INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in “sectioning” the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from “photographs” if essential to the understanding of the dissertation. Silver prints of “photographs” may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106
HAGAN, Danford Lee, 1935-
A COMPARISON OF HUMANISTIC AND SCIENTIFIC
EPISTEMOLOGIES IN EDUCATIONAL RESEARCH
AND EVALUATION.

The Ohio State University, Ph.D., 1973
Education, general

University Microfilms, A XEROX Company, Ann Arbor, Michigan

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED.
A COMPARISON OF HUMANISTIC
AND SCIENTIFIC EPISTEMOLOGIES
IN EDUCATIONAL RESEARCH AND EVALUATION

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of the Ohio State University

By

Danford L. Hagan, B.A., M.M.

* * * * *

The Ohio State University
1973

Reading Committee:

Robert R. Bargar
Ross L. Mooney
Gregory L. Trzebiatowski

Approved By

Advisor
College of Education
ACKNOWLEDGEMENTS

The assistance of my committee members for their intellectual and professional guidance is gratefully acknowledged.

The conceptual approach for the evaluation and communications models in the final chapter was developed with the assistance of my major advisor, Robert R. Bargar.

I am also deeply indebted to two lovely and helpful people Judith Hagan and Charlotte Wiant for helping to bring order into what might have been chaos.
VITA

May 25, 1935 .... Born - Somerton, Ohio

1958 ............. B.A., Ohio Wesleyan University, Delaware Ohio

1965 ............. M.M., Miami University, Oxford, Ohio

1968-1973 ........ Research Associate, College of Education
The Ohio State University, Columbus, Ohio

PUBLICATIONS

"A Survey of Ohio Public School Administrators", with Robert Bargar,
Investigation of Factors Influencing the Training of Educational
Researchers, Bargar, Robert R., et. al., Project No. 3191,
USOE, May, 1970.

"Challenge of Change", In-service "Elevision Programs for Teachers,
Produced by Ohio State University for State of Ohio Department
of Education - Script author and Co-Producer.

FIELDS OF STUDY

Educational Research and Evaluation
Professors, Ross L: Mooney, Robert R. Bargar:, and James K. Duncan

Educational Systems and Technology
Professor Gregory L. Trzebiatowski
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>VITA</td>
<td>iii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Dissertation Purpose</td>
<td></td>
</tr>
<tr>
<td>Communication Disjointedness Between Humanistic and Scientific Discipline</td>
<td></td>
</tr>
<tr>
<td>Conflicts Between Humanistic and Scientific Measurement</td>
<td></td>
</tr>
<tr>
<td>Rationale of the Dissertation</td>
<td></td>
</tr>
<tr>
<td>Dissertation Strategy</td>
<td></td>
</tr>
<tr>
<td>Chapter Outline</td>
<td></td>
</tr>
<tr>
<td>II. HISTORIC EFFECTS OF AN UNFORTUNATE ANALOGY</td>
<td>34</td>
</tr>
<tr>
<td>Introduction</td>
<td></td>
</tr>
<tr>
<td>Statement of the Basic Argument</td>
<td></td>
</tr>
<tr>
<td>In Search of Viable Models</td>
<td></td>
</tr>
<tr>
<td>Models as False Analogies</td>
<td></td>
</tr>
<tr>
<td>III. PURPOSES AND HISTORIC ANTECEDENTS</td>
<td>60</td>
</tr>
<tr>
<td>Introduction</td>
<td></td>
</tr>
<tr>
<td>Toward an Eclectic Approach</td>
<td></td>
</tr>
<tr>
<td>Historic Antecedents</td>
<td></td>
</tr>
<tr>
<td>Summary</td>
<td></td>
</tr>
</tbody>
</table>
IV. DEFINING SCIENTIFIC PARADIGMS

Introduction
Social Science Interpretations
The Hypothetico-Deductive System
Nomothetic Science
Summary

V. CRITICISM OF THE METHODS AND UTILITY OF RESEARCH

Introduction
Criticism from within the Profession
The Controversy of Utility
Criticism of the Nomothetic Concept
Criticisms of Methodology
The Form of the Basic Argument
The Probability Issue
Significance Tests: A Problem of Utility

VI. OPEN AND CLOSED SYSTEMS

Introduction
Definitions of Closed and Open Systems
Information Roles Model
Systems Concepts in Education
Implications of the General Systems Theory in Education

VII. TOWARD A HUMANISTIC FRAME OF REFERENCE

Introduction
Setting Aside the Old Myths
Research and Evaluation as Communication
Critical Differences Between Research and Evaluation

VIII. TOWARD A NEW PARADIGM

Introduction
The Great Issues
A Rational Framework for Analysis and Evaluation of Problems in Education

APPENDIX A

242
A COMPARISON OF HUMANISTIC
AND SCIENTIFIC EPISTEMOLOGIES
IN EDUCATIONAL RESEARCH AND EVALUATION

By

Danford L. Hagan, Ph.D.

The Ohio State University, 1973

Assoc. Professor Robert R. Bargar, Advisor

Abstract

It has become painfully obvious that educators cannot make enlightened decisions, plan effectively for change, or meet the new demands by society for social accountability, without a vastly improved capacity for providing evaluative data for decision making. Unfortunately, the relevance and range of usefulness of the most commonly used assessment techniques — primarily research methods and psycho-metric measurement — have proven to be severely limited in regard to comprehensive decision making.

Research has remained a rather isolated activity, conducted by university professors or foundation staffs, and has only rarely been called upon to assume a dynamic role within the mainstream of education. Evaluation has been similarly isolated, being
primarily associated with the tests and measurement movement. Attempts to evaluate effects of large scale reform or innovation efforts have been frustratingly unsuccessful.

The main contention is that the lack of success of research and evaluation has not been due to weak efforts, but rather, to logical flaws within the methods used. In some respects, these flaws actually create barriers to progress, rather than permitting research and evaluation to fulfill their basic purposes.

The rationale which is developed is based upon a logical contingency analysis between the perceived needs of decision makers for evaluative data, and the actual forms of information that the traditional models are capable of providing. When decision makers really need information upon which to base value judgements, the typical probability estimates and statistical hypothesis tests which the traditional models provide, are of little utility.

It is proposed that researchers and evaluators give more careful attention to the empirical necessity for analyzing the interactions between educational programs and their clients as systems which must have viable communications for system integration and evaluative feedback. Because educational programs are inherently part of the values of social systems, information generated from research and evaluation must reflect upon the humanistic concerns of society.
# FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Communication Processes in Various Disciplines</td>
<td>10-13</td>
</tr>
<tr>
<td>2.</td>
<td>Logical Distinctions Between Disciplines</td>
<td>15</td>
</tr>
<tr>
<td>3.</td>
<td>A Representation of the Processing of Descriptive Data</td>
<td>28</td>
</tr>
<tr>
<td>4.</td>
<td>Havelock Paradigm</td>
<td>64</td>
</tr>
<tr>
<td>5.</td>
<td>Regional Laboratories Link Research and Practice</td>
<td>65</td>
</tr>
<tr>
<td>6.</td>
<td>Philosophical and Historical Convergence of the Scientific Method</td>
<td>102</td>
</tr>
<tr>
<td>7.</td>
<td>Frequency Distribution</td>
<td>136</td>
</tr>
<tr>
<td>8.</td>
<td>Mooney Figure</td>
<td>152</td>
</tr>
<tr>
<td>9.</td>
<td>Roles of Information in Hierarchy of System Models</td>
<td>157</td>
</tr>
<tr>
<td>10.</td>
<td>General Model of Educational Research</td>
<td>166</td>
</tr>
<tr>
<td>11.</td>
<td>Steps in Hypothesis Testing</td>
<td>167</td>
</tr>
<tr>
<td>12.</td>
<td>Hypothetics - Deductive Model as a System of Logic</td>
<td>168</td>
</tr>
<tr>
<td>14.</td>
<td>Berio Model</td>
<td>196</td>
</tr>
<tr>
<td>15.</td>
<td>Merio Model Variation</td>
<td>196</td>
</tr>
<tr>
<td>16.</td>
<td>System Analysis and Evaluation Model</td>
<td>223-224</td>
</tr>
<tr>
<td>17.</td>
<td>Information Unit Model</td>
<td>232</td>
</tr>
<tr>
<td>18.</td>
<td>Model for Program Information System Design</td>
<td>234</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

Communication processes are so pervasive throughout modern life that it may not even be possible to isolate the effects of certain types of communications within complex social environments. Yet modern man must learn to plan and manage social institutions, even as they become larger and more complicated. The quality of life—man's very survival—depends upon our ability to understand the way people interact with each other in organizational settings and the effects of organizations upon the fabric of life.

We see this problem at the global level as nations struggle to coexist peacefully and to overcome the vast discrepancies in comparative living standards.

We see the problem within nations as large corporations and social organizations strive to achieve a balance between the negative effects of commerce—industrial pollution and the dehumanization of workers, and the positive effects—the production of needed goods.

We see the problem within institutions as they attempt to achieve meaningful and viable relationships, both with their clients
and their own employees.

The understanding necessary to move toward solution of these problems depends upon our ability first, to appreciate the humane qualities necessary to sustain vital interpersonal communications; and second, to build upon this quality so that large scale information processes which are necessary for communications within and between complex organizations do not systematically destroy these personal, humane qualities.

Western man in particular frequently takes great pride as he ponders his newly developed skills in creating powerful and complicated information processes. He can transmit telecasts from the moon or around the world simultaneously via satellites. The billing department of a large utilities firm can bill thousands of customers in the space of a few hours. Corporation executives are enabled to monitor world wide operations in a matter of minutes through teletype, or in almost instantaneous communications through facsimile reproduction. Scientists have been able to calculate the movement of matter both at the sub-atomic level, and at the interstellar level with a precision that was undreamed of only a generation ago.

These examples are almost trite in their obviousness. Yet, we cannot overestimate the power of information in the modern world. An industrial empire is only as strong as the communication linkages
between the various components. Scientific theories are only as sound as the information upon which they are based. The cohesiveness and character of a nation are dependent upon the strength and quality of the communications which permit people to share a common understanding. The marketing strength of an industry is dependent not only upon the quality of its products, but also upon the ability of its advertising agency to communicate the desirability of the products to the buying public.

These information triumphs have permitted modern man to gain great understanding and control of the physical universe. Science and technology by definition have developed the power to measure and describe physical events ranging from the minuteness of sub-atomic particles to the vastness of the galaxies. Although many accomplishments of science and technology have been very beneficial to man, particularly in the medical and biological areas, scientific information processes have not necessarily contributed to greater understanding in the psychosocial realm—nor were they expressly designed to do so.

One of the principal themes to be developed in this dissertation is that the primary information processes of science and technology—problem conceptualization, hypothesis testing, research methods, measurement techniques—are in most instances not directly applicable to the social sciences. If education is to contribute positively toward
the attainment of the most humanistic goals of modern society, the fostering of individual growth and the sense of worth and personal responsibility of each child, then the information processes at the personal and interactive level, as well as at the various organizational levels, must be capable of providing dramatic distinctions from those most commonly used in scientific and industrial communications.

In later chapters I will develop further the position that information processes in the psychosocial realm—the world of personal values and interpersonal communications—must reflect the uniquely humane qualities of life as experienced by individuals. There has been a nearly universal tendency in the history of the social sciences—in psychology and sociology in particular—to equate "scientific" information processes with the methodologies of the physical sciences. Scientific inquiry has been taken to mean the uniform description, precise measurement, and experimental control of variables. Efforts which have received the most credence have been those using the greatest degree of experimental control and statistical quantification.

The point that I wish to make here is that information processes in the psychosocial realm and in the world of science and technology though analogous in some respects, must be distinctly different in the manner in which individuals and their characteristics are "measured" and described.
Dissertation Purpose

The central purpose of this dissertation will be to discuss issues related to the conflict between science and humanism and to present a paradigm for the role of research and evaluation as inputs into information systems for decision making in educational contexts. Three different types of information criteria will be discussed: (1) the need for information of sufficient quality upon which to base decisions, (2) the need for information of sufficient scope to provide for modern, comprehensive approaches to program evaluation, such as the system analysis or CIPP models, and (3) the need for information within a humanistic frame of reference which honors the unique qualities of human interaction in personal and social experience.

The issue that I am addressing is not that of the traditional struggle for primacy between the humanists and the mechanists for popular approbation. I have no quarrel with the precision of measurement in the physical sciences, or with the power of communications in the business world, so long as the activities are used as a positive force in society, and do not have detrimental ecological side effects, either for the natural or psychological environment.

At issue are the vast logical and philosophical differences between the information needs of educators, and the methods which are useful in scientific or industrial settings. The quality of inter-
personal communications—the effects of educational programs upon individual values—the search for optimal ways to nurture individual growth and creativity—these are but a few examples of the difficult societal goals for which educators must develop and evaluate programs. Regardless of the power and accomplishments of scientists and businessmen, I am not aware of any scientific methodologies or business techniques which provide an appropriate measurement, or even an adequate frame of reference for dealing directly with these issues.

On the surface, information needs of educators appear to be somewhat analogous with those of science and industry. Educators must assess and validate the worth of programs; they must meet the challenge of achieving efficiency and accountability within large organizational structures. However, I maintain that this apparent similarity of needs should be viewed very cautiously, and that in fact, viable information processes in education must encompass a frame of reference, an historical perspective, and a set of practical assessment procedures which are uniquely and distinctively different from those used in science and industry.

Before going forward with the main body of this dissertation, it should be helpful to take a closer look at the background of the two principle aspects of the dissertation purpose. First, a brief discussion of the setting of the conflict existing between science and humanism,
along with a chart setting forth the main areas of study within those two parts; second, an overview of the conflicts between humanistic and scientific measurement, and the need for changes in the information systems for decision making in education.

**Communication Disjointedness Between Humanistic and Scientific Disciplines**

In later chapters the manner in which educational researchers and evaluators have tended to emulate the methodologies and paradigms of the sciences will be presented in greater detail. At issue is the degree to which the fields of research and evaluation have borrowed the research and measurement constructs of the social sciences—which in turn were adapted from the older, physical and biological areas.

There has been an almost universal tendency within the social sciences to assume a unity of method and purpose which is allegedly applicable across all scientific areas. The purpose of this dissertation is to illustrate that this concept is illusory—that this unity does not exist anywhere within either the philosophical or logical traditions of the various scientific areas. The challenges which face educational research and evaluation are indeed similar to those facing many other disciplines, but only at the level of analogy. Educators do need to create viable information constructs, theories, and hypotheses with
which to build coherent knowledge systems with the capability of explaining and predicting behavioral outcomes, just as the scientists have been able to construct powerful methods for predicting and measurement.

But the constructs and the measurement processes cannot be the same in education as they are for the sciences. The basic purpose of a construct is to provide a conceptual or measurement tool which although somewhat abstract, still has the power to communicate vital information to practitioners within a particular paradigm. The concept of the molecule is a behavioral construct which has had this type of power as a conceptual and communicative tool within the sciences. The construct of mental measurement has served a similar purpose within education, although with less happy results. Within the social sciences in general, the construct of sampling from a larger population has served as a useful conceptual tool.

However, conceptual constructs and measurement constructs must be valid, reliable, and more importantly, serve as means to communicate theories and propositions effectively within professional disciplines or scientific paradigms. The contention here is that the social sciences have uncritically borrowed many of these communicative devices from the older sciences without seriously examining the underlying assumptions. The critical issue is that each professional discipline
each practicing scientific paradigm deals with conceptual and measurement problems which are unique. At the practical, functional level, each area must utilize its own ideographic forms of communication which serve to integrate and unify knowledge.

Ruesch has performed a study of the variety of communication processes which occur within many different areas of study. Figure 1 depicts his analysis of nearly 40 different disciplines or fields of specialization. His contention is that each area by definition has its own unique characteristics and conventions for describing phenomena and communication. As he suggests, some areas stress interpersonal communications which involve all of the uncertain richness of human symboling, while others, such as engineering utilize very precise forms of communication such as computers and various measurement instruments.

