploration Focused on Television and the Establishment*, Jossey-Bass, San Francisco, California, 1968.

Fairweather, G. W., *Methods for Experimental Social Innovation*, John Wiley and Sons, Inc., New York, 1967.

Falton, B. J., ed., *Educational Innovation in the United States*, Phi Delta Kappa, Bloomington, Indiana, 1966.

Farquhar, John A. and Barry W. Boehm, *An Information System for Educational Management: Vol. 1 -- Design Considerations*, The Rand Corporation, Santa Monica, California, R-930, December 1971.

Flanagan, J. C., "Case Studies on the Utilization of Behavioral Science Research," in *Case Studies in Bringing Behavioral Science into Use. Studies in the Utilization of Behavioral Science*, Vol. 1, Institute for Communication Research, Stanford University, 1961, pp. 36-46.

Fleigal, F. C., et al., "Attributes of Innovations as Factors in Diffusion," *American Journal of Sociology*, Vol. 72, No. 3, 1966, pp. 235-248.

Foshay, A. W., *Curriculum for the 70s: An Agenda for Invention*, National Education Association Center for the Study of Instruction, 1970.

Gideonese, H. D., "Research Development and the Improvement of Education," *Science*, 1 November 1968.

Glaser, E. M., and C. G. Wrenn, *Pulling Research, Experimental and Demonstration Findings to Use*, Office of Manpower Policy, Evaluation and Research, U. S. Department of Labor, Washington, D. C., 1966.

Glaser, E. M., and H. L. Ross, *Increasing the Utilization of Applied Research Results*, Human Interaction Research Institute, Los Angeles, 1971.

Glaser, E. M., and S. Taylor, *Factors Influencing the Success of Applied Research*, Institute of Mental Health, Department of Health, Education, and Welfare, Washington, D. C., final report on Contract No. 43-67-1365, January 1969.

Glaser, E. M., et al., *Utilization of Applicable Research and Demonstration Results*, Final report to Vocational Rehabilitation Administration, Department of Health, Education, and Welfare, Washington, D. C., Project RD-1263-G, 1966.

Goodlad, J. I., et al., *Behind the Classroom Door*, Charles A. Jones Publishing Co., New York, 1970.

Greenwood, Peter and Daniel Weiler, *Alternative Models for The ERIC Clearinghouse Network*, The Rand Corporation, Santa Monica, California, January 1972.

Gross, N., et al., *Implementing Organizational Innovations: A Sociological Analysis of Planned Educational Change*, Basic Books, Inc., New York, 1971.

Guba, Egon, *Methodological Strategies for Educational Change*, ERIC Document Reproduction Service, Washington, D.C., 8-10 November 1965.

Guetzkow, H., "Conversion Barriers in Using the Social Sciences," *Administrative Science Quarterly*, Vol. 4, No. 1, June 1959, pp. 68-81

Hanushek, Eric A., *The Value of Teachers In Teaching*, The Rand Corporation, Santa Monica, California, RM-6362, December 1970.

Havelock, R. G., *A Guide to Innovation in Education*, Center for Research on the Utilization of Scientific Knowledge, Institute for Social Research, University of Michigan, Ann Arbor, Michigan, 1970.

Havelock, R. G., et. al., *Major Works on Change in Education*, Center for Research on Utilization of Scientific Knowledge, October 1969.

Havelock, R. G., *Planning for Innovation Through Dissemination and Utilization of Knowledge*, Institute for Social Research, Center for Research on Utilization of Scientific Knowledge, University of Michigan, Ann Arbor, Michigan, July 1969.

House, R. E., Thomas Kerins, Joe M. Steele, "A Test of the Research and Development Model of Change," *Educational Administration Quarterly*, January 1972, pp.1-14.

Husén, T., et al., *Educational Research and Educational Change: The Case of Sweden*, John Wiley & Sons, Inc., New York, 1967.

Innovation and Experiment in Education, A Progress Report of the Panel on Educational Research and Development, The President's Science Advisory Committee, Washington, D. C., March 1964.

Jackson, K., *The Politics of School Community Relations*, The ERIC Clearinghouse on Educational Management, University of Oregon, Eugene, Oregon, March 1971.

Janowitz M., *Institute Building in Urban Education*, Russell Sage Foundation, New York, 1969.

Jencks, Christopher, et al., *Inequality*, Basic Books, Inc., New York, 1972.

Joly, J., "Research and Innovation: Two Solitudes?" *Canadian Education and Research Digest*, Vol. 2, 1967, pp. 184-194.

Katz, E., "The Social Itinerary of Technical Change: Two Studies on the Diffusion of Innovation," *Human Organization*, Vol. 20, 1961, pp. 70-82.

Kiesling, H. J., *Input and Output in California Compensatory Projects*, The Rand Corporation, Santa Monica, Calif., October 1971.

Killingsworth, C. C., "How to Pay for Higher Education," address to Economic Society of Michigan, mimeographed, 1967.

Klitgaard, R. E., Carl Sewell, T. S. Donaldson, and G. R. Hall, *Are There Unusually Effective Schools? An Exploratory Study*, unpublished manuscript, The Rand Corporation, Santa Monica, Calif., January 1973.

La Pure, R. T., *Social Change*, McGraw-Hill, New York, 1965.

Levine, Robert A., *Public Planning: Failure and Redirection*, Basic Books, New York, 1972.

Lindeman, John, et al., *Some Aspects of Educational Research and Development in the United States - Report for the OECD Review*, Syracuse University Research Corp., December 1968.

Lippitt, Ronald et al., *Perspectives on Educational Change*, Appleton-Century-Crofts, New York, 1966.

Lippitt, Ronald, Jeanne Watson and Bruce Westley, *The Dynamics of Planned Change*, Harcourt Brace and Company, New York, 1958.

Mackie, R. R., and P. R. Christensen, *Translation and Application of Psychological Research*, Technical Report 716-1, Human Factors Research, Inc., Goleta, California, 1967.

Maguire, L. M., et al., *An Annotated Bibliography on Administering for Change*, Research for Better Schools, Inc., October 1971.

Maguire, L. M., *Observations and Analysis of the Literature on Change*, Research for Better Schools, Inc., June 1970.

Mann, F. C., and F. W. Neff, *Managing Major Change in Organizations*, Foundation for Research on Human Behavior, Ann Arbor, Michigan, 1961.

March, James, *The Technology of Foolishness*, unpublished paper, Stanford University, 1972.

Marcum, R. Laverne, *Organizational Climate and the Adoption of Educational Innovations*, Utah State University, Logan, Utah, 1968.

Mansfield, Edwin, "Speed of Response of Firms in New Techniques," *Quarterly Journal of Economics*, 1963, Vol. 77, pp. 290-311.

Miles, M. B., *Innovation in Education*, Bureau of Publications, Teachers College, Columbia University, New York, 1964.

Miller, D. R., *Planned Change in Education*, Operation PEP: A Statewide Project to Prepare Educational Planners for California, May 1968.

Miller, R. I., *Perspective on Educational Change*, Appleton-Century-Crofts, New York, 1967.

Mort, P. R., and F. G. Cornell, *American Schools in Transition*, Bureau of Publications, Teachers College, New York, 1941.

Nasatir, D., *Resistance to Innovation in American Education*, paper prepared for Institute to Government and Public Affairs Conference on Educational Innovations, UCLA Lake Arrowhead Conference Center, 17-20 December 1965.

National Academy of Sciences and Social Science Research Council, *The Behavioral and Social Sciences: Outlook and Needs*, Prentice Hall, Englewood Cliffs, N. J., 1969.

Oettinger, A. G., *Run, Computer, Run--The Mythology of Educational Innovation*, Harvard University Press, Cambridge, Massachusetts, 1969.

Planning for Creative Change in Mental Health Services: A Manual on Research Utilization, No. (HSM) 71-9059, National Institute of Mental Health, Rockville, Maryland.

Planning for Creative Change in Mental Health Services: Information Sources and How to Use Them, No. (HSM) 71-9058, National Institute of Mental Health, Rockville, Maryland, 1971.

Planning for Creative Change in Mental Health Services: Use of Program Evaluation, No. (HSM) 71-9057, National Institute of Mental Health, Rockville, Maryland.

Profiles of Excellence: Recommended Criteria for Evaluating the Quality of a Local School System, Office of Professional Development and Welfare, 1966.

Programming Interpersonal Curricula for Adolescents (PICA)--Project Year Three: A Laboratory Model, Institute for Behavioral Research, Inc., October 1971.

Purpose and Process: Readings in Educational Research and Development, Committee on Education and Labor, House of Representatives, January 1972.

Rickland, M., *Traveling Seminar and Conference for the Implementation of Educational Innovations*, Technical Memorandum Series 2691, SDC, Santa Monica, California, 1955.