As Ruesch points out, the psychological qualities of communications in the social sciences cannot be logically the same as the automata of the mechanistic fields, or the physio-chemical transactions within biological areas. There are simply too many distinct and unique qualitative features of communications within the psycho-social realms for these processes to be equated in any significant way with the information processing within the scientific-technical fields.
## FIGURE 1

COMMUNICATION PROCESSES IN VARIOUS DISCIPLINES

<table>
<thead>
<tr>
<th>Discipline or field</th>
<th>Subject and area of Specialization</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>The technological disciplines</strong></td>
<td></td>
</tr>
<tr>
<td>Cybernetics</td>
<td>Steersmanship and feedback in biological and social systems</td>
</tr>
<tr>
<td>Mathematical theory of communication</td>
<td>Information theory</td>
</tr>
<tr>
<td>Engineering</td>
<td>Computers, automata and control devices</td>
</tr>
<tr>
<td><strong>History and scientific philosophy</strong></td>
<td></td>
</tr>
<tr>
<td>History of language</td>
<td>Development of language over the centuries</td>
</tr>
<tr>
<td>History of instruments of communication</td>
<td>Development of communication technology</td>
</tr>
<tr>
<td>Epistemology of various fields</td>
<td>Assumptions made in scientific procedures</td>
</tr>
<tr>
<td>Unified theory</td>
<td>Communication as general systems theory</td>
</tr>
<tr>
<td><strong>The fine arts</strong></td>
<td></td>
</tr>
<tr>
<td>Painting &amp; sculpture</td>
<td>Expression of inner events through shape, color, movement, texture and sound, resulting in the creation of nonverbal signals and signs to which others respond</td>
</tr>
<tr>
<td>Dance</td>
<td></td>
</tr>
<tr>
<td>Theatre</td>
<td></td>
</tr>
<tr>
<td>Music</td>
<td></td>
</tr>
</tbody>
</table>
### FIGURE 1—Cont.

<table>
<thead>
<tr>
<th>Discipline or field</th>
<th>Subject and area of Specialization</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>The useful arts</strong></td>
<td>Shaping of the material environment embodies the assumptions and conventions of a particular period.</td>
</tr>
<tr>
<td>Architecture</td>
<td>Architectural structures and objects become symbols used in communicative exchange of people with each other and with posterity</td>
</tr>
<tr>
<td>Decorative arts</td>
<td></td>
</tr>
<tr>
<td>Cabinetmaking</td>
<td></td>
</tr>
<tr>
<td>Handcrafts</td>
<td></td>
</tr>
<tr>
<td>Interior decoration</td>
<td></td>
</tr>
<tr>
<td>Fashion design</td>
<td></td>
</tr>
<tr>
<td><strong>The social games</strong></td>
<td>Stylized message exchange by participants who assume certain roles and abide by well-established rules. Behavior then can be analyzed as if it were a game or a play</td>
</tr>
<tr>
<td>Law</td>
<td></td>
</tr>
<tr>
<td>Sports</td>
<td></td>
</tr>
<tr>
<td>Politics</td>
<td></td>
</tr>
<tr>
<td>Special social occasions</td>
<td></td>
</tr>
<tr>
<td><strong>The clinical disciplines</strong></td>
<td>Communicative difficulties inside family</td>
</tr>
<tr>
<td>Child and family psychiatry</td>
<td>Disturbed communication of patient</td>
</tr>
<tr>
<td>Adult psychiatry</td>
<td>Understanding of communications of the patient</td>
</tr>
<tr>
<td>Psychoanalysis and psychotherapy</td>
<td></td>
</tr>
<tr>
<td>Mental hygiene</td>
<td>Community organization</td>
</tr>
<tr>
<td>Neurology and neurosurgery</td>
<td>Pathology of organs of communication</td>
</tr>
<tr>
<td>Discipline or field</td>
<td>Subject and area of Specialization</td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>The psychophysiological disciplines</strong></td>
<td></td>
</tr>
<tr>
<td>Neuropsychology</td>
<td>Correlation of organs and functions of communication. Model construction</td>
</tr>
<tr>
<td>Psychophysiology</td>
<td>Sensory processes and performance</td>
</tr>
<tr>
<td>Psychopharmacology</td>
<td>Influence on communication through drugs</td>
</tr>
<tr>
<td><strong>The psychological disciplines</strong></td>
<td></td>
</tr>
<tr>
<td>General psychology</td>
<td>Perception, transmission, decision-making, memory</td>
</tr>
<tr>
<td>Clinical psychology</td>
<td>Verbal behavior; complex patterns of behavior</td>
</tr>
<tr>
<td>Ethology</td>
<td>Communications of animals</td>
</tr>
<tr>
<td><strong>The social disciplines</strong></td>
<td></td>
</tr>
<tr>
<td>Social psychology</td>
<td>Communication in small groups</td>
</tr>
<tr>
<td>Sociology</td>
<td>Mass communication</td>
</tr>
<tr>
<td><strong>The language disciplines</strong></td>
<td></td>
</tr>
<tr>
<td>Speech and linguistics</td>
<td>Phonetics, language codes</td>
</tr>
<tr>
<td>Signifies</td>
<td>The meaning of signs and signals</td>
</tr>
<tr>
<td>Cultural anthropology</td>
<td>Symbolic systems, verbal and nonverbal</td>
</tr>
<tr>
<td>Discipline or field</td>
<td>Subject and area of Specialization</td>
</tr>
<tr>
<td>---------------------------</td>
<td>-----------------------------------------------------</td>
</tr>
<tr>
<td>The commercial disciplines</td>
<td></td>
</tr>
<tr>
<td>Propaganda and advertising</td>
<td>Influencing people to act</td>
</tr>
<tr>
<td>Political thought reform</td>
<td>Influencing people to believe</td>
</tr>
<tr>
<td>Business organization</td>
<td>Improvement of efficiency of communication</td>
</tr>
</tbody>
</table>
As Ruesch points out, the psychological qualities of communications in the social sciences cannot be logically the same as the automata of the mechanistic fields, or the physico-chemical transactions within biological areas. There are simply too many distinct and unique qualitative features of communications within the psycho-social realms for these processes to be equated in any significant way with the information processing within the scientific-technical fields.

Figure 2 illustrates what I contend is a logical analysis comparing the necessary contingencies between the basic purposes of an area of study, the nature of its paradigmatic rules, its constructs, theories, and laws, and its utility. My contention is that the basic purposes of problem-solving paradigms within the physical, non-living universe (area one in the Figure 2 matrix), have very little utility for the psycho-social realms (cell 4.4 in the matrix). If practitioners within a particular paradigm are devoted to the pursuit of valid and reliable information concerning the phenomena within their purview, then they must by logical necessity develop and pursue those information and communication processes which fulfill their paradigmatic goals.

Following is a brief overview of the terms used to distinguish between the four areas in Figure 2.

**Basic purposes** -- have traditionally been the analysis of physical phenomena in quantitative terms, or in the case of algebra
## LOGICAL DISTINCTIONS BETWEEN DISCIPLINES

<table>
<thead>
<tr>
<th>AREAS OF STUDY</th>
<th>BASIC PURPOSES</th>
<th>NATURE OF PARADIGM</th>
<th>CONSTRUCTS THEORIES, AND LAWS</th>
<th>UTILITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. PHYSICAL, NON-LIVING UNIVERSE</td>
<td>1.1 To describe, categorize, and measure the composition of the physical universe.</td>
<td>1.2 The goals are to develop knowledge systems which are as nomological and as taxonomic as measurement and rationality permit.</td>
<td>1.3 Theories are developed - and laws of cause and effect are discovered - by experiment or other objectively rational forms of hypothesis testing.</td>
<td>1.4 Knowledge systems are highly rational. The achievements of scientific technical knowledge systems need little iteration - witness the exploration of space, nuclear weaponry, satellite communication, etc.</td>
</tr>
<tr>
<td>2. BIOLOGICAL UNIVERSE</td>
<td>2.1 To describe and understand the composition of living organisms in an objective and systematic universe.</td>
<td>2.2 The disciplines strive to be nomological, but are also concerned with difficult to resolve issues of explanation, and evolutionary complexity of life forms.</td>
<td>2.3 Objectivity and scientific precision are hallmarks, but biologists frequently deal with issues and constructs which do not fit highly rational nomothetic theories.</td>
<td>2.4 Achievements in biology, medicine and agriculture have contributed greatly to the improvement of health and living conditions in modern societies.</td>
</tr>
<tr>
<td>3. PSYCHO-SOCIAL REALMS: SCIENTIFIC</td>
<td>3.1 To analyze and describe the objectives, behavioral aspects of human experience.</td>
<td>3.2 The disciplines strive to develop nomothetic theories by studying those behaviors which are objective and quantifiable.</td>
<td>3.3 Disciplines have striven to isolate those forms of personal and social behavior which are amenable to predictive, or cause and effect relationship.</td>
<td>3.4 It is difficult to evaluate the benefits of the experimental and quantitative social sciences. In education and psychology, for example, their utility is subject to much controversy.</td>
</tr>
<tr>
<td>4. PSYCHO-SOCIAL REALMS: HUMANISTIC</td>
<td>4.1 To be increasingly sensitive to and aware of the richness and variety of the widest possible range of human consciousness and cultural experience.</td>
<td>4.2 Humanism has been primarily concerned with personal freedom and creativity. Formal structures are relevant only insofar as they are related to these basic issues.</td>
<td>4.3 Humanists are interested in the arts, humanities, and issues related to social and economic freedom. Personal morality, creativity, and social values resist quantification, and nomothetic theories.</td>
<td>4.4 Many individuals have been strongly committed to improving the quality of the human experience. A revolution appears to be aborning - it remains to be seen how many join the humanistic vanguard.</td>
</tr>
</tbody>
</table>

**FIGURE 2**
and calculus, the manipulation of abstract symbols.

**Nature of paradigm** -- has involved a long and complex tradition of ever more intricate and precise methods of reducing and analyzing physical phenomena into isolable parts at the micro level. At the macro level, science and technology have been able to develop elaborate systems for controlling and measuring huge complexes of phenomena, but the emphasis is upon control and measurement.

**Constructs, Theories and Laws** -- Kuhn has noted that scientists tend to ignore their recent history as soon as they have developed procedures or instruments which suit their purposes better than those in the immediate past. There is a fascinating dichotomy between the historic and philosophic interests of the sciences in comparison to those of the humanities and the arts. In the latter fields a fetish is made of glorifying the past. In science it is almost a matter of pride to live in the future and to demean the past. You will note that I am using history and philosophy in the same context, because it is the philosophy of science to ignore history, in the sense that the past is diligently recorded and revered in other fields of study.

**Utility** -- in the area of actual achievements and benefits, the range is from very concrete examples -- from the physical scientific areas to the abstract and mental examples from the humanistic area. The question remains, does a result have to be physically evident in
order to be worthwhile or beneficial?

These examples are cited to illustrate my contention that there is really very little historical or philosophical justification for the social sciences to look to scientific methods as a means for enlightenment or problem solving in areas of concern which are uniquely human--personal and social values--the arts--the whole cognitive domain.

Conflicts Between Humanistic and Scientific Measurement

Without information of sufficient comprehensiveness and quality it is difficult in education to agree upon where we have been, to accurately describe the terrain upon which we presently stand, or to make rational choices from alternative ways of getting where we would like to go. The central concerns in this dissertation revolve around the problems which educators face as they attempt to apply systematic information processes to these questions.

I feel that there are two basic reasons why the issue of systematic information processes in education is so crucially important. The first reason is viewed from a somewhat negative perspective; the second I consider to be a serious struggle between the growing trend toward mechanistic measurement processes in education--and those who have consistently maintained that the highest humanistic values cannot be properly served by these processes.
The first reason—which I believe is being criticized more and more by humanistically inclined educators—is that social scientists, researchers, and evaluators have tended to emulate the information processes developed by scientists and engineers for the wrong reasons. We have correctly attributed the startling successes of science and industry to the manner in which these fields have been able to develop information with great power to measure, predict, and control events. But we have incorrectly assumed that many of these same measurement techniques, such as parametric statistics, probability laws, formal hypothesis tests, and operations research techniques, could be uncritically applied to educational problems.

In a very real sense, technical and scientific information are, by definition, accurate, valid and reliable. As expressed by Kuhn, within any branch of "normal" science, there are commonly accepted problems, research procedures, and measurement techniques. These almost universally accepted practices establish what Kuhn calls a "paradigm", a concept discussed in depth in Chapter Four. Published reports within a particular paradigm are typically accepted as valid information, because there are systematic screening processes to eliminate reports which appear spurious, and in fact, scientists and research engineers are loathe to publish findings unless they have a high degree of certainty about the accuracy of their information.
Many of the measurement processes which social scientists and educators have borrowed from the physical and biological sciences have been used as though situations and events in social realms could meet the same assumptions and be measured in the same manner. Texts and research articles in such diverse fields as biology, economics, agricultural research, psychology, and educational research use exactly the same research and measurement techniques, as though human behavior, manufacturing methods, and fertilizers could somehow be lumped together into the same types of quantitative phenomena. A considerable portion of later chapters will be devoted to a rationale which creates a clear distinction between types of events which can be expressed validly in quantitative terms, and those in the human psycho-social realm which cannot be so expressed.

The issue to be developed is that events and phenomena within the scientific and technical realms are measurable and describable within the commonly accepted practices of their respective paradigms, but the social sciences make a serious mistake when they become dazzled by the successes of these fields and enamored with their measurement processes. I intend to develop the position that:

1. Only certain types of physical events and processes can be validly measured with these standard paradigmatic processes,
2. The problems within the typical science and engineering paradigms are properly confined to the limits of closed system functions. That is, the inputs, processes, and outputs of the system under question can be measured, predicted, and controlled by standard scientific or engineering measurement practices.

3. The measurement and information problems faced by social scientists in the psychosocial realism involve complex combinations of qualitative and value decisions. These are open system concerns. That is, (1) what is the quality of the interaction between systems, or sub-systems within a system—or (2) what is the value of one system in comparison to another?

4. Scientists and engineers have been eminently successful when dealing with closed system problems. However, when dealing with open system problems, such as defining the effect of a particular industrial practice upon the mental health of workers, or upon the ecological balance of the environment, closed system measurement and information practices are frequently dysfunctional or inappropriate.

This rationale is not intended to suggest an anti-science or anti-measurement stance. To the contrary, those scientific principles which have been proven to contribute positively to human progress, such as open-minded objectivity, accurate observation and measurement, and careful experimentation, should be further developed and expanded in order to improve educational practices. Rigorous and accurate assessment practices are an absolute necessity if educators are to develop the comprehensive and systematic information processes necessary to create and develop social institutions with more positive than negative results.
The problem is that the most predominant measurement and research procedures in education have been borrowed from the physical sciences, and that these processes, while valid for certain scientific and industrial practices, simply do not provide useful information in educational settings. The physical sciences have developed an array of both descriptive and experimental information in widely diverse realms, ranging from electron microscopy at the sub-molecular levels, to radio telescopes at the interstellar space level. The life sciences have achieved results similarly spectacular, though perhaps not on such a widely contrasted scale. In fact, as Lewis Mumford suggests, we have generated so much scientific and technical information that we are approaching the irrational—more information is being generated than can be usefully stored and retrieved.  