Rittenhouse, Carl H., "Innovation Problems and Information Needs of Education Practitioners, *ERIC Report*, Vol. 1, Department of Health, Education, and Welfare, Education Resources Information Center, Washington, D. C., May 1970.

Rogers, E. M., and F. Scoemaker, *Communications of Innovations*, The Free Press, New York, 1971.

Rogers, E. M., *Diffusion of Innovations*, The Free Press, New York, 1962.

Rogers, E. M., and L. Svenning, *Managing Change*, Operation PEP (a statewide project to prepare educational planners for California), U. S. Office of Education, Department of Health, Education, and Welfare, September 1969.

Rosenau, L. H., and J. Hemphill, *Utilization of NIE Output*, Far West Laboratory for Educational Research and Development, Berkeley, California.

Smith, Louis M. and Pat M. Keith, *Anatomy of Educational Innovation*, John Wiley and Sons, New York, 1971.

Stufflebeam, Daniel L., *Proposal to Design New Patterns for Training Research, Development Demonstration/Dissementation and Evaluation Personnel in Education*, U. S. Department of Health, Education, and Welfare, Washington, D. C., December 1970.

Taylor, J. B., *Introducing Social Innovation*, paper presented at meeting of American Psychological Association, San Francisco, California, September 1968.

The Management of Innovation in Education, Center for Educational Research and Innovation, 1969.

Watson, G., ed., *Concepts for Social Change*, National Training Laboratories, 1967.

Watson, G., and E. M. Glaser, "What We Have Learned About Planning for Change," *Management Review*, Vol. 54, No. 11, 1965, pp. 34-46.

Weiler, Daniel, "The Dissemination of Educational R&D Products: Where Do We Go From Here?," unpublished paper presented to NIE Symposium on Educational R&D, Washington, D. C., December 1972.

Wirt, E. W., and M. W. Kirst, *The Political Web of American Schools*, Little Brown, Boston, 1972.

York, L. J., *Arrangements and Training for Effective Use of Educational R&D Information*, Far West Laboratory for Educational Research and Development, Berkeley, California, December 1968.

Comments on Pincus Paper by David S. Mundel

In this paper, John Pincus sets out several propositions about the non-market structures and the resulting incentive systems of public school systems. These characteristics of public education relate to (1) the adoption of innovations and (2) the implementation of innovations within the schools, and thus have serious implications for the design of research and development programs which are designed to improve or change educational performance. This paper may have important positive effects on the reform of educational R and D itself, because much of the past and current R and D effort seems, at least to this observer, to be virtually unaware of the system toward whose reform it is oriented.

The first proposition which appears explicitly and implicitly throughout the paper is that public schools unlike private firms in competitive market situations are reluctant to innovate. This reluctance is compounded by both the monopoly and non-market characteristics of public school systems and the special qualities of education. Although this proposition is probably true, its importance is somewhat less than Pincus is willing to believe. Firstly, private firms in competitive and other markets are not all that willing to change their traditional practices. Secondly, given the inertia of the educational system and its politics and the 'special character' with which elementary and secondary education are publicly labeled, it seems unlikely that a competitive market system of education will come into existence in the near future. Therefore, the 'high' responsiveness of the competitive market is unavailable and thus somewhat uninteresting.

In order that we can expand our discussion, let me suggest some propositions which should be considered along with John's first.

It is not all that useful to understand that one institution is less or more willing than another to adopt an innovation unless one can introduce the more willing institutional form. Given the low likelihood of doing this, what is useful is to understand our ability to (1) marginally change the existing institutions' willingness to adopt and implement innovations and (2) change the innovations so as to change their treatment by an existing institution.

A second major proposition of the paper is that the fact that schools play a shared role in the education process alters their institutional character and innovative practices. Other influences -- heredity, family and peer group environment, communications media, etc. -- are also contributors to an individual's educational performance. Although these other factors may be found to be statistically the prime or principal determinants to educational performance, they may not be. There may be important aspects of the patterns in which schools apply their resources, attention, and incentives among students that would make the statistical appearances inaccurate assessments of reality. These patterns, themselves, may result from the same characteristics which limit the schools innovation and implementation.

Let me make this case concrete by digression into the problems of street cleanliness. I would imagine that if I assigned a student to predict the determinant of the presence of street litter, he would collect information on the dependent variable -- pieces of

cter per block -- and several independent variables such as (1) istence of street cleaning activities (2) the presence of multi- e unit residences and (3) income and education of the area's popula- on. The results of this study would probably show that street eaning has very little, if any, impact on the presence of litter. l of us would question the interpretation of this result as prov- g that the prime determinant of the presence of litter was the mily characteristics of the residents of the neighborhood and at consequentially no reform in our street cleaning procedures s likely to change our observations. I think we should be simi- rly questioning about the statistical interpretation of our wide nge of production function studies of the educational process.

The implications of this <u>second proposition</u> on the design of R and D program are:

1. We should look at areas of the schooling process which have not been looked at extensively by existing research in our search for the parts of schooling that matter.

2. We should not approach schools as if they have or can possibly have little if any impact because the likely results of believing this proposition are that we will find that, in fact, schools don't have any affect and that consequently we shouldn't attempt to influence or to change them.

One of the more important 'incentives for innovation in the blic schools' might be an improved knowledge about what they rrently do and why they do it.

A third proposition put forth by the author is that there are veral active constraints which limit the impact of an innovation iented R and D policy. The constraints include (1) bureaucratic ones, 2) external, social and political pressures and (3) the necessity or elite consensus.

I think that these and other aspects which limit innovation and change should not be looked at as if they are fixed and unalterable constraints within the existing school systems. It is true that they are factors which limit innovation and change in some, possibly most, schools, but it is not true, nor should we operate on the belief that it is true, that these constraints are fixed and unalterable, even within the current organizational form of public education. A research and development policy which took into account these factors and operated in such a way as to live within them and ameliorate their strength would seem potentially most fruitful. Such an R and D strategy is succinctly mapped out in the following paragraph of the paper.

> If we can identify the kinds of innovations that are likely to be adopted by school districts that follow such behavioral styles, we may be able to identify ways that R and D products can be oriented in order to gain acceptance. As soon becomes apparent, the three conditions favorable to adoption of innovations in the present setting are themselves rather restrictive. Advocates of substantial innovation in the public schools aren't likely to be very satisfied by a R and D dissemination strategy which takes these conditions as operative constraints. Therefore, it is useful to examine the ways in which R and D dissemination policy could take advantage of the existing structure of market and bureaucratic incentives and also to examine the ways in which these incentives could be modified by conscious R and D policy.

A fourth proposition of the paper regards the difficult connection between innovation and subsequent implementation. Failures to implement are due to a wide variety of factors. R and D organizations provide little guidance and in many cases seem unwilling to realize that implementation is their problem as well as is design and development of innovations. The schools themselves may know little about implementation of change, and thirdly, individuals within

the schools (teachers, administrators, and students alike) may simply not accept the obligations to change their behavior patterns in ways which are consistent with the desired implementation of an innovation.

The consequences of these propositions for the design of an R and D program are many fold. They include studying the implementation process more fully through case studies of innovation and implementation and encouraging the collaboration of R and D performers with educators. The author failed to mention a major possible path toward improving implementation. The R and D program itself should create within itself a system of incentives which reward development and innovation efforts which work actively on the implementation of their recommended reforms and which achieve implementation. This system of incentives should not be simply the review of implementation plans within a project grant decision process. Rewards or bounties for a successful and rapid implementation of change should be established.

A fifth proposition which is implicitly developed and supported in the paper is that principal avenue of approach toward stimulating innovations should be the development of incentives which cause schools to become more open to a variety of innovations. These incentives would come largely from the introduction of changes in the market structure within which schools operate. Other avenues to innovation and implementations are also mentioned. These include changing R and D policy so that innovations appeal more effectively to existing school situations and creating policies which permit the public to assess more accurately and completely what

schools are doing and what they are not doing. All three strategies should underlie a well designed program. The choice of one avenue versus another should be based on the likelihood of and need for immediate, intermediate and long run effects which might result from different approaches. An R and D strategy which concentrates entirely on the introduction of changes in the market structure may if it is successful have long run effects which are dramatic and desirable. It may have, however, few short run and intermediate effects on what happens within the schools. This failure may reduce the ability of the R and D policy to have the long run effects and may jeopardize the existence of the entire R and D program itself. The reduced probability of long term effects comes from my perception that an environment of meaningful rather than faddish change within the schools and within education in general needs to be created before a large scale structural and institutional change can and will be implemented and be successful. One method of establishing an increased willingness to change may be to stimulate individuals and organizations to undergo a series of small changes with some frequency. Another important stimulus is the development of information about the relative performance failures and successes of the schools as they are. Until it is more widely believed that schools have failed (if they, in fact, have) it is unlikely that the market structural changes which are often spoken about and at times supported will be instituted.