The second reason why I see the problem of the development of viable and systematic information processes in education as so crucially important at this time revolves around the basic issue of competition for power and resources. In their surge for influence and the efficient development and distribution of products, modern organizations have somehow lost their ability to perceive the need for humanistic interfaces between individuals and institutions. This may seem a strange indictment, since humans have created institutions and appear to be operating them. Yet, modern civilizations are rife with examples
in which either the mechanisms developed by institutions, or the mechanics of the institutional processes have increasingly taken precedence over the needs of individuals. As Mumford has stated:

Though we are too close to it to make a completely objective judgement, it has become obvious that our own culture has fallen into a dangerously unbalanced state, and is now producing warped and unbalanced minds. One part of our civilization—that dedicated to technology—has usurped authority over all the other components, geographical, biological, anthropological: indeed, the most frenetic advocates of this process are proclaiming that the whole biological world is now being supplanted by technology, and that man will either become a willing creature of his technology or cease to exist.5

There was a time when certain types of social organizations, such as religious and educational institutions, were able to wield considerable influence and command public respect, in comparison to scientific, business, and military establishments. Now, the powers of the former are on the wane, while the latter types are waxing to new zeniths. Again, I feel that the increasing power of the scientific and the technical is due largely to the control and effectiveness of information processes. Two factors, sheer force in the market place, and control of systematic information processes such as advertising and the mass media, have permitted certain corporate and military organizations to command unfair advantages over educational institutions and social agencies, who cannot command similar domination.
The problem is even more complex than that of competition between various types of institutions for public approbation. There has been a growing tendency in the Western world (well documented by humanists such as Mumford, Illich, Barzun, Reich, and numerous others) that the very processes of establishing individual and societal values have become more and more institutionalized. The military, the large corporations, the mega-universities, have tended to usurp the abilities of individuals to make decisions for themselves and to prescribe standards of performance and behavior for both individuals and groups.

As Reich states:

The great question of these times is how to live in and with a technological society; what mind and what way of life can preserve man's humanity and his very existence against the domination of the forces he has created. This question is at the root of the American crisis, beneath all the immediate issues of lawlessness, poverty, meaninglessness, and war...  

Rationale of the Dissertation

The dissertation is in the form of a series of logical arguments (recognizing that what appears logical to me may seem illogical to others), in which a rationale is presented for the apparent inadequacies of traditional approaches to formal information gathering and decision processes. The central theme will be that there is both historic and
current conflict between humanistic and mechanistic approaches to deciding social policy. Statements and positions of numerous eminent educators, philosophers, and scientists will be presented and discussed, with a view toward presenting the arguments of eminent critics of mechanistic philosophies.

A review of several of the most common formal research and evaluation processes will be presented, with a rationale for why they appear to be inadequate for decision making within a humanistic frame of reference. Further arguments will be set forth, discussing why these traditional approaches appear to be so tightly bound to the logical empiricist and positivist schools of scientific philosophy, that their adaptation for educational and social decision problems appears to be unfeasible.

In conclusion, case studies will be discussed which illustrate the inadequacies of certain traditional approaches. A value based model for developing information within the context of matrixes of social issues will then be offered as an alternative approach. Following is an overview of the sequence of steps which delineate the basic rationale.

1. Educators have historically been the guardians of the status quo, and the purveyors of basic cultural values. However, the pressures of modernization and the widening conflicts between diverse cultural interests are forcing educators to develop new programs for widely
varying constituencies. Concurrently, educators are asked to be rational and accountable in respect to both new and old programs, with a degree of articulateness and comprehensiveness that has never before been expected.

2. Educators as a profession have never seriously attempted to evaluate or assess programs in a systematic fashion. Formal research and evaluation training have been prominent in higher education, but the effects of this training have rarely filtered down through into the mainstream of education.

3. In a manner of speaking, this lack of application of research and evaluation methodologies on a wide scale has been fortunate, because when put to the test on a more wide scale with the advent of Federal funding of programs, these approaches were found to be sadly lacking in their ability to generate useful and relevant information for decision makers.

4. A careful, logical analysis of the formal information providing and decision models which are available from other fields for educators to adapt indicates that there are no neat, mechanistically simple approaches to solving problems within the matrix of complex social issues. Rather, the effects of science and technology have very likely created more problems for modern societies than they have solved. At any rate, problems are proliferating faster than solutions
can be found.

5. When applied in educational settings, many aspects of the traditional scientific and industrial problem solving models are not only irrelevant but are actual dysfunctional to the extent of damaging humanistic educational goals.

6. There is a critical need in education for concerted and cooperative efforts to develop information and decision processes for decision makers which are both valid and humanistic. These processes must allow decision makers to ask questions related to humanistic goals, deliver relevant information to the appropriate constituencies, and must permit democratic participation in deliberating and deciding critical issues.

Dissertation Strategy

The principal strategy of the dissertation will be to examine the logical contingencies which must occur between educational program goals and the necessary evaluative information necessary to judge the value of both the goals themselves, and the programs designed to meet the goals. One of the major difficulties facing decision makers today is that there are neither information methodologies or communication channels adequate to create the necessary linkages between the various components which compose the decision making matrix.
For example, goals are difficult to identify and define in terms which are capable of being assessed with present research and evaluation methodology. Decision makers frequently are unable to communicate with the important constituencies whom they are intended to serve in a manner which builds mutual understanding and trust.

The strategy here is to carefully analyze the purposes and the capabilities of the various research and evaluation processes which are available for decision making, to test for logical congruence between the types of information which these models provide and the information needs which actually exist. The assumption which is deduced is that there is indeed a vast discrepancy between the epistemology of scientific-industrial decision models and the value constructs inherent in educational goals. The key issue is that scientific values have traditionally excluded the concerns which are of crucial importance in educational decision contexts.

The basis for logical contingency analysis is that any social process or program may be analyzed into stages. Stake has used this type of analysis in educational evaluation to observe the congruence between program goals and observable outcomes. The schema proposed by Stake (see figure 3), is that the components of an educational process may be categorized into the stages of antecedents, transactions,
FIGURE 3

A REPRESENTATION OF THE PROCESSING OF DESCRIPTIVE DATA

Descriptive data

```
     Intended Antecedents  ----> Observed Antecedents
          ^                ^
          |                |
LOGICAL CONTINGENCY    EMPIRICAL CONTINGENCY
          |                |
     Intended Transactions  ----> Observed Transactions
          ^                ^
          |                |
LOGICAL CONTINGENCY    EMPIRICAL CONTINGENCY
          |                |
     Intended Outcomes  ----> Observed Outcomes
          ^                ^
          |                |
     Observed Antecedents  ----> Observed Antecedents
```

CONGRUENCE
and outcomes. Standards are established for each of these phases, which is then judged with the appropriate criteria. The basic principle of this type of analysis is that there are certain types of logical contingencies which must occur between each phase, which are judged either through conventional wisdom, logical analysis, or value criteria. The problem in education, of course, is to arrive at the most valid and objective type of judgements. This is also the cause of our dilemma.

**Chapter Outline**

**Chapter One.**—Presents an overview of the basic issues facing decision makers within the broad context of social issues. The position taken is that these issues are necessarily associated with important political, social and moral issues—-that they cannot be separated into the objective, value free constructs upon which scientific methodologies are based.

**Chapter Two.**—Presents a statement of the basic problem: that social concerns and the problem areas related to understanding the psycho-social dimensions of human behavior are much too complex and heavily weighted with value overtones to logically fit into the problem solving framework provided by scientific methodology.

**Chapter Three.**—Relates some of the basic purposes and definitions of research and evaluation provided by contemporary
methodologies. This is followed by a brief discussion of historic antecedents and philosophical issues related to the science-humanism conflict. The discussion is focused upon the sudden emergence of psychological and measurement constructs in education at the beginning of the 20th century.

Chapter Four.--Three positions are presented in this chapter: (1) statements of social scientists which define their allegiance to the scientific method; (2) a delineation of the constraints imposed upon scientific inquiry by the traditions of logical empiricism, and positivism, together with determinism and reductionism; and (3) some criticisms of leading educators in regard to the relevance of science for studying educational and social problems.

Chapter Five.--The basic framework for the logical contingency analysis between humanistic and scientific paradigms is presented. The principle argument is that since both the antecedents, transactions, and outcomes of the scientific method are basically antithetical to humanistic knowing and valuing, the usefulness of research and measurement methodologies based upon the positivist-empiricist tradition is highly dubious. The principle quantitative methodologies, parametric statistics, non-parametric statistics, and correlation techniques, are presented and analyzed in terms of the actual quality of information they are able to yield for the decision maker. Since the information yield is
only in the form of statistical estimates based upon predetermined
rules for calculating a calculus of probabilities, it is concluded that
these models are of highly questionable relevance for settings in
social environments.

Chapter Six -- A rationale is presented for considering various
types of systems as either open or closed, depending upon the nature
of information transactions between the system and its environment.
The position presented is that as scientific-technical information pro­
cesses have evolved within the framework of closed systems, they are
logically incompatible with the communication processes needed for
open systems. Open and closed systems are defined, together with
some dimensions of open system information requirements.

Chapter Seven -- The rationale for considering the primary pur­
poses of educational research and evaluation as means for providing
the communication needs of decision makers is presented. Because
educational problems occur within a complex matrix of varying social
factors, research and evaluation processes in order to be of maximum
utility to decision makers, must help establish viable communication
linkages between educational programs and their clients.

Chapter Eight - It is concluded that the critical role to be played
by research and evaluation methodologies is to provide decision makers
with evaluative data upon which value judgements can be based.
A model is proposed which categorizes decision making functions into logical phases composed of needs analysis, weighing of alternative strategies, and program monitoring and evaluation. A general paradigm for considering the critical elements of information systems necessary to establish communication linkages between program designers and clients is also developed.

The critical issues are that researchers, evaluators, and decision makers must consider information processes within the context of communication linkages necessary to develop viable, integrated system processes. Unless evaluative information help decision makers to apply value judgements, communicate with important constituencies, and fulfill humanistic educational goals, the mission of research and evaluation is not fulfilled.
REFERENCES


3. Ibid.


5. Ibid.


CHAPTER 2

HISTORIC EFFECTS OF AN UNFORTUNATE ANALOGY

Introduction

In this chapter a basic outline will be presented of a rationale which I hope will delineate clearly and unequivocally that scientific-technical decision processes are critically deficient as primary models for problem solving in education. I hope to show that the issue is much more serious than a set of polite, philosophical differences between differing points of view. There is increasing concern from persons representing many diverse disciplines, that the combined forces of science and technology are influencing a thought structure that is not only becoming a dominant value system, but which is also influencing monolithic institutions in a way that threatens personal freedom.

Writers have attached various names to this phenomenon (Mumford's term "megamachine" has already been mentioned). Tesconi and Morris have devoted an entire volume, the Anti-Man Culture to a discussion of the effects of this force upon the purposes and practices of
education. Their descriptive term is "bureautechnocracy" which they define as

a pattern of social management wherein the hierarchized, pyramidal, depersonalized model of human organization (bureaucracy) is linked with standardized, rationalized means (technology) with the overall aim of achieving control, flexibility, and efficiency in reaching some commercial or social objective.¹

In describing their views of the effect of bureautechnocracy upon the modern psyche:

The marriage of structure and process, thus bringing together of two powerful social inventions into a single instrument, has captured the American imagination as the single best method of dealing with virtually every problem that confronts us as a people. As the imperium of methodologies, it is now shaping its master into its own image. It is converting human beings into bureautechnocratic instruments.²

These authors are particularly concerned about the effects of bureautechnocracy upon the traditional values of schooling. They cite numerous instances of how the school curriculum and decision methodologies are increasingly based upon industrial concepts; how the principle of mass production produces as many educated students as efficiently as possible by using assembly line techniques; and that emphasis is on the "technology" of instruction—with stress on accountability by using industrial methods to measure performance.
There has been an increasing tendency for the purposes and practices of education at the most fundamental levels to be geared to the constraints imposed by tests and measurement. At a recent session of the American Educational Research Association, two noted speakers spoke of this problem. Ebel and Hastings, although speaking at different sessions, took a very similar stance. They both used the example of the effect of the testing movement upon the manner in which teachers have tended to organize their study units in both scope and sequence according to what is demanded for testing. They concluded that the very structure of curriculum has increasingly tended to emphasize cognitive achievement to the exclusion of other Values.

Even more severe have been the side effects from the emphasis on achievement tests, both for purposes of grades, and for sorting students into ability categories. Although Western Europeans have used tests to classify students for generations, Americans are only beginning to realize that classifying students by tests is of questionable value in light of our democratic ideals. Our methods of categorizing and labeling have not been so overtly discriminatory as in the European examples. We have meant well in our use of diagnostic tests, I.Q. scores, and achievement measures, believing that by so doing, we would become as precise, efficient, and accountable as our
scientific and industrial peers. Although our methods of labeling students do not overtly place them into different tracks, such as college preparatory, or trade school, as in the European examples, students are categorized nonetheless, and the curriculum content and the institutional structures are designed according to these categories.

For example, the curriculum is divided into segments which are only available to students who have achieved specified levels of mastery. At first glance, this seems very reasonable, and is in fact one of the basic premises of modern instructional technology, for it appears eminently sensible to only permit students to study those subjects for which they have an adequate background. But there are many serious flaws in the technology of testing, and in the manner in which educators have striven to use industrial approaches to the design of instruction. Two of these problems will be discussed briefly.

First, educators have almost exclusively relied on statistical norms and averages as a basis for their diagnostic and achievement tests. Aside from the many technical and philosophical issues involved in the statistics of mental measurement, the adverse social side effects are surely severe enough to give educators pause. I will discuss many of these issues in a later chapter, for educators must not only be concerned with the adverse side effects, but even more importantly, with the inherent logical issues and problems of validity and relevance.
There is a growing body of criticism concerning the misuse of tests and the over-emphasis on cognitive learning, especially as found in the lower grades of the elementary schools. Many schools are dropping the use of formal grades and are using checklists with written comments. Some large cities, such as New York, have dropped the use of I.Q. tests altogether because of the way they tend to be misconstrued by teachers and lay persons alike. In discussing the reasons for searching for alternatives to I.Q. tests in New York schools, Loretan remarks:

As one teacher said, 'Once you know a child's I.Q., you tend to see him through it... What is worse, the general public thinks I.Q. means brain power you either were born with or not... Yet we know that an I.Q. test tells only how the pupil performs on a limited number of tasks at the time he takes the test. (author's emphasis)'

Loretan cites several studies conducted in the New York schools in which I.Q. scores were raised as much as 40 points for individual students. He concludes that test scores are more truly a reflection of the combination of good teaching, motivated students, and adequate learning materials. He quotes Chauncey and Dobbin:

No intelligence test opens the window in the student's skull through which psychologists and scientists can ascertain the amount of latent brightness or intelligence he has. Nor can any test trick a person psychologically or otherwise into revealing how much brilliance or stupidity he possesses.
Educators are going to have to face the evidence that not only are traditional tests and measurement practices inadequate in themselves, but that these methods have actually served to drive the wedges of disintegration more deeply into societies. Educational development will follow a very tenuous and controversial road unless these issues can be reasonably and logically examined—and then resolved.

The second critical problem is related to the moral question of using the schools as an instrument of elitism and social discrimination. Again, I must emphasize that this problem has not resulted from the conscious intent of educators, although it must be said there are always special interest and pressure groups in society who do not hesitate to influence the schools toward these ends. No matter how inadvertently, when the school curriculum emphasizes cognitive achievement over other social values, when the content of instruction is organized according to academic mastery, and when learning opportunities are only available according to the relative position students attain on the scale of mastery levels, then it is the institutional structure of schooling which becomes the primary mechanism for determining the availability of learning opportunities.

Thus, the scientific premises of measurement, categorization, and statistical prediction have tended to shape the criteria by which students are classified and filtered through the educational system. The industrial
principles of efficiency, mass production, and rational allocation of resources according to scientifically derived rationales, have tended to determine the content and design of instruction. These deductions are witnessed by the heavy emphasis placed on statistical norms, the manner in which students are placed on scales derived from quantitative distributions of scores, and the ever increasing values administrators place on such technical constructs as accountability, achievement measures, and behavioral objectives.

Even without the many new demands for efficiency and accountability in modern education, educators would still be faced with increasingly complex management and decision making problems. School systems and universities have become so large and cumbersome that certain types of cost analysis and operations research are necessary simply to meet the needs of fiscal accounting. However, it is one thing for school systems to meet industrial standards in keeping track of materials and for balancing budgets, but another matter entirely to make value judgements concerning the welfare of students, or the worth of various programs to society.

In times past, community prestige and material success were not so closely related to academic accomplishment as they are today. It mattered little, fifty or a hundred years ago, if a man were a school drop-out or in the bottom fiftieth percentile of his class. If he didn't fit into the scheme of schooling, there was plenty of work available, or
cheap land to farm. Many of the land barons and industrial magnates of the period were unschooled, or barely literate. This is not to underestimate the high regard for professional training and the liberal arts which have always been held by the genteel and professional classes.

However, modernization has changed this hurly-burly, democratic approach to success and achievement. Now, it is not only the wealthy and professional classes who demand quality schooling for their children. Both because of changing national goals, and because of the sheer complexity of the modern world of work, social and material success depend upon schooling much more than in the past. Job status and community prestige depend to a high degree upon the quantity of schooling obtained.

For better or for worse, we have moved toward a meritocracy where everyone does not have an equal chance for the available success indicators as in the past. For it is not only the amount of schooling that is obtained, but also the type of degree and the prestige of the institutions attended, which determine entrance to the various social and economic strata. As a result, there is fierce competition for the most prestigious types of schooling. Parents and students become bitter when high marks are not obtained, and this is the inevitable result of a system where students are graded and sorted according to academic achievement. Even the positions at the bottom of the meritocracy require certain minimum standards such as a high school diploma or specialized training.
Ivan Illich has stated that the effect of obligatory schooling and the institutionalization of all available channels for permitting the individual to enter economic and cultural life is "neither reasonable or liberating." He states:

It is not reasonable because it does not link relevant qualities or competences to roles, but rather the process by which such qualities are supposed to be acquired. It is not liberating or educational because school reserves instruction to those whose every step in learning fits previously approved measures of social control.  

The result of this emphasis on schooling is that an hierarchy of caste systems is created which distributes people from top to bottom according to the amount of schooling they are able to obtain. Illich points out that this is indefensible morally, because only the right combination of background and academic skills permit one to rise to the top levels of a meritocracy. But this does not resolve the basic moral issues: whether or not meritocracies based solely upon academic prestige and qualifications are socially desirable, and the extent to which the schools themselves should participate in the competitive, academic sorting processes for entry into levels of the meritocracy system.

We cannot pretend that this is an issue easily resolvable using the problem solving methodologies which appear to be most readily available—those provided by science and industry. Nor can we assume that issues related to public education will ever be resolved with a
broad consensus. The needs and expectations of modern pluralistic societies vary so widely that public education can hardly be expected to completely satisfy such diverse constituencies. But decisions must be made, the merits of conflicting social values must be carefully weighed, and public resources must be judiciously allocated according to our basic democratic principles.

But who will define these issues? How will alternative programs be described? What will be the processes for evaluating the positive and negative values of alternative approaches? How will social issues be debated? On what merits—what criteria, will alternatives be judged? How democratic will the decision processes actually be? How will the decisions be implemented? And how will the decision results be monitored? As difficult as these questions are to answer, and this list certainly doesn't include all of the important issues, answers must be found within a context which recognizes both the traditional humanistic concerns of educators, and the diversity of modern pluralism.

**Statement of the Basic Argument**

The statement of the basic argument is that the purposes, the methods, and the results of scientific-technical problem solving paradigms, as presently constituted, are worse than inadequate for use by educators; they are seriously at cross purposes with humanistic
educational objectives. Although the argument would be simpler and easier to handle if the manner in which humanistic goals and bureau-technocratic goals are at variance could be specified in discrete categories, this is admittedly extremely difficult. However, several attempts will be made as the argument is pursued.

Humanism is richly embodied in many religious, cultural, and artistic traditions. The American founding fathers exemplified humanism and strove to develop a system of government which would guarantee the basic rights of humanism to succeeding generations. Although there were serious flaws in their efforts, every generation has perhaps made some progress in insuring these rights to more and more citizens. Some modern figures in humanism I can cite are Tolstoy, Schweitzer, Gandhi, de Chardin, du Nuoy, and Norman Cousins. Hopefully, each of us knows numerous individuals who exemplify the humanistic qualities of patience, understanding, and both personal and synectic creativity.

There are many, many more sources of humanistic thought than those listed above. In many respects, the history of the arts, humanities—the complete spectrum of cultural history—is a testament to the struggle of evolutionary man to employ his creative faculties to the highest. Quite obviously, history is also replete with examples of archetypes who have cared more for their own personal power or aggrandizement than for the refinements of culture.
In the stream of historic events, the creators of humanistic thought have only represented a small trickle. Christ and his followers were a ridiculous band of hippie paupers compared to the might of Rome. The wisdom of the Middle Ages was preserved in a few libraries and a small number of monasteries—a pittance compared to the power of the feudal monarchs, or to the bureaucracy of the Catholic church itself. The arts of the Renaissance, the Baroque, the Romantic eras—were puny in comparison to the power of Napoleon, or of the British empire. Who was Albert Schweitzer compared to a large university research hospital? Gandhi was a skinny little man who actually believed that by starving himself further, he could bring the British empire to its knees.

Historically, then, humanistics have, in number, represented a small minority of individuals. But even a brief look at the names of these individuals mentioned above should make it clear how far-reaching and earth shaking the effects of the few can be. Our world today, however, is compounded and confused by many more influences, needs, and possessions than that of any humanistic of the past.

Thus, the large question is: "How can the wisdom and insight of the humanists be brought to bear to creatively shape today's institutions and bureaucracies?" It is the viewpoint of humanists that the megamachines, the bureautechnocracies, and the materialism of modern science and technology are not presently leading modern societies
toward the most desirable ends: that, in fact, compulsive consumption, unlimited growth, the disintegration of the modern psyche—are inevitable unless we somehow learn to consciously bring these problems under control.

How can these issues be best resolved? Are they so complex that only dictatorships or fascist governments can bring the disparate factions into a consensus? Are we to depend upon elitist bureaucracies to solve the problems and mandate the solutions? We are well aware of examples resulting from these approaches, both historically, and from the present. If we were to choose a democratic route toward solving these problems, reason and logic would suggest that a well-educated and properly informed citizenry should decide the issues democratically. But even this latter solution is almost hopelessly naive given the complexities of modern societies. Only the most simple-minded of high school civics teachers or Fourth of July speakers could seriously suggest it.

Using our own country as an example, many issues, such as the potential danger of nuclear reactors, or the benefits of voucher systems for public education, are unresolvable with present knowledge. Other problems which face educational decision-makers in particular might be:

1. Which is more beneficial, spending $200,000 on a computer-assisted instruction system, or hiring 15 teachers to perform the same instruction?

2. What is the effect of the traditional grading system upon the
self-concepts of students, and the attitude of parents, who also happen
to be taxpayers? (3) Where is evidence that students in modern
societies may learn more useful information outside of the formal structure
of schooling, particularly at the high school and college age? Should we
therefore utilize the instructional benefits of modern technology and
disband formal schools altogether, say above sixth grade? (4) In
reference to the busing to achieve racial balance issue, (a) how valid is
the information that suggests black students really do achieve better in
integrated schools? Is this effect due to other factors? (b) In a given
community, there are other factors to be weighed, such as the expense
of busing, the moving of children to schools in neighborhoods other
than their own, and the increased crowding which results in the white
schools. (c) Are the tentative results to be obtained from busing worth
the accompanying disadvantages? (d) Are there more desirable alterna-
tives than busing?

These problems are mentioned to illustrate my basic position
that all of the important problems related to decision-making in the
schools have serious social and moral dimensions. Since there are no
simple mechanistic solutions handily available, educators must learn to
look to the most appropriate sources for assistance. I submit that the
wisdom of the humanities, and the finest attributes of the reasoning
processes of enlightened citizens should be brought to bear, even
though, whereas scientists and engineers have been able to develop approaches to solve their problems with elegance and pristine power, the resolutions of social issues are bound to be inelegant and somewhat "messy", as have been the ecological side effects of modern technology itself.

Every decision-maker in a position of social or educational responsibility is well aware that his problems are difficult to conceptualize, awkward to resolve, and that many choices, no matter how they are decided, are bound to create distress, and perhaps tragedy. The researcher who has only been trained in the most sophisticated of experimental procedures and principles of statistical analysis has little to offer the decision-maker who is facing real world dilemmas. There are no laboratories for deliberating social variables, where the clutter and contamination of reality can be eliminated. Nor are there social measurement constructs which can be neatly isolated and manipulated.

Nor can the decision maker be very happy with the evaluator who is a self-styled expert in the standard manuals of cognitive achievement and norm-referenced tests. It does the administrator little good to know that the reading level in school A is .2 years below school B. The evaluator himself feels awkward and indisposed when a group of teachers ask for assistance with assessing the self-concepts of children, or of evaluating the differential effects of an open classroom learning
environment as compared to each other.

I must emphasize that I am not chastising individual professionals, nor do I wish to belittle the profession in general for not having developed the capability to respond to comprehensive and difficult assessment techniques which have never before been demanded. I am simply trying to point out that it is illogical to expect the scientific Weltanschauung to provide a satisfactory frame of reference for the moral, social, and cultural milieus. Nor can the world of science and industry provide useful problem solving methodologies, when in their 400 year history, they have almost completely excluded the study of moral, educational, and cultural problems from their span of experience.

On one occasion several years ago (1954) Ross Mooney stated:

The most critical problem today is that of how men are to deal with men--how nations are to treat nations, how groups are to treat groups, how man is to treat himself. We have the physical power to destroy man if we will, or to serve him abundantly; the critical question is what man wants to do with himself."

Mooney relates these problems to the concerns of educators by stating, "Teaching men how to deal with men and doing research to improve such teaching is the primary obligation of universities now."

Although this was written nineteen years ago, we might say that the need today is even more acute. One might also ask, "How have the academic disciplines addressed themselves to this problem, if at all?"
Mooney has held the position for many years that one of the most critical and abiding problems in modern education has been the gradual disintegration of institutions as a result of the increasing separateness of the academic disciplines. He has stated recently (1971) that:

> these domains, over time, had slipped apart so far that, in the unexamined inner depths of Western man, disintegration had set in, and we were split without an inbuilt discipline for integrating life.\(^9\)

It is difficult, perhaps impossible, to articulate just how or why the more formal aspects of education have tended to disjoin, rather than integrate the image of modern man. Contributing factors may be due to a combination of reasons—the modern tendency toward increasing specialization, the academic reward systems whereby professors and researchers are encouraged to publish, no matter the degree of minutia or social irrelevance of the subject matter. Perhaps, the strongest contributing factors have been the somewhat narrow interpretation of the benefits of scientific method by the social sciences.

Over the years the social sciences have continued to adopt their interests and problems to the constraints imposed by their research models, rather than creating new approaches to study problems as they emerged. The result has been that their units of measurement and their experimental constructs have tended to become increasingly abstract. Every scientific construct is an artifact, an abstraction, a creation of
man's inventiveness. Because scientists in the physical and biological fields have been able to invent and discover constructs which are highly isomorphic with reality, Western thought has gradually lost sight of the basic abstract nature of scientific descriptions, so that now, modern man thinks that the language of science is synonymous with reality. This modern penchant for naming, describing, measuring is inculcated into children from a very early age with the result that the intellect of modern man has been split from his feeling, intuitive nature.

As a result, certain scientific disciplines have actual rules forbidding the use of terms couched in esthetic, emotional, or institutional language. Much of the literature of the social sciences, and certainly the bulk of research reports are completely devoid of rich, descriptive language, and perhaps of rich, descriptive concepts as well. The interests and the subjects of these reports tend to become as asceptic from the dirt and the glory of real life as the most barren and sterile laboratory.

In Search of Viable Models

The issue is not whether or not educators and social scientists should use models, for all formal organizational procedures must follow some type of model. If traditional models do not exist which adequately solve pressing problems, then new ones must be developed. A model need
not be constraining or rigidly systematic. It might just as well be heuristic, or open-ended; that is, a search for problem solutions not necessarily rooted in any preconceived notions of what the outcomes might be.

Decision makers in several areas of the social sciences, including educational planners, have increasingly used the concept of model building, which has been borrowed from the physical sciences and engineering. Undoubtedly, the use of models to depict educational processes, makes it easier for educators to conceptualize problems and to perceive the structures and interrelationships which exist in problem contexts. As Trzebiatowski states:

... models reduce complex problems to a manageable size, thus making it possible to examine individual segments without losing sight of the whole or losing sight of the relationship between segments. For educational planners, the ability to reduce complex problems to manageable size is perhaps the most useful function of models.\(^{10}\)

However, there are several potential dangers involved in the modeling of social systems, which are related to the difficulties in defining system parameters and in delineating educational problem contexts to be discussed later. As Trzebiatowski suggests, "Most of the faults of models and model building seem to involve an overemphasis on the model itself, to the detriment of the subject matter which it represents."\(^{11}\) The great challenge facing educators is to
conceptualize and develop models for research and evaluation which are reasonably congruent with the real problems faced by decision makers.

The degree to which a model is able to accurately portray the reality it is supposed to represent is termed isomorphism. Hill and Kerber state that the basic concept of a research model is that of isomorphism. They state:

> The science of logic states that two sets are isomorphic if the following two conditions are satisfied: (1) there exists a one-to-one correspondence between the respective elements of the sets, and (2) certain structures of the sets are preserved. These two conditions are necessary if isomorphism of the sets is to exist.  

These authors describe three types of models: (1) iconic, which is a pictorial or physical representation of exact, features of a system, such as a model car, a blueprint, or a full-sized mock-up; (2) the analog model, which is a diagram or a flow chart of a system; (3) the symbolic model which uses symbols to designate properties and relationships within a system. They suggest that educational research employs analog models drawn from other fields of science. They define the analog model as, "a set of elements which is isomorphic to a defined set of elements found throughout methods of inquiry associated with certain philosophical orientations."  

However, educators and social scientists who use models are susceptible to the danger of a type of mechanomorphic fallacy, because
of the vast differences in nature between physical phenomena in science and industry, and human and social processes. Iconic models are real, though miniature representations of reality and are of course, impossible in the social sciences. Analog models of physical or industrial processes are commonly used in science and industry with great accuracy and reliability. Symbolic models are used in mathematical and computer applications with exact precision as a common, every day occurrence.

To summarize, blue-prints, industrial flow-charts, and mathematical formulae can be translated from the model stage to reality with almost perfect isomorphism, for all practical purposes. In contrast, models of individual behavior or of social institutions are subject to all of the ambiguities, conflicting interpretations, misunderstandings, misconceptions, and human disagreements as to what is "real", as many as there are different viewpoints and philosophies in the world. Later chapters will discuss some of the aspects of this problem in greater detail.

**Models as False Analogies**

Karl Deutsch has stated that "Men have tended to order their thoughts in terms of pictorial models since the beginnings of organized thought." He suggests that men commonly use either natural events
or technological examples in their immediate culture as examples or
guides for thought. The assigning of mystical spirits to natural objects,
or thinking of human functions in terms of mechanical processes are
examples of this mode of thought. Deutsch also suggests the example
that Harvey was dependent upon the technical development of mechanical
pumps before he could conceptualize the analogy of the heart as a system
of valves and pumps as a description.