In brief, Pincus has set out several important propositions about innovation and the implementation of reforms in schools that should provide guidance to current efforts to reform education

research and development itself. The propositions appear accurate but the policy guidance which Pincus derives from some of them seems inappropriate and potentially ineffective.

General Discussion of Weiler and Pincus Papers and Clark and Mundel Comments

These papers were discussed together because of the complex interrelation between the incentives for change inherent in public school systems and the vehicles used for disseminating innovations. Generally speaking, it was agreed, over the short term, that it was important to identify a client's incentive structure and then to develop products and means of disseminating these products which build on this incentive structure. For the long term, it may be preferable to alter the incentive structure of the school system but this takes considerably longer than the NIE or education R & D has before they will be called upon to justify their activities and expenditures. Thus, given these latter considerations, the short term strategy seems to be the most logical to follow during the early stages of the NIE's development.

One of the most important and problematic aspects of the dissemination process is the determination of successful or effective implementation of innovations. Inherent in both papers is the assumption that the implementation of an innovation was the coyping of some template. There was no attempt to question the desirability of this for education. Should innovations be adopted without modification? Can they? If not, then how does one maintain the integrity of an innovation? Further, it would seem that these questions would become more important, if and when the Federal government assumes a greater responsibility in the establishment of dissemination procedures and guidelines.

Should this occur, the need for assessing the effectiveness of innovations will also increase. At present, there are few generalizable criteria for evaluating the effectiveness of innovations. If the Federal government expands its role in the dissemination process, then this deficiency will become critical not merely after products are disseminated but also in the selection

of products to be disseminated and the identification of clients. Implicit throughout this entire process are internalized normative evaluative criteria which determine an individual's decisions in these circumstances. What is important here is to determine what these implicit values are and then to decide whether or not mormative standards are necessary and/or advisable.

As in the other sessions, the discussion shifted to the role of the teacher in the adoption and implementation of innovations. More specifically, the question was raised as to how one would go about changing incentives here without trespassing on parental incentives which could, in turn, lead to over-regulation of schools. At present, the incentive structure of schools tends to discourage teachers' initiating innovations on their own. How could one encourage innovation without instituting some kind of controls over the kinds of innovations being adopted? Further, who set these standards -- the NIE, the individual school systems, the parents, or the administrators? Who should set such standards?

With regard to these latter issues, the role of parents was discussed briefly. At present, school systems seek to average out parental demands within the system. However, if parental demands could be aggregated by types this might establish countervailing pressures on the bureaucracy in the direction of change. Again, one might question whether this would really change the balance of power in the public schools or merely perpetuate the same interests and influence over a larger population? Moreover, how would it be possible to arrive at some kind of a consensus with regard to educational goals and policies given the potential size and heterogeneity of parental groups? Lastly, what impact would increased parental participation have on the organizational and programmatic concerns inherent in the dissemination process.

EVALUATING THE NATIONAL INSTITUTE OF EDUCATION

by

Senta A. Raizen

Paper prepared for the
Symposium on Educational Research and Development

Washington, D.C.

December 11-12, 1972
revised January 23, 1973

EVALUATING THE NATIONAL INSTITUTE OF EDUCATION

Senta A. Raizen*

The Rand Corporation, Santa Monica, California

The introduction in 1965 of the Planning-Programming-Budgeting system into government has brought in its wake increasing demands for evaluating the effectiveness of government programs. As experience with assessment of social programs, and particularly educational experimentations, has accumulated, knowledgeable researchers have come to advocate that evaluation be made an integral part of program development (Light and Smith, 1970; Smith and Bissell, 1970). But to my knowledge, it is unique for a government R&D agency to consider at its inception by what standards its accomplishments should be judged. Perhaps in the case of the National Institute of Education (NIE), which was created to solve educational problems, this early concern with evaluation is an expression of the current disenchantment with R&D, or of the erosion of the formerly deep-rooted American faith in education as the solution to most social problems. Indeed, questions have been raised as to the impact that a government R&D agency can have on education, in view of our limited knowledge about education and R&D (Cohen, 1972). Further, some critics do not hold out much hope for evaluation as a consequential means of influencing policy or practice (Fox, 1967). Since the creation of a new government agency such as the NIE is itself a form of social action, however, attempts to assess its effectiveness will inevitably be made. If the NIE can guide these attempts by developing legitimate standards for its evaluation so that results will be both useful and actually used, one of the NIE's first successes may be to provide an acceptable pattern for other government agencies.

---

*An earlier version of this paper was presented at a Symposium on Educational Research and Development sponsored by the NIE in Washington, D.C., on December 11, 1972. The views expressed are those of the author. They should not be interpreted as reflecting the views of The Rand Corporation or the official opinion or policy of any of its governmental or private research sponsors.

Rossi and Williams (1972) note a number of problems and risks in developing and using evaluative results: conceptual and methodological problems, scarcity of competent people, and bureaucratic and political risks. Bureaucratic and political impediments are likely to operate particularly strongly in the case of evaluating overall performance of an agency. This is not only because, "...neither social service delivery systems nor government programs are organized to generate information about their effectiveness" (Rivlin, 1971, p. 64), but also because, as the new director of the NIE himself stated in a paper appraising the evaluation of federal manpower programs (Glennan, 1969, p. 45), "(1) Most programs and most agencies are reluctant to be evaluated; (2) if they must be evaluated, they will seek to find evaluation designs that have the greatest probability of supporting the status quo."

This paper is an effort to clarify the problem of evaluating the NIE and begin the process of developing satisfactory performance criteria. It assumes that R&D carried out by a federal agency can contribute to education, and that it is therefore in the interest of all -- the government, the R&D communities, the education professions, students and parents, and taxpayers -- to see the agency succeed. Unfortunately, the paper has had to be prepared in the absence of NIE-generated programs and of organizational structure for the new agency, thus limiting the discussion of evaluation methods and criteria to rather general and abstract suggestions. These need to be developed in greater detail as agency plans and organization are formulated.

I. THE CONTEXT FOR EVALUATION OF THE NIE

To understand just how difficult the task is, we must take a look at the context within which the NIE is going to operate. The old five "w's" of the newspaper reporter -- *who, what, why, when, where* -- can help us define this context.

Let me change the order slightly and deal with the *what* first. This is essentially a question of defining the objectives of the NIE. The legislative charter of the NIE is not much help since it is framed so broadly that it merely transforms the overall question of whether, and how much, the NIE is improving education into four questions, namely

whether and how much, the NIE is:

"(A) helping to solve or to alleviate the problems of, and achieve the objectives of American education;

(B) advancing the practice of education, as an art, science, and profession;

(C) ...strengthening...the scientific and technological foundations of education; and

(D) building an effective educational research and development system."*

In order to *evaluate* anything, as the word implies, one must know clearly what is being valued, that is, one must define the desired directions of change. Lord Rothschild (1972) advocates that the client of R&D -- either the government agency or the ultimate users -- decide specific program objectives. It would appear, therefore, that one of the first tasks for the NIE is to translate the general goals in its legislation into operational objectives that reflect consumer needs. For the NIE, this is a rather more complex undertaking than for an agency like the NIH, on which the NIE is superficially modeled. There are clear consensus goals for R&D in health: curing cancer, reducing the incidence of dental caries, eliminating stroke and heart disease. While the choice of strategies and resource allocations for R&D to attain these goals may often be difficult, at least they are undisputed social goods and clearly perceived. However, as we move into the domains of mental health and human development where goals are less easily defined, we find the cognizant agencies having greater difficulty justifying their programs. In the case of education, there appear indeed to be consensus goals, at least at the level of public rhetoric: ability to get a "good" job -- preferably one yielding financial success,** learning to get along with others,** achieving some acceptable level of reading and other cognitive skills by the end of high school, escaping poverty, making the schools work better for the children of the poor. But by now we know that all these goals have

---

*These four missions are quoted from *Education Amendments of 1972*.

**Obtaining a better job (44 percent), getting along better with people (43 percent), and financial success (38 percent) were the top goals for education of their children chosen by adults in a recent Gallup reported in *Phi Delta Kappan*, September 1972, p. 33.

one of two (sometimes both) characteristics: They are not equally valued by different client groups and therefore often in competition for resources with each other and with other highly valued educational goals, or they cannot be attained solely or even primarily through education. For example, investing the sizable resources necessary to make the bottom ten percent of students reach reading norms would either require withdrawing resources from other instructional areas now considered important or conflict directly with another current goal -- holding the line on steeply rising educational expenditures. Attaining a good job requires a set of attitudes, abilities, and credentials to which education can contribute, but it also requires that there be enough good jobs to go around, a function of the labor market rather than educational processes.