In describing the classical model of mechanism, Deutsch states
that it is a strictly metaphysical concept, since "Nothing completely
fulfilling these conditions has ever been on land or sea."\textsuperscript{15} These
principles of mechanics do work admirably in industry and certain
physical sciences, of course, as long as inanimate objects or mechanical
processes are under consideration. However, it is a false analogy to
equate the information needs of educational processes with the roles and
functions of information gathering techniques used in science and
industry. Any model is an abstraction, but the test of usefulness or
isomorphism is the degree of accuracy with which a model is capable of
representing reality. The scale model of a plane in a wind tunnel is
both useful and isomorphic because the effects of airflow over wing
surfaces can be realistically tested, although on a much smaller scale
than the real plane. An electrical wiring diagram seems very abstract,
because it doesn't look anything like a transistor radio, for example,
yet the diagram can be perfectly transformed into an actual radio with nearly perfect isomorphism.

Perhaps social scientists have been justifiably attracted to scientific problem solving models because of the inherent strengths scientists have been able to build into them over the years. These models do provide a frame of reference, a model of operations, methods for observing and measuring, and processes for analyzing and interpreting data. But the challenge facing educators is to create research and evaluation models which honor the humanistic and democratic goals of education.

The traditional models were developed by scientists to observe and understand the physical universe. The methods and procedures of the physical and biological sciences were never intended to be applied to the understanding of the complexities of human learning and social interactions. Educators and social scientists might tend to feel that by developing new processes to replace the traditional models of science, that they would be surrendering to chaos, or giving up their rigorous standards.

Careful analysis would reveal that these fears are groundless, and that the wiser course would be to recognize that where the traditional models have failed, they should either be substantially improved or abandoned in favor of new approaches. Where dichotomies exist
between societal needs, and the actual ability of social scientists to provide viable problem solutions, then researchers have the responsibility to try new approaches which promise better alternatives. Granted, it takes courage and conviction to give up old ideologies in favor of new. Also, there is something to be said for diligently trying to make the best of what is available, rather than capriciously adopting every new research fad which happens along. However, one could hardly say that the views of the eminent humanists whom have been cited are faddish, nor that the debates which have been raging for 50 to 100 years about the validity of physical science approaches in the social sciences are recent developments.
REFERENCES


9. Ibid., p. 544.


11. Ibid., p. 85.


13. Ibid., p. 15.


15. Ibid., p. 339.
CHAPTER 3

PURPOSES AND HISTORIC ANTECEDENTS

Introduction

It is my contention here that the modern educational researcher or educator should be considered as a professional information specialist who provides the decision-maker with the information necessary to make the best possible decision from among available alternatives. This is not a particularly new or unique view, but one which is shared by many authors, such as Weiss, Suchman, and Rossi in the field of research and evaluation of social programs; and Cuba, Stufflebeam, Provus, Staken and Denny in educational evaluation.

Formal information processes in education occur in two principle traditions: educational research methodology, and various evaluation models and processes. As in so many scientific and academic disciplines, both research and evaluation have developed many separate areas of study and specialization. In educational research, there are schools of specialization which include experimental design, statistical analysis, psychometrics, data processing, etc. The field of evaluation
has not developed many distinctly different areas, but has certain
disciplines of specialization such as measurement, norm referenced
tests, and evaluation design. Educational research and evaluation
methodologies have both borrowed heavily from other disciplines such
as experimental psychology, sociology, and statistics.

Guba and Stufflebeam have defined evaluation as the process of
"Stimulating, aiding, and abetting insightful action" on the part of
educational decision makers. Stake and Denny have defined evaluation
as, "The task of gathering information about the nature and worth of
educational programs in order to improve decisions about the manage-
ment of those programs". Another definition of evaluation is that
given by Provus, who says it is the process of

(1) defining program standards; (2) determining whether a discrepancy exists between
some aspect of program performance and the standards governing that aspect of the program;
and (3) using discrepancy information either
to change performance or to change program
standards.

Definitions of educational research have been the subject of
considerable controversy which will be discussed at length in a later
chapter. Cronbach and Suppes have set forth some interesting
distinctions concerning the nature of educational research. They suggest
that there are two ways of categorizing research activities: conclusion
oriented research, and decision oriented research.
Cronback and Suppes term basic research "conclusion oriented", and applied research "decision oriented". They propose the following distinctions, "In a decision-oriented study the investigator is asked to provide information wanted by a decision-maker... The conclusion-oriented study, on the other hand, takes its direction from the investigator's commitments and hunches." In further defining this distinction, they state:

The distinction between decision-oriented and conclusion-oriented research lies in the origination of the inquiry and the constraints imposed by its institutional setting, not in topic or technique... The conclusion-oriented study is not planned with an eye to a definite and useful result. The main benefit is in the unforeseen ideas it adds to society's intellectual capital.

These authors make no pretense about the direct utility of conclusion-oriented studies, believing that "The great majority of conclusion-oriented studies pertinent to education speak to questions that do not enter directly into the prevailing view." They go on to say: "A conclusion-oriented study is not performed for the mass of educators; it is performed for the enlightenment of the investigator and the small community of specialists thinking about the same problem."

This question of the utility of research is probably one of the greatest sources of controversy in education, yet one which is hardly ever mentioned in research texts, research courses, and in the journals,
where the prevailing view seems to be that the scientific method is both the basic source and the most important method for discovering new knowledge. Klausmeier has also articulated the distinction between two types of research, basic and applied:

Basic research is conducted to generate knowledge and is not concerned with whether the knowledge may be useful in improving educational practice. Development-based research, referred to by some as applied research, is specifically concerned with developing a substantive or procedural output designed to achieve specified objectives and with ascertaining how well it achieves the objectives under varying conditions with children or instructional personnel of varying characteristics. 

Klausmeier further defines the differences between these two approaches by the methodologies they each employ: "The prevalent form of basic research is the controlled experiment and its variants. Most controlled experiments designed to extend knowledge about learning are conducted in the laboratory, or under laboratory conditions." In describing applied research, he states: "Development based research encompasses identifying a deficiency or problem area in some component of an instructional system; identifying or formulating objectives of that component; and developing, testing, and refining a new component to achieve the objectives." 

Regardless of the point of input in the decision process, most researchers undoubtedly feel that their work has potential for helping
to solve educational problems, and in fact, one of the central themes of the educational research paradigm is that research is the most important source for new knowledge. This assumption is implicit, if not actually paramount, in most of the social sciences. As an example, Havelock's view is that "the university is the primary source, storage point, and cultural carrier of expert knowledge in all fields, basic and applied." It must be said in fairness to Havelock, that he does qualify this remark somewhat by suggesting that even if the university is not the basic source of knowledge, it is at least the principal carrier of knowledge.

Havelock has developed an extensive summary of the role of research in the generation of new knowledge, following the basic paradigm below:

![FIGURE 4](image)

Another illustration of the projected role of research in the dissemination and utilization of new knowledge is depicted in the following model developed by the Rocky Mountain Educational Laboratory. In describing the crucial role played by research in this model, they state:
In filling the gap between research and practice, the research and development centers, colleges and universities generate research. If these research findings are to be meaningfully brought together and applied in the classrooms, then developmental work on appropriate and relevant products must be done . . . The laboratories can take the research findings related to educational problems, develop the appropriate products and consequently facilitate the applied programs of other agencies . . .

FIGURE 5
REGIONAL LABORATORIES
LINK RESEARCH AND PRACTICE

Research

Development Field Test Demonstration

Regional Laboratories

State Dept. Title III
College & University

Practice

Elementary & Secondary Schools

Colleges Universities R & D Centers
Thus, the epistemological purposes of educational research have been construed as: (1) the primary source of new knowledge in education, (2) the most logical starting point in the development and diffusion processes, and (3) the rational foundation for decision making, whether considered from the basic (conclusion-oriented), or the applied (decision-oriented) perspective.

As already mentioned, educational evaluation has most frequently been allied with the tests and measurement movement for the purpose of the selective function of assigning students to various programs, and for categorizing them into labeling systems, such as bright, average, or dull. As Bloom, et al., have stated:

... education has been viewed as having a fixed curriculum, a graded set of learning tasks, and a mixed group of learners to be classified at each major time unit in the system. Examinations or other evaluation procedures are used to make critical and often irreversible decisions about each student's worth and his future in the educational system. These decisions and classifications frequently affect his entire career.14

In further describing some of the historic and current purposes of evaluation, these same authors state:

The purpose of evaluation, as it is most frequently used in the existing education systems, is primarily the grading and classifying of students. It is designed to find those who have failed (D or F), those who have succeeded (A or B), and those who have
gotten by (C). As testing and other forms of evaluation are commonly used in the schools, they contribute little to the improvement of teaching and learning, and they rarely serve to ensure that all (or almost all) learn what the school system regards as the important tasks and goals of the education processes.

In discussing various criteria for evaluative information, Jung states: "In the process of evaluating, one must have an objective assessment of the way things are, an operational definition of what is desired in accordance with what is possible, and an operationally defined theoretical model of how and why things do, and may, operate in the system." Jung further clarifies his views:

Educational evaluation is most vulnerable in its lack of operationally defined theoretical models. Without such models, it is difficult to determine what caused a change, or lack of change. Without them, there can be no accountability for negative side effects, or reward for constructive spinoffs. We need them most, in education, to struggle toward an answer to John Goodlad's question, 'What kinds of human beings do we wish to produce?'

Although it is very important to sort out the potential needs and purposes of evaluation, conclusion-oriented (basic) research, and decision-oriented (applied) research, the discussion is somewhat academic until the utility of these various efforts for the requirements of decision-makers can be analyzed. From my own analysis, I must conclude that (1) neither research or evaluation have achieved high levels of credibility across the profession; (2) although debatable,
the impacts of research on the mainstream of education have been incredibly small; and (3) the basic perspectives and methodologies of evaluation seem so much more amenable to the needs of educators, that evaluation as a profession seems to be eclipsing the former primacy held by research.

The reasons for these conclusions are as follows:

1. Research has developed and maintained a primary allegiance to the premises and methodologies of the physical sciences, resulting in

2. sets of rigidly positivist and empiricist procedures for framing basic questions, gathering evidence, hypothesis testing, and decision rules.

3. These rigidly prescribed rules, which must be adhered to in order to be considered "scientific" have resulted in the systematic exclusion of the problem areas, values, and processes which are most important to educators.

4. Further, the basic premise of the scientific method is that only information which is obtained in laboratory or experimental settings is considered valid and reliable.

5. Although many evaluators are actually transfers from other fields such as research, educational psychology, and sociology, they have tended to be more adaptable to the practical exigencies of decision makers in real world contexts.

I prefer to believe that the wisdom and expertise of all researchers and evaluators can be synectically applied to solve the large, social dimensions of educational problems. Little can be gained by battling about which method is best, or which branch of the profession
would yield the most useful information, if followed exclusively. It seems to me that we are all involved in an evolutionary process of social development. The problems we face are new, and of a size and complexity that we have never had to face before. Surely, the solutions lie in the direction of eclecticism and cooperativeness.

Suchman, with a background in sociological and medical research, has a solution to the problem of the conflict between evaluation and research by the neat semantic ploy of combining the terms. His position is:

Evaluation is an important component of administration, whether such evaluation be formal or informal. If we view the administrative process as a 'cycle' which includes the following special activities: (a) decision-making, (b) programming, (c) communicating, (d) controlling, (and) (e) reappraising, it becomes apparent that evaluation is an essential tool of management. Since the major focus of administration is the organization of resources and activities so as to achieve some desired objective, and since we have defined evaluative research as the study of the relationship of planned activities to desired objectives, we place evaluative research at the heart of the administrative process. In fact, we would argue that evaluative research constitutes the methodological and empirical backbone of any attempt to build a field of administrative science or practice theory. (emphasis added)

Suchman further states that, "From this point of view, evaluation becomes programmatic research whose major function is to aid administrators or program operators to plan and adjust their activities."
... Evaluation is, thus, an integral part of any operating system."

He stresses that evaluation should play a major role in the major aspects of programmatic activity, which he terms as research, planning, demonstration, and operation.

Kaplan has lent his insight also to the problem of methodologies which are too dogmatically pursued:

I am pleading, in short, for a certain catholicity of outlook in behavioral science ... An artist may see the beauty in the products of other styles without therefore sacrificing his own; esthetic judgment may be all-embracing even though personal taste remains limited. And co-workers in science surely have at least as much to learn from one another as do those in the arts ... what is striking in behavioral science is how unsympathetic and even how hostile to one another such schools often are. Their internecine struggles bring into play the tactics of defensive incorporation and exclusion ... (author's emphasis)

Kaplan takes the position that any definition of a particular methodology as being "scientific" is basically illogical, for "if a definition of 'scientific method' is specific enough to be of some use in methodology, it is not sufficiently general to embrace all the procedures that scientists may eventually come to find useful." In fact, Kaplan's eclecticism is so pervasive that he uses the term "inquiry" in preference to the "scientific method" or "research".

Following is Kaplan's description of the pitfalls of narrow methodologies:
There is no need for behavioral science to tighten its immigration laws against subversive aliens. Scientific institutions are not so easily overthrown . . . The more realistic danger is that some preferred set of techniques will come to be identified with scientific method as such . . . there is also at work a very human trait of individual scientists. I call it the law of the instrument, and it may be formulated as follows: Give a small boy a hammer, and he will find that everything he encounters needs pounding. (author's emphasis)

Mooney has tirelessly worked to bridge the gulf between the physical science ways of knowing, and the types of understanding needed to help bring about more creative, humanistic approaches to science. In his view:

The social sciences are at a stage in their growth where they tend to be focused inward on their own self-realized domains without an outward reach to a more encompassing referent. Like early adolescents, they have been busy in the discovery of themselves as something present in the world, and with the feeling that they are differentiated, separated, and to be distinguished in some special way which they need to comprehend and evolve. They have sought to realize borders for separated special fields, and to concentrate on what is to be composed inside of each. Most of their energy to date has been used on the internal frame of reference with little left to see or discipline themselves by something greater yet.

In discussing the shortcomings and the potential dangers of the classical science approaches he states:
In the social sciences, much of what has been done has come out frayed, scientific ceremonies have been used as something separate and apart, to be applied wherever they would seem to fit, as a magic rite.  

Mooney takes the position that the sciences will have little utility for contributing to the evolution of man's creative and social consciousness, as long as they are restricted to the world that is external and outside of the sphere of personal awareness. He portrays the goals of an "aware" science:

The aim in a science fully formed is to bring man's consciousness into use in working out ever more relatedness between man's inner composition and the universe . . . The scientist of man is faced with the need to validate his science by the way it fits to nature at both ends, i.e., within himself and other men. . . The role of the scientist of man is that of the self-development of mankind.  

Mooney, perhaps more than any other individual, has succeeded in articulating the over-arching concerns of educational development in its most humane sense. Rather than research methodologies which emphasize the physical, objective, quantitative dimensions of life, he stresses the need for increased understanding of the personal and subjective:

'The problem,' then, is to find those roots that go into knowing what a man is as a species-specimen. The problem is to make clear what 'being human' means when it is fitted to life-development within the species frame. The roots
then go into the nature-made. What educators need is a nature-base, a science formed in resonance, not with social accouterments, as such, but with the germinating source from which life in man emanates. Death comes to men, and those establishments of men, who lose their relevance to this.\(^{26}\)

The critical issue seems to be somewhat analogous to that of a person suffering from a mania, depression, or phobia, who seeks psychiatric or psychological counseling. How, under the sun, will he receive relief, will his disintegrations be made whole, unless the therapist himself has basic understandings to bring about healing? Given a choice, who would seek therapeutic assistance from someone who was himself suffering disintegration or delusions? Of course, when people are ill, they are frequently confused and distressed, and hardly know where to turn for help, unless they are fortunate enough to have guidance readily at hand.

**Toward an Eclectic Approach**

The educational decision-maker finds himself within a complex matrix of social mores, political forces, regional conflicts, diverse and contradictory educational philosophies, all of which tend to vary according to the individual community. The manner in which those involved in the critical decision processes understand the issues and are able to cope with problems determines the relative degree of
success. If the researcher and evaluator are to help decision-makers, they must understand these problems as well as, or better than, the decision-maker. They too, must possess insights and understandings into the complex issues of human and cultural developments, as well as a wide variety of different types of social and learner interactions. Not a few, but most of the real problems facing decision-makers are related to intricate and subtle value questions. Many decisions are fraught with potential for social conflict, where the decision maker might actually commit political blunders, be he right or wrong.