Thus, the question of *what* becomes one of defining important and relevant problems, relevant in the sense that they actually fall within the domain of education and are amenable to R&D approaches. The NIE can itself contribute to the validity of judgments about its programs by phrasing its objectives to imply appropriate rather than inappropriate criteria. To elaborate the job example, reduction of unemployment is an inappropriate criterion for R&D in education, but additional understanding of who is unemployed because of lack of skills, and the degree of success of new educational systems that deliver the needed skills are both relevant to assessing a program concerned with R&D in career education.

An equally important concern is: *who* should evaluate the NIE? Or, perhaps more realistically, who will evaluate the NIE? This is a complicated question that can be answered by a simple declarative sentence: The NIE is an agency of the federal government in the Department of Health, Education, and Welfare charged with carrying out R&D in education. Each of the nouns in this sentence corresponds to a set of evaluators. Any agency of the federal government will be judged by the Executive Office (currently that means largely the OMB) in the context of the President's annual budget formulation and by the cognizant committees of Congress in the context of legislation (authorization committees) and financing (appropriation committees). Any component of HEW must also account

to the Secretary, the Assistant Secretary for Planning and Evaluation, and -- in the case of an education agency -- the Assistant Secretary for Education. An R&D agency will be judged by the R&D communities that are or would like to be its client groups; an agency created to improve education will receive critical attention from all those who have a stake in that enterprise.

To appreciate the special difficulties faced by the NIE in regard to the last two groups, a comparison with the NIH is again illuminating. The R&D communities that interact with the NIH are well-defined and share a common set of belief structures and interests, not only because they are drawn from the disciplinary bases of the biomedical sciences, but because the first director of the NIH made it his prime responsibility to establish the desired relationships. The case of the NIE is quite different; in addition to the established educational research community, researchers from many different disciplines and applied fields -- from the natural and behavioral sciences to the humanities, from operations research to communications technology -- can and do claim that they can contribute to, and therefore assess, R&D in education.* These disparate groups hardly understand each other's languages; instead of sharing a common outlook, they are ignorant of each other or, if brought into contact, often sharply at odds in defining R&D problems, in preferred R&D styles, and in assessing outcomes. As to those who have a stake in education, the NIE is faced with two facts: powerful and vocal professional organizations (some two million strong) who consider themselves in an expert position to judge the effectiveness of educational R&D, and -- in great contrast to medicine -- the widespread belief on the part of the ultimate consumer (student, parent, employer, taxpayer) that he knows quite as much about education, having gone through it, as the professional.

As one considers the kinds of questions these different overseers are likely to ask, one comes to the *why*, the purpose of evaluation.

---

*Insofar as the NIH is having to concern itself increasingly with delivery of health services, it will have to involve a wider spectrum of performers, and its problems will be more like those of the NIE.

Levine and Williams (1971) note two such purposes: to affect resource allocation and to improve R&D strategies. The governmental overseers are likely to be most concerned with the first purpose, the R&D communities with both allocations and strategies, and the education professionals and consumers with the eventual results of allocations and strategies. Resource allocations are usually the result of many considerations, however, and objective assessment of the benefits of a particular program or agency as compared to others with claims on the federal dollar often play only a minor part. Evaluation cannot, and should not, serve as a substitute for good judgment. Nevertheless, the HEW leadership and OMB will expect evaluation to produce information on the potential importance of each NIE program, potential payoff, and likelihood of success. Congress will have similar concerns and, in addition, will be sensitive to geographic and institutional distributions of funds and second-order benefits. The NIE and its advisers would do well to construct ahead of time some evaluation criteria responsive to such questions; I shall try to suggest a few later in this paper.

The various R&D communities will (whether asked or not) assess the quality of the R&D output, with implications for improvements in R&D strategy; a concern with who receives funding for what purposes will hardly be divorced from this assessment, but the criteria for appropriate distribution will no doubt be quite different from those of Congress. The judgment of consumers is likely to be influenced by governmental and R&D performer groups insofar as their evaluations receive public notice -- though the influence often may take opposite directions for different consumers. But consumers will react with much greater intensity to programs having a direct impact on them, either as practitioners or as recipients of education. This reaction can be justified, as in the case of frustration with performance of inner-city schools in the face of ESEA Title I efforts, or capricious, as witness the furor over the "new math." Although the NIE will not be able to discount the unsolicited judgments of clients and consumers, these are not likely always to provide the optimal input for improving

its R&D strategies and operations.* There is another group of observers, the Advisory Council, that should be specifically charged with the responsibility of synthesizing judgments from all the evaluating groups -- self-appointed and solicited -- in order to make its own assessments and provide feedback to the NIE on needed changes to improve its performance.

The *when* is a critical problem for the NIE. Federal resource decisions are made in the course of the annual budget cycle; the present climate for educational R&D is not likely to permit growth or perhaps even stability without evidence as to achievement for dollars invested even within the first couple of years, unreasonable as that may seem given the difficulty of some of the problems and the long-range effects of most educational interventions. This implies that, no matter what other considerations go into choice of programs, there must be some activities designed to yield short-term successes, and some which are convincing demonstrations that progress is being made toward solving some difficult problems. Again, it will be necessary to spell out appropriate indicators for such successes and demonstrations, so that rational assessment can inform the decisions that are going to be made in any case by the executive branch and Congress.

One would hope for more leadtime for judgments from the R&D communities, education professionals, and ultimate consumers on the efficacy of a new agency's programs and operations, but here also history has predisposed many of us (for we all fall into one of these three groups) toward impatience, if not skepticism. Signals as to its competence will have to be given early if the NIE is to avoid a premature -- and negative -- assessment of "more of the same." Thus, the NIE faces severe time constraints, yet planning for long-range evaluation may be as important for its future as concern for immediate survival. Some of the most significant contributions of R&D to education are likely to be efforts resulting in the design of improved products, practices, and perhaps entire new systems of delivery for education, and such efforts may well consume five to ten years, with valid assessment stretching even beyond.

---

*The NIE might, however, develop consumer-oriented evaluation procedures for products developed under its sponsorship to clarify the pur- and appropriate applications of those products.

Therefore, while the NIE will have no choice about short-range accountability, it must explicitly build toward a demonstrable record of achievement measured on a time scale appropriate to design efforts.*

In journalistic practice, the *where* tends to come at the bottom of the list as providing the least important bit of information. In the present context, it might be considered synonymous with *who* if we interpret it as meaning where assessment questions will be asked. I wish, however, to consider instead *where* they might be answered, or better, *how* they might be answered. The remainder of this paper will, therefore, deal with the where-how of evaluating the NIE in the climate of the existing constraints.

II. THE DIMENSIONS OF EVALUATION

The dimensions of evaluation are implicit in who is rendering judgment with what purpose. They can be subsumed under four general headings: technical quality of the R&D, choice of the questions or problems being addressed, effectiveness of program output, and distribution of funds and of second-order benefits. Each of these dimensions has associated with it a series of questions that can help us define relevant criteria and perhaps even some appropriate methodologies.

Technical Quality**

The caliber of the R&D supported by the NIE is of most direct concern to the R&D communities, although it will, in the long run, affect the judgments of other groups as well, as quality begins to impact on the agency's ability to address problems these groups perceive to be important. Some questions useful in structuring any assessment of R&D quality are:

- What are the fields (and subfields) of activity?
- On what basis are they selected?

---

*Iterative engineering characterizes successful design. Whether the design is to be for a hardware system such as a moon launch or for a service system such as design and implementation of an innovative curriculum, the time span needed tends to be measured in decades rather than in single years.

**This section draws on some unpublished work by John Wirt of The [illegible] Corporation, who kindly made it available to me.

- What are the objectives in each field and subfield?
- What styles of R&D (research, policy analysis, development/design, experimentation, evaluation) are being supported?
- Is the mix of styles appropriate to the objectives in each field?
- What is the quality of the performers being supported?
- What is the mix of performers?
- Is this mix appropriate to the objectives in each field?
- What contributions are being made to the knowledge base in each field?

Of the four different aspects of evaluating a mission-oriented R&D agency considered in this paper, the methodology for quality assessment is probably the best developed. Criteria for choosing fields and subfields have been established in such recent examinations as *Priorities for Space Research, 1971-1980* (1971), and *Physics in Perspective: Recommendations and Program Emphases* (1972). They generally include intrinsic, extrinsic, and collateral criteria. Intrinsic criteria measure inherent quality: "ripeness" of the field, availability of new techniques, recent discoveries that have posed new significant questions, prospects of opening up further areas of inquiry, propensity of the field to attract able researchers. Extrinsic criteria are concerned with contributions to other fields, to policy, to progress in practical applications to social goals. Collateral criteria deal with coherence of R&D activities within the agency, coherence in the context of overall activities in a field, consistency and reliability of results, appropriate balance of R&D styles.

Relatively well-established practices for applying these criteria to R&D programs exist. The raw material for evaluation is aggregate information on proposals, performers, progress reports, final reports, and review information on R&D projects supported. The method usually involves some form of peer-group review, often through specially convened panels, sometimes via a two-tier system consisting of specialist subpanels and an overview panel.* For the NIE as a whole, the Advisory

*The NIE has itself applied this type of assessment to the programs the Regional Laboratories and R&D Centers.

Council could function as the overview panel, although completely independent reviews should also take place to assure objectivity and credibility. In actual application, the procedure often resembles an adversary model in "which there are claims and counterclaims, arguments and counter-arguments, and each side advanced by an advocate who attempts to make the best possible case for his position [*sic*]."* Guttentag (1971) points out that this is a quite appropriate model for evaluating programs in actual social contexts, but it requires advocates deeply versed in the case -- and that is, of course, the catch for the NIE. Criteria will be applied differently by educational researchers as opposed to behavioral scientists, by economists as opposed to computer scientists, and so on. Whose advice should the NIE elicit to improve its programs? Whose opinion will it have to take into account, whether the assessment was elicited or not? Whom will the other groups, particularly those that control NIE's resources, listen to on questions of quality of its programs? The responses that the NIE formulates to these questions -- the relationships that it chooses to build, as in the case of the NIH -- will play a major role in its development, perhaps even its survival. And there is no substitute for staff competence and judgment in shaping these relationships.

An auxiliary mechanism coming into more frequent use to help structure technical evaluations of R&D programs is the commissioning of state-of-the-art reviews. Such reviews can be considered the research component of evaluating an R&D program, for they investigate ("gather evidence" on) the content of each field, its strengths and weaknesses, the record of progress. While panel evaluation should take place periodically, perhaps annually or biennially, state-of-the-art reviews ought to be carried on

---

*Guttentag (1971) quoting from an unpublished paper by M. Levine. Cain and Hollister (1969) also discuss evaluation as "an attempt to raise the standards of what is admissible as evidence in a decision process that is inherently likely to remain adversary in nature. Higher standards of evaluation will lessen the role of 'hearsay' testimony in the decision process, but they are not meant to provide a hard and fast decision rule in and of themselves...if standards for the acceptance of evaluation results are viewed in terms of the 'rules of evidence' analogy, we can begin to move toward the judicious mix of rigor and pragmatism that is so badly needed in evaluation analysis."

continually, field by field, and each field should be reviewed every few years. The NIE itself should sponsor the reviews as one of its research activities. Staffing (whether by in-house researchers, outside consultants and grantees, or a combination) is critical; the individuals charged with conducting reviews of a field must be able to commit enough time, be technical experts, have wide-ranging interests in order to avoid biases, and have highly developed critical and analytical faculties; the actual authors of resulting papers must also have lucid writing styles. Some questions that can be addressed by state-of-the-art reviews include:*

- What are the principal findings and results in the field being examined? In each subfield?
- What are the principal non-findings?
- What is the technical reliability of results? Are achievements repeatable? Consistent?
- What R&D problems are currently receiving the most attention? Why?
- What problems are important but not being worked on? Why not?
- What are the principal impediments to more rapid technical progress? Lack of data? Lack of theory? Lack of facilities or appropriate settings? Lack of instrumentation?
- Who are the major contributors to the field?

Findings developed by such reviews of fields and subfields should be published in professional journals and other media, for, if well done, they can provide milestones not only for NIE planning and evaluation, but for the wider community of researchers, professionals, and interested laymen. This would be a useful service for the NIE to perform, quite apart from supplying input for assessing the progress being made through its support of R&D.

Choice of Questions or Problems

The NIE might receive high marks for the technical quality of the R&D it supports, and yet be condemned on the basis of not coming to

*A detailed list, much of which is applicable, can be found in ndix C of *Physics in Perspective* (1972).

grips with the really important problems of education as commanded by its charter. One observer (Timpane, 1970, p. 565) comes to rather pessimistic conclusions as to the ability of R&D to deal with priority problems: "If a problem area proposed for experimentation is unpopular and/or unimportant, experimentation should not and/or will not be done; but if it is popular and important, action will not wait for experimentation.... In the competition for funds, short-term attention to action demands is likely to offer greater promise of political reward than research." In the face of such political exigencies, will the NIE be able to address important problems? And in the absence of clear consensual goals and sufficient understanding of problems to allow parceling out the educational components, how can one assess whether the problems the NIE does select to work on are the right ones? These questions will not be satisfied by an evaluation of the kind just described, which is concerned with research and technical problems, for clearly the word "problems" in the legislation and in the view of most of the NIE's overseers (excepting only some components of the R&D community) carries a quite different meaning. It refers to the publicly perceived educational problems, for example, the failure of the schools to teach reading, and not uncommonly even includes non-educational problems thought to be solvable through education, such as drug abuse or environmental deterioration. Insofar as resource allocations are made on the basis of assessing various programs against each other, the matter of problem choice is at least as crucial as quality. However, despite a sizable body of literature on decisionmaking, there are no sure-fire methods for selecting problems or for deciding whether those of highest priority are being addressed. But again, asking some specific questions will help clarify what information is relevant to such an assessment.

- Who thinks the problem or question is important?
- Why is it considered important?
- Are major policy or funding initiatives regarding the problem anticipated?
- How many individuals does the problem affect?
- What is the nature of the injury or disservice done to the individual or group affected?

- What are the overall societal effects of this injury?

The questions themselves imply some methodologies for developing the needed information: opinion surveys, including the opinions of affected populations; collecting opinions of leaders; recording policy as expressed in major federal and state legislation, proposed and enacted; analysis of data from the census, schools, courts, and other sources; statistical and case studies of affected populations. An improved knowledge base should make possible some judgments on relative importance of problems, though the nature of the judgment will still be influenced by the perspectives of the evaluating groups. But problem importance is only one consideration in choice of problems; the second is concerned with feasibility. A problem may be very important, but knowledge and resource constraints may make it a poor choice for the NIE's R&D program. Any major program initiative (except field-initiated basic research) should be subjected to an examination addressing the following questions:

- What are the components of the problem or question that are appropriately addressed through R&D?
- What components of the problem can be ameliorated through educational intervention?
- Has enough R&D progress been made to make further progress likely? To allow needed development and design of alternative educational systems?
- Is there a base of exemplary practice to serve as focus for research? For development? For directed experimentation?
- Are competent people available and interested in working on the problem?
- What other agencies, federal or non-federal, are working on the problem?
- Are the available financial resources appropriate to the likely effort needed? On the part of the NIE? On the part of other agencies that could be engaged to cooperate on the problem?
- If directions for solutions are found or educational alternatives developed, will they be implementable?

Although these ought to be planning questions, they are also relevant to evaluation, particularly if it is to be useful for improving the NIE's R&D strategies. Answers will not always be available at the time programs are initiated; therefore, it should be part of program operations to develop them as a program proceeds. It is the evaluator's function to assess the validity of problem choice in the light of planning rationale and of progress being made toward improved understanding and design of ameliorating interventions.

The two components of problem choice require very different types of information: the first -- on problem importance -- should be as broadly elicited as possible; the second -- on feasibility -- depends on expert knowledge of the state-of-the-action concerning a problem. Review papers similar to those recommended for assessing the state-of-the-art of a field are appropriate here, but with a different focus: to collect and synthesize information on all activities attempting to develop solutions for the problem. Again, quite apart from their importance in making evaluation of problem choice a more rational activity, state-of-the-action reviews would be an invaluable source of information for researchers and decisionmakers of all sorts. Evaluating the NIE's decisions as to the feasibility of R&D approaches to various problems established as important could be carried out by similar panel methods as suggested for the quality evaluation. One product of such an assessment could be suggested changes of problem choice.

Effectiveness of Program Output

Assuming technical quality and appropriate choice of important problems, what are suitable measures for assessing the results of the NIE investment in R&D in education?

First, any evaluation must clearly focus on the fact that the NIE's mission is research and development,* not large-scale action programs, the more common subject of evaluation. Therefore, relevant

*This does not preclude development of strategies and tools for dissemination and implementation of the results of its R&D efforts; it does preclude wholesale funding of adoptions of innovations.

criteria will measure progress in three areas: (1) contributions to the knowledge base needed to deal with educational problems, (2) contributions to policies that further educational objectives, and (3) development and testing of products and processes designed to improve delivery of education. Aggregate measures such as national or city-wide reading scores, dropout statistics, or distribution of different population groups in institutions of higher education are not appropriate in the first few years of the NIE's existence; they may become so provided that NIE-initiated policies or educational interventions become widely implemented, and that the phenomenon being measured is to a substantial degree subject to modification through education.