The needed skills and competencies for both the information providers and the information users in these difficult educational contexts include the most profound qualities of human insight, sensitivity to critical moral and political values, the ability to frame the right questions, identify the most crucial decision criteria, and to obtain the necessary information from which to judge the range of alternatives. Most importantly, educators must understand human development, both from the standpoint of the critical interactions of the learner, and from the standpoint of the evolution of cultures.

Although these understandings do not necessarily exclude the traditional scientific problem solving methodologies, they are certainly much closer to the types of knowledge possessed by lawyers and politicians than they are of the scientist or engineer in his laboratory.
I would emphasize that Western political and judicial processes are themselves much more worthy models for decision making in education and the social sciences than science and technology provide. Modern democracies have gradually evolved decision making processes which, though imperfect, are based upon principles of human rights, freedom, and dignity of the individual person which frequently seem to be disregarded by many of the precepts inherent in science and technology.

These traditional values - which comprise the humanistic vision of man - seem to have been either willfully or inadvertently excluded from the scientific-technical world view. Many social critics such as John Gardner, Ralph Nader, and Vance Packard, have taken the position that the increasing proclivity for the institutionalization of the dictates of science and technology in disregard of these traditional values is degrading the quality of modern life. Regarding the pervasiveness of this tendency, Tesconi and Morris state:

> We have shown how the whole mentality of science gets converted into an ideology now beyond criticism. . . Anything that can be known through science, we are told, *should* be known. And it doesn't matter whether you have to turn human beings into things in order to discover it, if it is scientific, it not only should be, but *must* be known. . . knowing itself has become a bureautechnocratic industry in its own right.  

These same authors are less than enthusiastic about the potential of scientific or technical as a suitable frame of reference for problem solving in education. They ask:
Could science as an epistemology possibly break out of its narrow bed and become human? We believe the answer must be 'No'. And we build our case on three final charges which we now specify: (1) science, by definition, is incapable of studying the individual; (2) science, by definition, converts the experienced world into a nonhuman form; (3) science brutalizes man... When science turns to the study of man, it necessarily must remain true to its own logic. It cannot deal with individual entities in their uniqueness; it must deal instead with classes of beings which are to be understood entirely in terms of shared properties.28

**Historic Antecedents**

The fascination which scientific-technical decision and problem solving models hold for educators is difficult to understand. When it is considered that educators have considered themselves the genteel guardians of the status quo and the professors of eternal verities for so many hundreds of years, the sudden preeminence of scientific paraphernalia such as statistics and experimental methods seems a paradox. In contrasting the differences between humanistic and scientific epistemologies, it may be useful to analyze some important antecedents.

**The Influence of Psychology**

The plight of educators in respect to the onslaught of the scientific world view was in many ways analogous to that of a pastoral, peaceful nation before a modern technological Blitzkrieg. Because education had remained the province of the classics and the liberal arts for so many years, the basic precepts were considered virtually inviolate. John
Dewey's experimental school and the empirical questionings of traditional methods by J.M. Rice were considered to be almost profane. The conventional wisdom of education in the early twentieth century was not open to either the objectivity, or the searching inquiry of modern science. Because the traditional bastions did not expect to be assaulted, their defenses were weak, or nonexistent. Blind tradition had led to logical lethargy and philosophical flabbiness.

The humanists and traditionalists were remiss by hiding a Maginot Line of untested assumptions and ossified conventions. Therefore, the establishment was utterly unprepared to cope with the logical strengths of the scientific method, the obvious utility of measurement, and the practicality of empiricism. Moral justification actually appeared to be on the side of technics and science, for in keeping with the spirit of the new age, technology appeared to lead the way to progress. The only resistance that could be effectively mustered against this popular force was inertia.

It must be remembered that the public at this time was spellbound with the wonders of the new age so that it was perhaps inevitable that educators, particularly those interested in psychology and mental measurement, would succumb so quickly to the scientific-technical world view. From this vantage point it may be unfair to chastise educationists of this period for embracing the measurement and behavioral constructs of the psychologists so uncritically.
behavioral constructs of psychology so uncritically. These precepts which were so quickly and wholeheartedly adopted were in themselves untested assumptions. Perhaps it has taken until now to gain the necessary insight and perspective to critically examine the adverse effects of mechanistic measurement and behaviorist dogma upon humanistic educational values.

Matson traces these influences to the eighteenth and nineteenth century traditions in psychology of associationism and utility: "... association in its classical form represented an attempt to translate the postulates and methods of Newtonian mechanics, as precisely as possible, into psychological terms." He further states: "During the course of the eighteenth century, both Hume and Hartley strove to erect upon the principle of association an exact natural science of the mind which would be 'positive' and objective as the axioms of the new mechanics." In describing the utility concept in psychology of the period he describes:

... the human mind was not merely to be likened to a machine; it was in simple fact a machine, neither more nor less - a delicate mechanism whose clockwork operations were automatically triggered by physical forces from the outside (sensory stimuli), and kept going by forces no less physical on the inside. 31

This tradition appears to be continuing today, as we so frequently here the human mind likened to technical concepts such as "mechanism" or a "blackbox" which mysteriously but predictably transforms inputs
into outputs. In quoting one of the movement's leading figures, James Mill, Matson describes:

... but it was the elder Mill who undertook systematically, in his *Analysis of the Phenomena of the Human Mind*, 'to destroy the illusion of psychical activity ... to reduce every thing to constant and in some sort mechanical relations between elements which should be as simple as possible'.

Matson summarizes the premises of the associationist and utilitarian schools as:

Thus, in the forefront of the Utilitarian campaign to produce a natural science of man and society was the postulate of mechanistic reductionism: 'the analytical breakdown of all objects of investigation into their simplest observable parts or elements'.

Allport, as well as Matson, has criticized the influence of the psychology of animal behavior upon twentieth century psychology. Matson's summary of this movement is:

The development of animal psychology was, of course, part a pervasive movement of thought within the life sciences during the latter half of the nineteenth century ... (leading) to a 'triumphant materialism, or mechanism' in which the problems of mind were no longer to be tolerated as such, but were to be reduced to the form of general physical problems ... The scientific psychology which arose with the new century to perform this evangelical role was that of behaviorism.

Although it is very difficult to estimate with precision the reasons why the behaviorists adopted such a rigid, mechanistic viewpoint toward explaining human behavior, it may very well have resulted from the constructs of logical positivism. As Allport describes this problem:
Perhaps the simplest way to characterize the positivist view of man is to say that he is regarded as a reactive being. What he does is determined by outer forces or by inner drives. Like traditional natural science, positive psychology sees movement as caused and determined by pressures. Man is like inanimate objects (including machines) and like elementary organisms. (author's emphasis)

Allport states that positivism does "not pretend to be synoptic in its view of man", but rather, is limited to the study of those fragments of behavior which are amenable to empirical observation and objective analysis: "Attention is devoted to the partial, the physical, the quasi-mechanical, the regular, the logical, because these aspects can be controlled."\(^{35}\)

Allport postulates that psychologists have tended to adopt the positivist-reductionist approach in order to fit their studies to "the canons of experimental research". Thus, they have assumed that a system of nomothetic theories, laws, and generalizations can supposedly be generated by inductively assimilating the behavioral "bits" which may be studied and analyzed in isolation from the effects of the whole psyche:

Since positivism seeks nomothetic generalizations about behavior it is likely to regard curiosity about the internal order of mind—in-particular as subjective and 'unscientific'. It somehow seems more scientific to send a platoon of white rats through a maze than to occupy oneself with the complex organization of a concrete personality.\(^{36}\) It is more respectable to pursue averages and probabilities for populations than to study the life-style of one person. Such preference is not hard to explain in a culture that is
In retrospect, it may not be possible to unravel the separate threads of humanistic and mechanistic thought which comprised educational practices of this period. Certainly, there must have been many educators who shared the beliefs of humanists such as Dewey, Montessori, and Parkhurst, who striving valiantly to create more open and creative learning environments. Yet the fact remains that the mainstream of public education seems to have almost universally adopted the psychological precepts of behaviorism and mental measurement. Nor can educators take refuge in saying that they were unaware of the implications, or that the psychologists seduced them with hidden motives. Behaviorists such as Watson, Hull, and Skinner have consistently been forthright and explicit in defining the purposes of the behaviorist credo.

According to Matson, "Watson did not hesitate to take the step of repudiating all reference to consciousness or purpose as irrelevant to the scientific appraisal of man." He relates:

For Watson "psychology, as the behaviorist views it, is a purely objective, experimental branch of natural science which needs introspection as little as do the sciences of chemistry and physics." In short, if psychology were to become a natural science, it must . . . 'become materialistic, mechanistic, deterministic and objective.'

The statements of Hull are no less straightforward and unabashed in rejecting personal values and subjective notions:
No less than Watson, Hull wore his mechanistic philosophy (composed of equal parts of objectivism and reductionism) on his sleeve. On no point was he more insistent than that a genuine theory of human behavior must require the expulsion of all traces of what he called 'anthropomorphic subjectivism' - i.e., any betrayal of the presence of a valuing human observer. . . The very notion that human behavior could not be reduced in all cases (individual, or social, moral or immoral) to automatic mechanical processes, identical for men and animals, was curtly dismissed by Hull as a defeatist attitude and a 'doctrine of despair'.

Nor has B.F. Skinner attempted to conceal the implications of his behaviorist orientation in regard to such humanistic values as personal freedom, independent inquiry, and autonomous behavior. The impact of Skinner upon the design of modern instructional environments cannot be overestimated. He is undoubtedly the most quoted psychologist in areas such as programmed instruction and behavior modification. Matson states in describing Skinner's approach:

Skinner has sought with . . . candor and absence of equivocation to outline a workable program for the effective control of human behavior. It is a program which, if seriously undertaken, may be expected literally to transform the character of man as well as of his cultural environment.

Skinner unequivocally states: "A scientific conception of human behavior dictates one practice, a philosophy of personal freedom another". He describes his conviction of the ultimate benefits of a science adopted completely to the study and control of the future:
To the extent that relevant conditions can be altered, or otherwise controlled, the future can be controlled. If we are to use the methods of science in the field of human affairs, we must assume that behavior is lawful and determined. . . that what a man does is the result of specifiable conditions and that once these conditions have been discovered, we can anticipate and to some extent determine his actions. 41

This is an astounding statement when it is compared for example, to the principles of Jeffersonian democracy, or the recent statements of educational humanists. There is simply not room here, however, to discuss the many important implications of Skinner's philosophy and its effect on educational psychology and instructional practice.

The critical issue in its most simple form for the purposes of this discussion is whether or not experimental psychologists and educational researchers dogmatically follow the same statistical methods and experimental principles as the physical scientists in order to discover the antecedent conditions determine the lawful behavior which Skinner supposes.

Matson further relates the manner in which Skinner defines his positivist - empiricist view of human behavior:

If we are to harness this scientific potential, Skinner insists, we must be prepared first of all to give up the habit of viewing behavior 'in terms of an inner agent which lacks physical dimensions and is called 'mental' or 'psychic'.

Quoting Skinner: "We cannot assume that behavior has any peculiar properties which require unique methods or special kinds of
Matson argues that in order to force scientific ways of handling data in this respect requires that:

"Everything in the behaving system must be reducible to physical terms... For such a science the ideal materials would be those derived from experimental studies of human behavior; but unfortunately direct work on human beings (quoting Skinner) 'is sometimes not so comprehensive as one might wish' - for obvious unscientific reasons. In these frustrating circumstances our information must be largely drawn from experimental studies of animals below the human level - studies which, if not quite conclusive, are perfectly adequate and appropriate (according to Skinner)."

The positivist-empiricist followers of Skinner are thus placed in the unenviable position of only considering those forms of human behavior as worthy of scientific consideration which fit the objective measurement, empirical observation schema. The strictures on the types of behaviors which may be studied by these methods is even more severe when the measurement and probability assumptions of statistical hypothesis testing must be met. These will be discussed in a later chapter.

**The Tests and Measurement Movement**

Until the early twentieth century, educators were rarely confronted with momentous decisions which required broad scale and comprehensive planning and assessment. Educational enterprises were usually conducted on a rather small scale. Community values related to goals, methods, and instructional content were comparatively homogeneous. Decision making was no less important than today, but was
simpler, simply because there were fewer alternatives. Conventional wisdom and public discussion did not require complex decision processes such as system analysis or operations research.

However, this picture changed suddenly with the advent of two strong forces: 1) the influence from psychology upon mental measurement and prediction of performance, and 2) the general fascination - derived from industrial technology, no doubt - with measuring performance and efficiency. During this period, a number of measuring "instruments" were quickly developed and adopted. Terman and Goddard translated and adapted the Binet mental measurement scale for American children. The Stanford revision, published by Terman, became the standard from which group and individual intelligence tests were compared for over 20 years.

Cattell and Thorndike became a focal point for the development and diffusion of standardized tests. Thorndike soon became a leading figure in the promulgation of achievement tests and mental measurement. Apparently the time was ripe for American educators to enthusiastically embrace the values of mental measurement and achievement tests, as R.L. Thorndike and Hagen have pointed out:

The work of Binet was eagerly seized upon in this country. His tests were translated and produced in several versions, of which by far the most influential became the Stanford-Binet . . . The testing movement seemed especially suited to the temper of this country and took hold here with a vigor and enthusiasm unequaled elsewhere.45
These authors are also impressed with the fecundity of the testing movement, stating that the period from 1915 to 1930 were the "boom" years, where: "The pioneers had shown the way, and in the hands of enthusiastic followers tests multiplied like rabbits." 46

The Growth of Empiricism

The beginnings of empiricism reflected the noblest of objectives, the desire to reform the dull and deeply entrenched curriculum and teaching methods of the nineteenth century. In describing the reasons why J.M. Rice, one of the first of the empiricists, was interested in reform, Cronbach and Suppes state:

Inspired by the teaching he observed in Germany, he found himself appalled by the mechanical, nonreflective, noncreative character of American schools and set out to bring them into line with the new emphasis on interesting, reflective classroom activity that he had encountered abroad. 47

Rice felt that it would be simple logic to single out one of the most obviously dreary practices, rote spelling drills, and compare the relative performance of students receiving differing amounts of drill.

In comparing the performance of 16,000 students, from 1895-1897, he found that achievement bore no relation to the number of minutes devoted to drill. These dramatic revelations were not enthusiastically received however:

The presentation of the results brought upon the investigator almost unlimited attack. The educators united in denouncing as foolish, reprehensible, and from every point of view indefensible, the effort to discover anything about the value of
teaching of spelling by finding out whether or not the children could spell.48

However, the fetters of tradition gradually gave way in some degree due to the innovative influences of leaders such as Dewey and Parkhurst. Because of the interest in trying new and different approaches, it soon seemed only natural to compare the performance of students receiving different types of instruction.

But there were some logical and practical problems in interpreting the data from these incipient experiments. How could the interpreter be sure that measured differences were due to the treatment, or to initial differences between groups? How could one control the effects of differences between groups, or random interference with experimental variables?

During the 1920's and 1930's researchers such as McCall and Monroe began to develop methods and statistical procedures which attempted to deal with these problems. McCall understood the logical necessity of random sampling and developed a statistic for estimating the factor of random error. Englehart states that McCall's use of these procedures actually antedated the great R.A. Fisher, particularly in regard to rotation designs and the calculation of probability of error. However, it wasn't until the late 1930's when Fisher's Design of Experiments, and Statistical Methods for Research
Workers, Snedecor's 1937 text, and Linguists volume on Statistics in 1940, that the way was paved for the rapid adoption of modern experimental and research techniques by the social sciences.