Second, while some objective criteria are available for assessing program effectivness in the three areas noted, efforts to develop benefit-cost ratios for purposes of resource allocations are not likely to be any more productive for the NIE's programs than for other R&D support activities. R&D is a risky activity, as Rivlin (1971, p. 51) comments by way of illustration: "The costs of finding a cure for cancer are inherently uncertain; they depend on unforeseeable outcomes of basic and applied research." Nor will it be either possible or desirable to project benefits solely in economic terms for most contributions to knowledge about educational problems or to educational interventions. Though there may be some specific initiatives for which this is appropriate, for example, efforts to increase educational programs designed to make migrant rural families economically viable, in general, there will be few instances in which enough empirical data are available to allow the application of cost-benefit analysis.* However, comparative operational costs of educational alternatives developed under NIE auspices are a legitimate evaluation criterion, as noted below.

Assessing progress in the knowledge base needed to resolve questions or problems in education is closely related to the quality assessment. The state-of-the-art reviews suggested there, if the same field is re-examined at periodic intervals, will serve as evidence of contributions

*See Rossi, Chapter 2, in Rossi and Williams (1972).

to crucial data, theory, and conceptual understanding ascribable to NIE-supported activities. In addition, use indicators are appropriate:

- What is the quality and quantity of literature resulting from NIE support?
- How frequently are findings cited in later work in the field? By researchers not receiving NIE support?
- Are advances in the understanding of a specific problem or question clearly discernible over a two-year period? A five-year period?
- Are the findings useful to the NIE's own programs? What is the level of direct application within the NIE?
- Are the findings being used by other institutions, federal and local? To what extent?

Insofar as these criteria involve judgments of quality of the R&D output, peer-group review is again an appropriate method; amount of usage should, however, be established independently through such means as citation indexes and can in itself help in quality assessment.

Questions to be asked in evaluating contributions to policy formulation also revolve around usage, but the documentation is likely to be much more difficult, since the basis for most policy decisions is usually multifaceted and not often fully explicated. The user clientele, instead of professionals in various disciplines and in education, will be the components of the executive branch of the federal government concerned with educational policy, Congressional committees dealing with education, state and local education agencies, and educational systems and institutions. The documents to be examined, rather than the scientific and professional literature, should include sponsored and enacted legislation at all levels of government, policy statements by educational decisionmakers from the President to local superintendents and editorial and similar non-professional literature intended to influence public policy. Unlike the somewhat similar search to establish problem importance, the required examination should -- if possible -- be carried out independent of the NIE funding, since its objectivity is [illegible]ly to be questioned otherwise.

It may be useful, however, for the NIE to sponsor retrospective studies like TRACES* and Project Hindsight** some five or ten years hence, to analyze use of the NIE output both in the knowledge base and policy formulation areas. The NIH, for example, is currently engaged in some examinations tracing the effects of their past efforts. The purpose of such studies should be to enhance program effectiveness rather than influence resource allocations through justification of past support, therefore -- unlike the examples just given -- the studies should also note instances of failure, particularly in the policy area, for example, where directions were taken in deliberate contravention to what appeared to be indicated in NIE-developed information, or where such information was ignored because of gaps in communication.

Evaluation of success in developing and testing improved products and alternative systems for education can build on a considerable history of such assessment. Educational innovations may consist of designing components that will help make existing systems work better, such as new curriculum programs, information systems accommodating tracking of individualized instruction, performance-based testing to credit experience-based learning; or it may put a number of components together in such a way that an entire new system results. Each of these should be assessed separately, for it is quite possible that some components may prove successful apart from the system for which they were designed. Indicators of success should be based on operational objectives; decisions as to implementation are also relevant criteria, but use criteria should be applied only after broad-scale implementation has actually been attempted. Again, retrospective studies may help highlight the sources of success and failure in development, testing, and implementation. Appropriate questions are:

- Has the developed product or system had the effect originally aimed for, as documented by testing?

---

*_Technology in Retrospect and Critical Events in Science_ (1968), prepared by IIT Research Institute.

**Office of the Director of Defense Research and Engineering (1969).

- For what populations, in what settings, does it have the desired effect?
- In what ways, desired and undesired, is the performance and behavior of participants changed by the educational innovation?
- Is adequate information being provided on how to install the innovation? On costs? On training prerequisites for staff? On special requirements (e.g., equipment, space, management arrangements)?
- Have the NIE innovations led to implementation funding by social action agencies such as OE or OEO?
- Are local school systems or other educational institutions investing their own funds in adopting NIE-sponsored products?
- What are the barriers to implementation?

If implementation actually does take place, additional criteria can be applied, such as number of users or sites, effectiveness of replication (is the product or process still recognizable after it is out of the hands of the original developers?), test scores and other performance indicators, distribution of use among target populations, and unintended side effects.

Assessment of the products of development and experimentation can in itself become a major R&D activity. Planning for appropriate evaluation should be part of the program development process, as emphasized by Crawford (1972) in his recent study of the impact of educational R&D products, but ordinarily the level of evaluation effort will be minimal at program inception and become greater as products come into use. Putting the matter another way, development of truly innovative educational curricula or practices is complex and time-consuming, impact even slower, therefore evaluation of development and experimentation must have an adequate time frame. Considering the high expectation for visible successes, however, which is likely to enter any outside evaluation of effectiveness, the NIE would be well-advised to invest in some short-term projects that could yield rapid payoff, for example, implementation manuals for adopting improved practices that have already been tested through natural experimentation or through demonstration funded other agencies.

Distribution of Funds and Second-Order Benefits

This dimension of evaluation is quite different in character from the other three: rather than being concerned with outcome, it focuses on process. In some sense, satisfactory performance along the other three dimensions should make this issue superfluous, but it must be considered separately because of its special interest to Congress. Apart from concerns with substantive contribution and allocation of educational R&D resources to yield optimal results, Congress attaches importance to the "fairness" by which R&D funds, prestige, and access to more subtle benefits (e.g., being part of an "in-group") are distributed. Questions of greatest interest usually involve geographic distribution of funds (and also of eventual benefits to practitioners and consumers), widely accessible opportunity to compete for funding (e.g., dislike of sole-source contracts), and openness of management procedures (e.g., 5 U.S.C. 522, *The Freedom of Information Act*). To some degree, the performer communities will share these interests, though their notions of fair distribution criteria will not match those of Congressional or departmental watchdogs. Williams (1971, p. 135) points out that public agencies have traditionally been sensitive to such questions and will attempt to establish a record of accountability and fiscal prudence, sometimes to the point where "administrative purity may become a public manager's greatest concern."

There will never be an adequate response to distributional questions, however, precisely because "fairness" is perceived differently by different overseers and clients, and because any concept of fairness is to some degree in conflict with quality and effectiveness criteria in the allocation of R&D support. The NIE must put quality and effectiveness first, but it should be open to judgment on the availability of information about any of its practices and rationales for them. This implies the existence of an effective management information system that permits quick access to data on number and origins of proposals; data on location and types of performers working on current grants and contracts; agency guidelines on requests for proposals, proposal evaluation, and property rights and licensing procedures for products developed with NIE support; monitoring procedures, and so forth. As important as forthright and prompt response to questions on the *what* of [illegible]l practice is the *why*. Therefore, any evaluation should consider

the validity of the reasons for various management procedures, the clarity with which procedures are explained to all concerned parties, and the effects of the procedures. Evaluation should also consider to what extent practices are designed ahead of time in pursuit of deliberate strategies for R&D management instead of representing the accretion of ad hoc decisions and responses to hostile criticisms that characterizes many government programs.

Assessing R&D Capability

The reader will note that the evaluation criteria and methods discussed so far address in a variety of ways the first three missions of the NIE as delineated in the legislation, but few are directly applicable to the fourth, "building an effective educational research and development system." (Although distributional criteria are sometimes made to serve this purpose, they are no more applicable for gauging the effectiveness of educational R&D than they are for gauging the effectiveness of R&D to develop alternative energy sources, despite the great differences in the spread of expertise in the two areas.) This omission is quite deliberate and derives from appraising past attempts at building R&D capability in vacuo, that is, without an existing core of quality R&D, before important problems amenable to R&D approaches are defined, and in the absence of any strategy for assessing the effectiveness of the R&D system's output.

If the NIE can perform successfully in regard to its first three missions, building R&D capability only as specifically required for program initiatives in regard to those missions, then it will indeed be developing an effective educational R&D system, and this will be evidenced through evaluation addressing the substantive missions. Criteria solely concerned with the R&D system itself, e.g., number of educational researchers trained, number of institutions active in educational research, number of new performers, are, in my opinion, not only irrelevant but misleading, for they may raise unwarranted expectations of performance. Such indicators will not be needed to assess the effectiveness of an R&D system that produces the substantive results sought in the NIE's author- legislation regarding problem-solving in education, advancing

its practice, and strengthening scientific and technological foundations; nor will they convince in the absence of substantive results.

III. THE USES OF EVALUATION

In considering the various ways in which the NIE should -- and will -- be evaluated, one must ask two further questions: (1) How useful will any evaluation be? and (2) How will evaluation results be used? While the second depends in part on the first, it also depends on political considerations that need to be examined separately from usefulness, for evaluation "cannot (and should not) replace politics, but it can, over time, facilitate better political decisions" (Williams and Evans, 1969, p. 130).

Usefulness of Evaluation

Any evaluation, to be useful for decisionmaking, must have three characteristics: it must be competent; it must be relevant; and it must be honest. Unfortunately, particularly where evaluation is to provide feedback for improving an agency's R&D strategies, these aims may be in conflict, as has been noted by Glennan (1969).

I have suggested several types of studies that need to be carried on fairly continuously in order to provide a substantive information base for evaluation and increase its caliber. This background work is unlikely to get done on a systematic basis unless the NIE itself sponsors a good portion of it. "Unless legislation or agency policy specifically earmarks funds, evaluation staffs will not be assembled nor the evaluation job done. Only when a flow of resources exists will a formal responsibility to evaluate be translated into significant evaluation activities" (Wholey, et al., 1971, p. 77). Thus, to obtain competent evaluation, agency commitment is necessary.

Wholey also points out that spending program funds on evaluation (often resisted by program managers who may view it as a threat) is justified if program decisions are likely to be influenced by evaluation. Relevance to decisionmaking, particularly within the agency, again requires agency involvement, as has been emphasized by nearly everyone who has [illegible]ned the field, including several of the authors already cited. But

both competence and honesty require objectivity, and that implies that evaluation should be carried out as an independent activity by outside experts. Perhaps the Advisory Council could play the role of sympathetic but impartial judge, but this precludes its functioning as a knowledgeable advocate of educational R&D, another possible role for the Council. In any case, no matter how the Council defines its functions, its credibility with outsiders as objective assessors of the NIE's performance will not be high, raising the old question: *Quis custodiet ipsos custodes?*

For the NIE's own needs, a possible resolution of the quandary is to emphasize competence and relevance in its self-initiated evaluations. To ensure these and the maximum attainable degree of honesty, a threefold strategy might be used in which the NIE Director and Advisory Council define the purpose of the evaluation, and the NIE funds the necessary background studies, but the actual evaluation procedures are carried out as much as possible by outsiders. The aim would be to provide maximum feedback for the NIE; however, a second purpose might also be served: if the NIE succeeds in obtaining competent evaluations based on relevant information for its own needs, these evaluations may find their way into the assessments generated by independent overseers and critics inside and outside government. It is to be hoped that such an information flow will take place so that completely independent evaluations can take advantage of the evaluative information base established by the NIE, and the NIE in its turn will welcome and use independent appraisals.

Using Evaluation Results

Let us assume for the present that such a climate for using evaluation results will actually exist. How could the results be used? There are three ways in which an agency or its overseers can attempt to introduce improvements based on evaluation feedback: allocating resources differently (both as to overall agency budget and internally, among the agency's programs), changing the management procedures, and reorganization. The four dimensions suggested for evaluation bear directly on resource allocation and on management procedures; changes in organization will usually be a consequence of changed resources and management. For example, an assessment of the technical quality of the R&D, if it includes

the suggested state-of-the-art and peer reviews, will uncover which fields are being overfunded and which are being neglected, in view of their potential contribution to the NIE's missions. Thus, priority judgments become feasible that are independent of proposal or other client pressure and less subject to proportional in(de)crementalism, the usual criteria for budget allocations. Assessments of problem choice, based on the subjective and objective criteria discussed for problem importance and on state-of-the-action-reviews, will also be useful in formulating priorities for budget allocations, for the NIE as a whole and for individual programs. The recent assessments of physics and space research already referred to have, in fact, been able to incorporate priority judgments based on alternative budgets and quantitative scoring. The assessment of effectiveness of output may lead to such suggested changes in management strategies as altering the emphasis on different R&D styles (e.g., less basic research, more development), changing the degree of directiveness and program control, designing new ways of soliciting proposals, changing proposal evaluation mechanisms, and adjusting monitoring procedures. Clearly, quality and problem choice assessments should also feed into the consideration of what management changes might be needed to improve performance. The implications for management of distribution questions have already been discussed.

If suggested changes in resource allocation or management procedures are substantial, their implementation may require changes in agency organization. Depending on the degree of reorganization needed, a separate assessment (perhaps two, one done by an inside and one by an outside group) may be useful to determine the most effective organization for administering the new budget and management procedures.

Application of evaluation results requires that:*

- o New policy directions are articulated clearly.
- o The agency is in a position to institute the changes.
- o Staff are capable of carrying them out.
- o Client groups are willing to adjust.

---

*See Williams (1971), Chapter 8.

The last three conditions are more likely to be met when "changes are modest and take place within the context of a particular ideology, operating primarily to improve efficiency.... These are changes that sometimes can be made by administrative fiat without necessarily arousing professional opposition.... [But] change in policy and agency ideology... could be experienced as 'revolutionary' and threatening by many of the existing staff [and clients] and therefore would likely be opposed or subverted. Such major changes might only become acceptable when an agency experienced a crisis or a keenly felt need to re-examine existing practices...extraordinary efforts on the part of leadership, perhaps including the introduction of new personnel, might be necessary" (Glaser and Ross, 1971, p. 54). In the end, whether any changes actually take place as a result of evaluation, whether the status quo is preserved despite indicated directions for improvement or whether changes take place independent of evaluation results will depend to a large extent on the motives of those individuals or groups responsible for generating the evaluations. The motivation is not often truth for its own sake; as Levine and Williams (1971, p. 31) say: "Ordinarily, however, decisionmakers [or those who wish to influence them] have preconceptions about answers to the questions addressed by an evaluation.... A decisionmaker with strong a priori views...will be a good customer for evaluation only when it supports these views." Further, no evaluation will be so free from flaws that it cannot be used or attacked to serve a particular group's purpose.* Only commitment at top management levels to base policy (where possible) on evidence supplied by evaluation results and to implement suggested changes will make evaluation a useful activity.

Besides attempting to ensure the competency, relevance, honesty, and usefulness of the evaluations and evaluation components that it sponsors itself, can the NIE affect in any way the climate in which it will be evaluated?

---

*Williams (1971, p. 123) states this as "the iron law of absolute evaluation flaws.... *The absolute methodological and logistical deficiencies in any evaluation make political infighting a near certainty when evaluation results threaten a popular program.* In short, 'questionable evaluation practices' can always be attacked on methodological grounds for political and bureaucratic purposes" [italics in original].

I believe it can, through assuring positive results of an evaluation that I have not as yet discussed, but that is probably the most important of all: the reactions to the day-by-day signals broadcast by the management and staff of the NIE in all its operations. Whether dealing with prospective performers and their institutions, with its official overseers in the legislative and executive branches, with education professionals or the consumers of education, or with the press and other media, the NIE will be subject to covert and continuing appraisal. Through their words and actions, the staff will project an image of competence or incompetence; of judgment and taste or mediocrity; of a dynamic and flexible enterprise likely to accomplish something, or another manifestation of government bureaucracy. No matter what the formal evaluation mechanisms set up by the NIE itself or by others to evaluate its performance, they will be permeated by the agency's image as created by the staff. There is no more important concern for the NIE, for its ability to carry out its missions and any judgment on its worth will ultimately depend on it.

BIBLIOGRAPHY

Cain, Glen G., and Robinson G. Hollister, "The Methodology of Evaluating Social Action Programs," in *Discussion Papers*, Institute for Research on Poverty, University of Wisconsin, 1969.

Cohen, David K., *The National Institute of Education: What Can Be Expected?*, National Institute of Education, Planning Unit, Report No. P106, Washington, D.C., September 1972.

Crawford, J. J., et al., *Evaluation of the Impact of Educational Research and Development Products*, American Institute for Research in the Behavioral Sciences, Pittsburgh, 1972.

Education Amendments of 1972, 92d Congress, 2d Session, House of Representatives, U.S. Government Printing Office, Washington, D.C., Report No. 92-1085, May 1972.

Fox, David J., "Issues in Evaluating Programs for Disadvantaged Children," *Urban Review*, Vol. 2, No. 3, December 1967, pp. 5, 7, 9.

Glaser, E. M., and H. L. Ross, *Increasing the Utilization of Applied Research Results*, Human Interaction Research Institute, Los Angeles, March 1971.

Glennan, Thomas K. Jr., *Evaluating Federal Manpower Programs: Notes and Observations*, The Rand Corporation, RM-5743-OEO, September 1969.

Guttentag, Marcia, "Models and Methods in Evaluation Research," *Journal of Theory and Social Behavior*, Vol. 10, No. 1, April 1971, pp. 75-95.

Levine, R. A., and A. P. Williams, Jr., *Making Evaluation Effective: A Guide*, The Rand Corporation, R-788-HEW/CMU, May 1971.

Light, Richard J., and Paul V. Smith, "Choosing a Future: Strategies for Designing and Evaluating New Programs," *Harvard Educational Review*, Vol. 40, No. 1, Winter 1970, pp. 1-28.