Summary

In a certain sense, the attempted adoptions of scientific-technical problem solving and measurement techniques by the social sciences was a social experiment itself. Until roughly 1900, education had been deeply steeped in methods and content which had not changed for generations. Education was generally construed as a somewhat passive conveyor of deeply ingrained traditions. It wasn't until the advent of the pressures of the scientific world view and the new industrial age that the old values and methods began to be swept away.

There is no simple analysis to this complex change process. Many of the old traditions linger yet, and this is as it should be, for some of our values have evolved over thousands of years and deserve to be preserved. Undoubtedly most of the early reformers were well-intentioned, and firmly believed that education could profitably be altered in the image of science and technology. However, we are only now beginning to seriously analyze the results of this gigantic social experiment of the twentieth century.
REFERENCES


13. Ibid., p. 7.


15. Ibid.


17. Ibid.


19. Ibid.


21. Ibid., p. 27.

22. Ibid., p. 28.


24. Ibid., p. 81.

25. Ibid., p. 76.

26. Ibid., p. 57.


28. Ibid., p. 179.

30. Ibid., p. 48.

31. Ibid., p. 49-50.

32. Ibid., p. 50.

33. Ibid.

34. Ibid., pp. 53-54.


36. Ibid., pp. 35-36.

37. Ibid., p. 36.


39. Ibid., p. 63.

40. Ibid., p. 72.

41. Ibid., p. 73.

42. Ibid., p. 73.

43. Ibid., p. 74.

44. Ibid.


46. Ibid.

47. Cronbach, *op. cit.*, p. 47.

48. Ibid.
CHAPTER 4

DEFINING SCIENTIFIC PARADIGMS

Introduction

The great strength of the scientific method is that for certain types of phenomena it has been able to provide comprehensive systems for defining problems and systematically deriving problem solutions. These comprehensive systems Kuhn terms "paradigms", which, "for a time implicitly define the legitimate problems and methods of a research field for succeeding generations of practitioners". He suggests that the paradigm serves two functions, 1) to suggest an area of study "sufficiently unprecedented to attract an enduring group of adherents away from competing modes of scientific activity," and 2) to be sufficiently open ended to leave all sorts of problems for the redefined group of practitioners to resolve. 1

According to Kuhn, a scientific paradigm provides a frame of reference for defining legitimate fields of study; it prescribes the appropriate methodologies for studying problems; and perhaps most importantly, it provides a system of checks and balances for interpreting research results and deciding which types of information deserve to
be recognized for paradigm inclusion. Depending upon the nature of the particular paradigm, logical systems are developed to organize theories and constructs. Some paradigms have been able to organize their systems with the logical precision of calculus, such as physics and physical chemistry. Others, such as clinical psychology and cultural anthropology cannot have such precise systems, but use forms of comparative observations for relating deductions. Of course, the most logically constructed paradigms are those which are able to isolate experimental variables or use well controlled processes for observation and measurement.

This chapter will follow three themes: 1) the manner in which educational researchers have tended to interpret the nature of scientific paradigms, 2) an overview of salient features of positivist-empiricist philosophies, and 3) critical testimony from eminent educators who question the utility and appropriateness of scientific paradigms for educational problem solving.

Social Science Interpretations

In the previous chapter it was noted that the behaviorist school of psychology has delineated its basic premises in clear, strong language. The following statements indicate that many social scientists wish to pledge their allegiance to positivist-empiricist paradigms in
forthright fashion. In this discussion, various areas of social science are mentioned in addition to educational research, because educators have tended to interpret scientific paradigms from the lead set by other disciplines.

Educational Research

Ausubel takes the stance that improvement in educational practice must logically follow the systematic research of such variables as "long range modification (improvement) of learning capacities", and "discovery of appropriate and maximally efficient practices and ways of organizing and presenting learning materials." There can be no quarrel with Ausubel's purposes; it is adherence to the scientific method as a basis for problem solving which must be questioned.

Ausubel summarizes his position as:

The failure of education to acquire status as an applied scientific discipline can be largely ascribed to two contrasting approaches to pedagogical knowledge. One approach has relied on empirically untested theoretical propositions, on dogmatic assertion... The other approach, going to the opposite extreme, has avoided coming to grips with the fundamental research problem of education as an applied science...

He further clarifies his scientific orientation by stating: "If the profession of education is open to attack, it is vulnerable on the grounds of failing to make the progress it could reasonably have been expected to make in providing a scientific basis for pedagogy."
Stressing the importance of science as the primal source of knowledge, Gilbert Sax in his educational research states: "This text takes the position that science provides the most reliable means for obtaining knowledge about empirical relationships and that education can be studied scientifically." In support of this statement, his rationale is:

Educational research presupposes that behavior is predictable and that events act lawfully in relation to one another. If this concept of determinism cannot be accepted, no science of education is possible.5

It is difficult to imagine a less equivocable statement of scientific allegiance than this.

Hill and Kerber also have taken the position that educational research models may be isomorphic with those used in science:

Research of a human machine system, like that of education, for example, can be carried out with the rigorous standards of scientific research if the 'models' approach is employed.6

The model which they suggest, of course, is the standard, hypothetico-deductive method. It is also noted that the research and statistical approaches which these authors propose in their text are entirely consistent with their positivist-empiricist view of science, being almost entirely concerned with experimental processes and statistical methods. As in most research texts, descriptive methods, survey techniques, and principles of obser-
vation are barely mentioned as being important components of edu-
cational inquiry.

The critical issues of the relevance of psychometric measures, understanding the social complexity of educational inquiry, or the virtual impossibility of controlled experiments in social settings are scarcely mentioned. The utility of the positivist-empiricist doctrine is accepted as a foregone conclusion.

**Educational Psychologists**

Educational psychology is perhaps the field most closely allied to educational research, both epistemologically, and in terms of mutual interests. Psychology is also generally the vehicle for translating experimental findings in learning research for education. Because of this close organic relationship, many procedures and areas of study, the same problem solving paradigm frequently serves both fields.

In describing the benefits of psychological research for education Travers has related:

Through the introduction of measurement techniques, scientists have been able to provide in the course of a century greater understanding of how people learn than was provided by wise men through all previous centuries, who did not have such measurement techniques at their disposal.

Not to be outdone in immodesty, Becker in the preface to a collection of articles entitled *Selections on Behavioral Psychology for Teachers*, states: "In a matter of six or seven years... the functional
behavioral approach has won acceptance, at least as an alternative, in most educational circles. The reasons he cites are:

1. An empirical body of evidence showing how behavior can be built and changed.

2. A methodology applicable to research in the animal laboratory, the clinic, the school, or the hospital, and

3. A cadre of applied-research workers who day by day are gathering strong evidence that operant learning principles and behavioral analysis methods can be used in all sorts of settings to deal with all sorts of problems.

One can only question the degree of objectivity with which the statements of Travers and Becker were made, and the sources of empirical evidence upon which they were based.

Sociological Research

Educators have not exchanged ideas with sociologists as much as with psychologists, perhaps because of a preoccupation with learning theory and with conditions which can be controlled within the classroom environment. In turn, sociologists have been concerned with a much broader social spectrum than that which is reflected in most educational research activities. This lack of interchange has been unfortunate, as sociologists have made considerable progress in survey research, constructing attitude measures, and in identifying differential effects of various social variables. It seems that although teachers have long recognized the wide differences between children
due to factors such as individual backgrounds, interests, and levels of motivation, research as a formal discipline is only beginning to recognize the complexity of these issues.

Although sociology as a discipline has been much more cognizant of behavioral variables resulting from socio-economic and cultural differences than have educators, sociologists have none­theless tended to assert a strong allegiance with physical science research and statistical methods. For example, one author, discussing the utility of mathematical models in a compendium of social science research articles states:

To the mathematician or the individual trained in the spirit of modern mathematics, the views just presented (referring to criticism of statistical methods in the social sciences) seem to be based on nothing more profound than a misunderstanding. 'Mathematics', said the American physicist Gibbs, 'is a language.' If this be true, any meaningful proposition can be expressed in a suitable mathematical form, and any generalization about social behavior can be formulated mathematically. Mathematics, in this view, is distinguished from the other languages habitually used by the social scientist chiefly by its superior clarity and consistency.¹⁰

Can there be a more clear statement than this of the rationale for the primacy given to statistical methods in the social sciences? This statement exemplifies the yearning for nomothetic precision expressed by the social sciences from Locke and Hobbes - through Thorndike and Watson - to Skinner and present day behaviorists.
If "any meaningful proposition can be expressed in a suitable mathematical form", then the social sciences would be able to formulate their postulates into a calculus with similar logic and precision as the Queen of Sciences, physics. According to this proposition, the language of mathematics, noted "chiefly by its superior clarity and consistency", could then replace the nebulous and imprecise forms of the arts and humanities, so that human behavior could be measured, predicted and controlled with the precision long dreamed of by Hull and Skinner.

Another example of the fervent belief placed in statistical extrapolations from the older sciences into social science is presented by Reichenbach:

"... the concept of probability can be applied in the social sciences in the same way as in the physical sciences. It would be a mistake to believe that the social sciences require a concept of probability different than the usual one. The methods of probability cover historical or sociological events in the same way that they cover physical events; and since the logical problems of probability have been accessible to a solution satisfying the requirements of an empiricist conception of knowledge, the social scientist can take over the usual statistical methods without philosophical scruples about the foundations of his science."

The "usual" probability methods which this author refers to are derived from classical gaming theory, and are only empirical in the sense that phenomena such as dice throwing and the tossing of coins do provide a basis for the well known normal distribution
concept. However, if this author has sufficient empirical data to assert that historical or sociological events obey similar laws of probability as physical phenomena - the data is well hidden and does not appear in the popular texts. Further, the positions taken by statisticians such as Hogben, Bayes, Savage, Salmon, Edgington, and numerous others, flatly contradict Reichenbach's statement.

His assertion is also naive from the standpoint of modern physics, where it is becoming well recognized that all behavioral phenomena do not obey the same laws - that sometimes different laws explain the same phenomena, such as the particle theory - wave theory problem. The issue of statistical probability will be discussed in greater detail in Chapter Five.

The Hypothetico-Deductive System

The preceding statements by practitioners from various areas of the social sciences were cited to illustrate the strong desire to follow the basic tenets of science. The social sciences have tended to adopt as the culmination of scientific practice a hypothesis testing procedure usually termed the hypothetico-deductive process. Englehart lists the basic steps of this model as: 1) a research problem which is 2) characterized by hypotheses deduced from theory; 3) proceeding to the empirical testing of the consequences deduced from the
hypotheses by observation or experiment, and 4) concluding with generalizations or laws useful in explanation and prediction, inferred from these tests. Because the social sciences interpret the logic of the scientific method in the strictest possible sense, the primary purpose of the model is to insure that theories are not contaminated with untested hypotheses.

The epistemology of the hypothetico-deductive system embodies the most salient features of the empirical traditions of science. Figure 6 is a synopsis of the primary philosophical influences which comprise the epistemology: 1) Logical Positivism, 2) Logical Empiricism, 3) Reductionism, and 4) Determinism. Although this discussion cannot do justice to the historical and philosophical implications of these antecedents to present day practice, it is vitally important that social scientists, and educators in particular, become more aware of the sequence of historic developments which have comprised present day interpretations of the scientific method.

Each of the four major components depicted in Figure 6 embodies an important aspect of scientific tradition which, taken together, compose the rules which prescribe paradigmatic credibility. Each of the major components and its sub-components will be discussed in succession.

**Logical Positivism**
PHILOSOPHICAL AND HISTORICAL CONVERGENCE OF THE SCIENTIFIC METHOD

Phenominalism
Nominalism
Value Free Data
Unity of Science
Empirical Rules
Objectivity Rules
Verifiability Rules
Physicism
Reduction by Analysis
Reduction of Psyche
Molar to molecular
Inductivism
Search for Verities
Linearity of Causes
Behavioral Laws
Cause by Mechanisms

Logical Positivism
Logical Empiricism
Hypothetic-O-Deductive System
Reductionism
Determinism

FIGURE 6
The basic features of this doctrine are that knowledge, in order to be considered properly scientific, must be obtained through the systematic gathering of objective, quantifiable information which can be validated through value-free instrumentation.

Phenomenalism refers to the assumption that reality can ultimately be best interpreted through the observation and measurement of overt behaviors; that there are no necessary conditions between essence or purpose and that which is observable. In short, knowable reality is that which is overtly mechanistic or behavioral.

Nominalism is the contention that reality can be ordered into discrete, measurable categories. Phenomena are not considered to have independent, or necessarily intrinsic features or abstract qualities which demand relative value associations. Thorndike’s statement that "Whatever exists, exists in quantity, and whatever exists in quantity can be measured", is an example of this viewpoint.

Value Free Data posits "the rule that denies cognitive value to value judgements and normative statements", according to Kolakowski. It follows from this premise that the most important forms of data are those which are observable and quantifiable in value-free terms.

Unity of Method refers to the assumption that since ultimate reality is both value-free and quantifiable, the scientific method is the preferred, indeed, the only valid form for acquiring knowledge.
Kolakowski provides an excellent summary of positivism:

Defined in the most general terms, positivism is a collection of prohibitions concerning human knowledge, intended to confine the name of 'knowledge' (or science) to those operations that are observable in the evolution of the modern sciences of nature. More especially, throughout its history positivism has turned a polemical cutting edge to metaphysical speculation of every kind, and hence against all reflection that either cannot found its conclusions on empirical data or formulates its judgements in such a way that they never can be contradicted by empirical data.  

Logical Empiricism

The developers of this school of thought had the intent of creating a philosophical system which would eliminate all metaphysical statements from its vocabulary. They took the position that the essential meaning of all cognitive processes could be reduced to rules of empirical verifiability. One of their main contentions was that even written language should be reduced to empirical rules, by analyzing statements in terms of strict logic.

Empirical Rules refer to the mandate that the meaning inherent in any form of logical statement, either mathematical, or linguistic, must be reducible to empirical observable conditions.

Objectivity Rules simply refer to the dictum that measurement and observation must be objective and value-free.

Verifiability Rules embody the great scientific edict - that observations or experimental results are not valid unless they
are replicable by independent or disinterested parties.

Physicism, according to Kolakowski, "was very popular with the founders of logical positivism. The language of physics was held to be universal, and only statements formulated in this language, or statements that can be translated into it, were regarded as meaningful."\(^\text{15}\)

In describing the effects of logical empiricism upon the social sciences and psychology in particular:

However, in actual scientific practice this rule (empiricism) proved very difficult to observe consistently. In psychology it led to behaviorism, which in the opinion of this school is the only scientific psychology. The older introspective psychology is dismissed as a tissue of irresponsible fantasies concerning the 'soul' and 'spiritual' faculties. Behaviorism denies that psychological statements have any sense more or less than other statements about observable modes of human behavior; in particular, statements about 'inner' experiences are devoid of scientific meaning if they refer to something other than behavior. \(^\text{16}\)

Reductionism

This has been the essential mode of inquiry in the physical and medical sciences. The basic premise is that a complex problem can best be understood by analysis into its component parts. Many of the great scientific discoveries have resulted from this method. For example, the medical scientists, Koch and Pasteur, were not able to develop appropriate vaccines for anthrax and rabies until they were able to isolate the viruses which caused these diseases.
Reduction by Analysis refers to the study of phenomena by reducing them into their most simple components. The assumption is not necessarily that the whole can be best understood through the analysis of parts, but that the analysis of phenomena into ever more refined components is a worthy goal in itself.

Reduction of Psyche is the logical extrapolation performed by the behavioral school of psychology in order to justify studying human behavior through the process of analysis and reduction.

Molar to Molecular - Since the scientific method does not provide a basis for inductive synthesis as a formal discipline, it has inevitably tended to reduce molar to molecular phenomena, rather than concentrate upon wholistic purposefulness.