Project Hindsight Final Report, Office of the Director of Defense Research and Engineering, Washington, D.C., 1969.

Physics in Perspective: Recommendations and Program Emphases (Bromley Report), National Academy of Sciences, Washington, D.C., 1972 (other volumes forthcoming).

Priorities for Space Research, 1971-1980 (2d printing), National Academy of Sciences, December 1971.

Rivlin, Alice M., *Systematic Thinking for Social Action*, The Brookings Institution, Washington, D.C., 1971.

Rossi, Peter H., and Walter Williams (eds.), *Evaluating Social Programs*, Seminar Press, New York and London, 1972.

Rothschild, Lord, "The Organization and Management of Government R. and D.," in *A Framework for Government Research and Development*, Her Majesty's Stationery Office, London, 1972.

Smith, Marshall S., and Joan S. Bissell, "Report Analysis: The Impact of Head Start," *Harvard Educational Review*, Vol. 40, No. 1, Winter 1970, pp. 51-104.

Technology in Retrospect and Critical Events in Science, Illinois Institute of Technology Research Institute, Vol. 1, December 15, 1968.

Timpane, Michael P., "Educational Experimentation in National Social Policy," *Harvard Educational Review*, Vol. 40, No. 4, November 1970.

Wholey, Joseph S., et al., *Federal Evaluation Policy*, The Urban Institute, Washington, D.C., 1971.

Williams, Walter, *Social Policy Research and Analysis: The Experience in the Federal Social Agencies*, American Elsevier Publishing Company, Inc., New York, 1971.

Williams, Walter, and John W. Evans, "The Politics of Evaluation: The Case of Head Start," *The Annals of the American Academy of Political and Social Science*, Vol. 385, September 1969.

Comments on Raizen Paper by Alice M. Rivlin

Senta Raizen has written a perceptive, realistic paper about evaluation of the National Institute of Education (NIE). It seems to me that her "dimensions of evaluation" are the right ones and that she has struck an appropriate balance between under- and over-selling the importance and the feasibility of evaluating research and development in education. I have no major quarrels with her observations or her emphasis, but would like to offer a few additional reactions.

Let me start with some skepticism about a standard scapegoat of education researchers: The alleged absence of a consensus on goals in education. One of the established rituals of the education research community is beating of the breast and moaning about how hard it is to evaluate research in education because no one knows that the objectives of education are. Part of the ritual is a statement about how much easier everything is in some other research area where the goals are clearer. Raizen picks health research for this comparison, alleging that it is more difficult for NIE to do its job than for the National Institutes of Health: "There are clear consensus goals for research and development in health: curing cancer, reducing the incidence of dental caries, eliminating strokes and heart disease. While the choice of strategies and resource allocations for research and development to attain these goals may often be difficult, at least they are undisputed social goods and clearly perceived." (p.2)

But I submit that there are plenty of specific goals in education analogous to curing diseases. No one disputes that it would be desirable to "improve reading comprehension" or "increase computational skills in mathematics." these educational skills are desirable and measureable and they bear the same relation to education as curing diseases bears to health.

"Health" is just as difficult to define as "education." Health presumably involves vigor, stamina, strength, resistance to infection, and a great many other things besides mere absence of disease. It is a multi-dimensional state and there is certainly room for dispute about what the dimensions should be and what weights should be attached to them. Curing disease is certainly desirable but it is not sufficient for attaining health.

Moreover, even if the cure of a specific disease is accepted as an overriding goal, this goal is not helpful in allocating basic research money in the bio-medical sciences. The basic need is to find out how the body works and how it interacts with the rest of the environment. There may be a long non-obvious path between such basic understanding and any identifiable health goal. In education the need is to find out how people learn and interact with their environment. It may also take a long time before such basic understanding can be translated into skill improvements or other educational goals.

Raizen goes on to note that the problem with proximate goals in education is that "they are in competititon with other highly valued educational goals for resources, or they cannot be attained solely, or even primarily, through education." But again these statements do not distinguish proximate goals in education from those in health.

Specific education goals are not "in competition" in the sense that a person with basic skills is less likely to get a job or someone who can read well is less likely to be able to do mathematics. They are only in competition in the sense that resources devoted to research and development on one specific skill cannot be devoted to research and development on another. But that kind of competition occurs in health and every other kind of endeavor. Resources specifically devoted to curing cancer are not also available to programs to reduce dental caries.

Moreover, whatever good health is, it surely cannot be attained by

medical care alone; just as education goals cannot be attained by formal education alone. Indeed, evidence seems to be mounting that diet, lifestyle, environment, and genetics are far more important to an individual's health than the availability of medical care.

Making more effort to define specific goals in education certainly will not do any harm, but I doubt that it will do much good either. The real problem is not that we lack goals but that education research has not told us much about effective ways to reach even widely accepted goals. NIE is supposed to change this situation.

As Raizen points out, many different kinds of evaluation of NIE will be appropriate and will take place over the years. The most important, however, will be the informal judgment exercised by the Executive, the Congress, and the public as to whether NIE is worthy of continued and expanding support. Positive evaluation will be reflected in a growing NIE budget.

This crucial type of evaluation will probably not be governed by the average qulaity of NIE's overall effort. The average quality could be quite poor and most projects could be failures, but NIE would still be given positive marks provided it could show a few spectacular successes. NIE needs a Salk vaccine -- something that really works, that solves a problem everyone knows about. The problem need not be important. Polio, after all, was never an important disease. It was dramatic because it was a killer and crippler of children. But even in epidemics the incidence was extremely low.

In education, as in health research, a few dramatic wins will make up for a lot of failures and more importantly, for slow undramatic progress on hard problems. One or two visible successes will demonstrate that success is possible, and give the public the feeling that education research is more than a luxury; it has potential for doing good. Indeed, if I were to draw some distinctions between education and health research I would point -- not to

distinctions that Raizen has made -- but to the marked difference in public attitude toward the two fields. The Congress and the public stand in awe of medical researchers. They believe that the men in the white coats with their rats and their test tubes know what they are doing and will, in the end, produce something useful. They have no such confidence in the education researcher.

The biggest question facing NIE will be what proportion of its resources to devote to basic long-term research and what proportion to research with immediate relevance to practical problems. In terms of maximizing short-run improvements in the health of the nation NIH has been a colossal failure. Resources devoted to improving health education, health delivery systems and the effectiveness of current clinical practice might well have brought major improvements in health status. Instead, NIH has devoted most of its resources to basic research in human biology, biochemistry, and genetics behind a facade of relevance created by the naming of institutes after particular diseases. In the last couple of years the balance has shifted -- basic research has given way to more immediate attempts to find disease cures -- a change which many scientists believe is disastrously short-sighted.

NIE will have no choice; it will have to be more current and relevant than NIH in order to survive. At least in the short-run, it will have to concentrate a substantial portion of its resources on research that stands some chance of improving American education in visible ways. Once NIE is established, however, one might hope that the balance would shift toward basic research in the learning process, that increasing amounts of resources would be devoted to finding out, for example, how the brain works and how people differ in the way they perceive and learn.

I would offer two suggestions to the policy-makers at NIE. First, do not waste much time at the beginning defining or refining the goals of education. Pick out a few proximate goals on which there is a high quality basic

research in learning with no obvious practical payoff.

Second, do not, as some have suggested, put much effort into understanding the current process of innovation in education. Looking at how new ideas have spread in the past is unlikely to be relevant to the future. If education research were producing demonstrably successful innovations these innovations would spread. Educators want, by and large, to do a good job and want people to know they are doing a good job. The hard problem, as President Johnson said, "Is not to do what is right, but to know what is right." The first job of NIE is not to examine the impediments to the spread of useful innovations, but to produce the useful innovations themselves.

General Discussion of Raizen Paper and Rivlin Comments

In discussing the potential criteria to be used in evaluations of the NIE, it was noted that the selection of R & D programs should be made with one eye toward satisfying the different clients and/or masters the NIE is expected to serve. Obviously, different audiences call for different approaches, priorities, and programs. If the NIE is to survive and grow, it must choose carefully the problems it wishes to study. Programs that are widely accepted and which have a high probability of success should have a higher initial priority. This would enable the institute to buy time for longer range programs which lack a ready-made and vocal constituency. Since these latter programs could take up to twenty-five to develop and implement, a balance between short, intermediate, and long term projects is most advisable especially in view of the imminent, and quite possibly hostile, evaluation the NIE will undergo in the near future.

As the last statement implies, the NIE will quite often find itself in the position of having to choose between that which is politically acceptable for survival versus that which is needed to improve the educational system. This could be costly for R & D efforts in important areas, and this suggests that the NIE ought to determine ahead of time what its short term and long term strategies are. But in order to do this, the NIE would have to define better who its clients are and what might satisfy them, and it would have to come to grips with its role, mandates, and limitations. It was not clear if more than intelligent guesswork can be possible here.