Inductivism is the logical basis for reductionism. By analyzing large amounts of isolated phenomena, supposedly the findings will eventually be synthesized into wholistic understanding.

Determinism

Determinism is a philosophy derived from Newton's billiard ball concept of the universe - that future events may be predicted by applying the basic laws of matter and motion which supposedly apply to all natural phenomena. Astronomy, laws of mechanics and physical science principles of measurement provide the empirical basis for deterministic assumptions.
Search for Verities - In spite of the development of the theory of relativity and the philosophy of indeterminism, most science and technology must still depend upon the essential causal connections which compose the practical laws of mechanics and physical phenomena. The overriding concern of most scientists is to find the primary causes with the stability and power to predict behavior.

Linearity of Causes - Until recently, most sciences have assumed that cause and effects were generally attributable to either linear or curvilinear explanations. Certainly, contiguity between cause and effect is assumed in most experimental models.

Behavioral Laws - By discovering the laws which control behavior, scientists have assumed that 1) it is possible to identify and specify the necessary conditions upon which to statistically predict future behavior, and 2) by discovering the essential laws of cause and effect, it naturally follows the future can be better controlled and planned for.

Cause by Mechanisms - Because the whole apparatus of determinism is based upon the analysis and prediction of behavior in terms of statistics or calculus, it is necessary that the mechanisms which determine behavior be analyzed and understood. The premise of underlying causes or mechanisms has been an important component of virtually every scientific philosophy.
The highest goal of scientific paradigms has been to achieve systems of nomothetic knowledge. A nomothetic net refers to a system of interrelated laws, theories, or propositions, whose relationships can be stated in terms of logic, as in a philosophical system, or in both logic and abstract symbols, as in calculus or plane geometry.

The purpose of the nomothetic net is to provide a knowledge system which both serves as a storehouse of interrelated concepts (which meet the necessary logical conditions), and to serve as a frame of reference against which new knowledge or refinements may be judged as worthy by the paradigm's gatekeepers. Generally, those knowledge systems in science which have attained the most eminence are also the most formal, such as physics and calculus.

Kyburg states that a formal knowledge system has two parts, 1) a calculus, and 2) an interpretation. The calculus provides the syntax of the system whereby propositions are dealt with, or tested mathematically, depending upon the type of language used. The interpretation aspect of the system is in relating of system concepts to the empirical world, as in physics, or the manipulation of abstractions, as in algebra. In order to accomplish these ends, according to Kyburg, the system of calculus must have four attributes: 1) a vocabulary,
2) formation rules, 3) axioms, and 4) rules of inference. 17

Kyburg's rationale is a concise summary of rules which have evolved into formal systems for studying problems (scientific paradigms), which create integrated knowledge systems (nomothetic nets), and test the validity of information (the hypothetico-deductive system).

Braithwaite, who has been one of the leading figures in the philosophy of science has also described the utility of the nomothetic net concept:

The use of calculus to represent a deductive system has the enormous advantage that it enables deductions to be effected merely by symbolic manipulation, and the correctness of these deductions can be checked automatically merely by inspecting the relationships between the symbols . . . 18

Another advantage of this type of system, as postulated by Braithwaite, is the highly specific manner in which related theories or propositions may be tested, both because of the precise nature of the symbolization, and because every proposition may be presented in the same logical form, such as algebraic symbols, or computer language. Braithwaite explains:

But the use of a calculus also ensures that thinking should be completely explicit, since no principle of deduction can be used in the deductive system which is not represented by a rule of symbolic manipulation in the calculus. 19

Here we witness another example of a scientific dictate which asserts that only certain types of information - those which may be represented by a rule of symbolic manipulation - may be consecrated
in the hallowed sanctuary of nomothetic science.

It is critical to the logic of nomothetic science that propositions and theories be reduced to the logical precision of calculus, for without nomological order, assumptions necessary for formal hypothesis tests cannot be achieved.

**SUMMARY**

The purpose of the preceding analysis has been to illustrate that there are numerous philosophical antecedents of the hypothetico-deductive problem-solving paradigm which must be taken into account before the model may be uncritically applied in social contexts. It is my position that the conditions which have led to this particular epistemology, while necessary in the fulfillment of physical science assumptions, actually serve as prohibitive constraints when applied to problems and issues in the psycho-social realm.

The scientific method justly deserves its respect in many areas of study. However, in relation to the needs, goals, and values of the humanistic perspective many serious issues of relevance and utility arise. The method, which has proven so powerful within certain parameters of inquiry is not in itself to blame. The fault lies in the misconstructions and the misapplications of the model in social problem solving.
1. Kuhn, op. cit.


4. Ibid., p. 13.


9. Ibid, p. XI.


15. Ibid.

16. Ibid.


19. Ibid.
CHAPTER 5

Criticism of the Methods and Utility of Research

Introduction

The hypothetico-deductive research model is based upon two principal conditions: 1) the control of the experimental environment so that cause and effect factors may be isolated, and 2) generating the necessary measurements so that statistical inference and probability estimates may be made. This chapter presents criticism of these doctrines which have been synthesized from several eminent sources, together with this author's own logical analysis of the flaws inherent when classical probability concepts are applied to the social sciences.

Criticism from Within the Profession

From the facade presented by the research journal, the texts, the graduate schools, and the professional associations, it would appear that the educational research establishment is a closed corporation. From the statements of Ausubel, Travers, Sax, and Englehart, mentioned earlier, it does seem that the research community is convinced of the
utility and viability of the classical research paradigm as the primary, authoritative source for generating solutions to educational research problems.

However, there are a few eminent and articulate critics who question the validity of such a dogmatic interpretation of the tenets of science. The criticisms fall into four categories: 1) the utility of the research paradigm, 2) the viability of the nomothetic net concept, 3) criticism of the methodology, and 4) the basic issue of the epistemology of science as an appropriate forum for debating social issues. The first three of these criticisms will be summarized in the remainder of this chapter. The fourth issue, an underlying theme of the whole dissertation, will be delineated further in Chapter Six.

The Controversy of Utility

As mentioned earlier, in discussing the impact of scientific methods on the educational world of the early 20th century, many psychological constructs together with tests and measurement concepts were almost universally adapted within the space of perhaps 25 years. It is difficult to estimate the esteem with which research is held across the broad spectrum of education. What appears obvious is that most practicing educators are either oblivious to the alleged attractions of research, or the research paradigm cannot be feasibly
applied in practical situations.

Campbell and Stanley have described how a "wave of enthusiasm for experimentation dominated education in the 1920's, but . . . turned into apathy and rejection during the 1930's."¹ Fox quotes a survey of research practices by Monroe in 1938, in which that author concluded that the direct contributions of controlled experimentation had been disappointing. Fox states, "Since Monroe made that statement 30 years ago, the direct contributions of controlled experimentation to education, if anything, are even more disappointing than they seemed to Monroe."²

No less an authority than Gage has asked, "Can science contribute to the art of teaching? It would be nice if the answer could be a resounding 'Yes', followed by a long parade of conclusive evidence and examples of richly useful findings. Unfortunately, that happy paper cannot yet be written in any honest way."³ Gage is certainly not a rank critic of the scientific method, but rather takes the open-minded approach that there may be valid ways for providing information about educational problems other than a strict application of traditional research methodology. Discussing the validity of various methods, he points out:

As for validity, it is not inconceivable that in the long run some nonscientific insight or artistic hunch may turn out to be superior to what can now be cited on the basis of research
evidence. That is, the truths propounded in the past by novelists, essayists, or skilled supervisors of teachers may eventually outstrip in scientific validity the results of research now available.\(^4\)

Gage also cites the review of research of J. M. Stephens, in which several hundred studies were reviewed in order to summarize research on various teaching methods. Stephens concluded that all research on school variables, not merely research on teacher behavior, has yielded negative results for the most part: "Stephens has looked at the research reports and the summaries of research and has concluded that practically nothing seems to make any difference in the effectiveness of instruction."\(^5\) This comment can be interpreted several ways: that schooling really does not make any discernible difference which can be detected with present assessment techniques; or, that the assessment techniques themselves are insensitive to the values in which educators are primarily interested.

Frymier seems to take this latter position: "In a review of almost 300 research studies along this line, it was found that 'no significant difference' is the most frequently reported research result. Other studies of educational change report similar observations."\(^6\)

Some authors such as Provus take an even more critical view:
For a long time, school research and evaluation activity has been dominated by the kind of experimental research conducted in universities to confer doctoral degrees on graduate students in psychology. The generation and testing of hypotheses by means of probability sampling, establishing treatment conditions; and drawing statistical inferences about results was in turn based on work done in the biological sciences at the turn of this century. Countless evaluations of this sort have been conducted in public schools to no effect.

In a similar vein, Guba and Stufflebeam question:

Why is it that educators are failing to provide evaluations which are at the same time useful and scientifically respectable? Why is it that evaluations which adhere to classical research methods provide information which is of only limited help in making decisions about programs, and why do the typical 'no significant difference' findings of so many of these evaluations contravene the experiences of those who are intimately involved in the programs?

Guba and Stufflebeam have perhaps exerted more influence than anyone else in attempting to lead the educational profession away from excessive reliance on classical research models, toward newer conceptions of program assessment. Their major contribution has been the "CIPP" model, which provides a framework for providing decision makers with data. However, the CIPP model, (which stands for gathering data related to the context, inputs, processes, and products or programs, respectively), does not in itself answer the basic questions: 1) how do people know what they value; 2) how can
differential schooling effects be determined; and 3) what are the

critical communication processes necessary to enable the decision
maker to apply value criteria?

Nonetheless these authors have played an important role in
articulating the serious shortcomings of traditional approaches to
research. Following is their statement regarding the need to criti-
cally examine the utility of research and evaluation:

Any professional area that is so much
avoided, produces so many anxieties,
im mobilizes the very people who might
want to avail themselves of it, is so
widely regarded with skepticism, is
incap able of operational definition even
by its most trained advocates (who in
fact render bad advice to the practitioners
who consult them), is ineffective in answer-
ing reasonable and important questions,
and has made little apparent effort to isolate
and ameliorate its most serious lacks, is
indeed on the critical list.10

Criticism of the Nomothetic Concept

Ebel has taken a very strong position in regard to the poten-
tial of educational research for solving problems: "Basic research in
education can promise very little improvement in the process of educa-
tion, now or in the foreseeable future."11 (author's emphasis) This
is a startling comment, especially since Mr. Ebel has been a presi-
Ebel's principle contention is that the nomothetic concept is invalid in education; therefore, research with its emphasis on experimental methods and statistical hypothesis tests has little value: "Basic research seeks eternal verities. Its hallmark is the carefully designed and well controlled experiment whose conclusions are rigorously tested for statistical significance."  

Ebel cites three reasons for his viewpoint:

1. The past performance of research is very poor.

2. The justifiable explanations of that poor performance call attention to serious basic difficulties that are unlikely to be overcome in the foreseeable future.

3. The process of education is not a natural phenomenon of the kind that has sometimes rewarded scientific study in astronomy, physics, chemistry, geology, and biology.

He asks: "Where are the illustrations of the kinds of rigorous laws of learning, and of teaching, we are supposed to get from basic research? Where are even the basic, operationally defined quantitative concepts which such laws might relate?" Ebel suggests that human behavior involves too many complex "antecedent and concomitant" factors to yield constructs which may be forged into the necessary statistical or logical connections appropriate for a nomothetic science.

He comments, "If the structure is complex, if the variables are numerous, if they interact complexly, if the lines of relationship have
critical regions of nonlinearity or discontinuity, then the nomo-
thetic net is likely to be sketchy or nonexistent."

In relation to the value of the scientific approach:

... the process of education is not a rewarding subject for scientific study. Science has been defined as the systematic study of natural phenomena for the purpose of understanding ... But formal education (i.e., purposeful instruction), which is what we are trying to improve, is not a natural phenomena. It is a human invention, a construction, a cultural institution designed and built by men.13

Gowin makes a similar point in asking, "Is education research distinctive?"14 The issue which Gowin raises is that scientific concepts are not valid models for educational constructs because educational phenomena are "man made (artifactual), not natural. They are therefore not likely to yield laws and other modes of invariance such as the natural sciences report in that domain." Gowin takes the view that since educational processes are purposive and vary depending upon many complex interactions of cultural and regional values, then research methodologies appropriate to assess or reflect upon these cultural attributes must be distinctive.

His rationale is, "If we recognize that educational phenomena are not the same kind of 'natural events' as those found in the natural sciences, then we have little or no reason to suppose that the theory of science which guides research practice in those areas where laws
will have been found will be the correct one for educational research. The fact that so few laws have been found through psychological and sociological research should lead us to have some doubts about this view of science as the basis for educational research.  

Shulman has also questioned the nomothetic basic of educational research stating:

If the object of "research" is the development of coherent and workable theories, researchers are nearly as far from that goal today as they are from controlling the weather. If the goal of educational research is significant improvement in the daily functioning of educational programs, I know of little evidence that researchers have made discernible strides in that direction.

Shulman expresses his impatience with the tendency of past research efforts to make unwarranted extrapolations from experimental to practical settings:

It should be no surprise that the history of behavioral science research in education is not particularly glorious. The differences between the human learning laboratory and the typical classroom are numerous. The differences between the animal learning laboratory and the classroom are far greater. Researchers have been all too quick to generalize even from the latter setting to classroom behavior.

Criticisms of Methodology

The discussion here is limited to general criticisms of the
shortcomings of the research method as it applies to nomothetic science. More specific criticisms of particular aspects of design and measurement will be presented at the end of this chapter.

Again, the issue hinges upon the question of whether or not the classical research paradigm is capable of asking the right questions and providing the relevant information to enable decision makers to render value judgements. Addressing this problem, McGrath states: "Some of the obvious shortcomings of research efforts in education stem from the mistaken view that the 'aim of science is alleged to consist only in the discovery of universal laws, a viewpoint aptly designated the nomothetic bias'." In describing the poor relationship between educational concerns and traditional problem solving paradigms, McGrath states:

The nature of the behavioral and social sciences tends to deny the nomological universe for their knowledge systems. In education, for all the concern with idiographic variables or individuality — most sharply illustrated with the concern for the individual learner with his individual differences and the need for his individualized program — researchers have spent much of their time in attempts to fit idiographic data and evidence into nomological knowledge systems. Further, they have attempted to utilize the classical ideal controlled experiment as the ideal method or nomothetic model of process for their research efforts.

Comparing the methods of the social and physical sciences,
McGrath states "the task of physical science consists essentially of evolving universals which may be confirmed by other inquiries. These universal laws are essentially nomological models which are characterized by norms rather than by the individualistic characteristics of the idiographic sciences." "On the other hand", McGrath continues,"the principal task of the behavioral and social sciences is that of evolving an idiograph of an individual for a more complete description of that given individual."

Frymier also stresses that the experimental approach typically only provides data comparing the average performance of groups: "First, by employing the experimental-control approach, what is actually being tested is a group solution (author's emphasis). That is, when we take the old program out and put the new one in, what we are actually presuming is that one group solution is better than another group solution." Frymier emphasizes that group data, which is based upon averages, or concepts related to the comparisons of other statistical parameters, simply do not yield the type of information most valued by educators - the variability and the uniqueness of individual students.

In discussing the flaws of experimental designs as they apply to education, Fox points out that, "What makes the attribution of cause possible is eliminating all factors other than the independent
variables as possible explanations of differences in the dependent variables."²² Fox stresses that "in reality, in the social disciplines we seldom have the ability to completely eliminate all other relevant factors... In fact, we have considerable difficulty even knowing what the relevant factors are for many research problems."²³

In attempting to control leakage or contamination of experimental controls from extraneous influences, Fox states:

... unlike his colleagues who study learning and behavior with animals, the educational researcher cannot be certain his subjects come from identical genetic backgrounds, cannot control their experiences before, after, or even during the experiment, and cannot control the variables extraneous to the experiment which will intervene.²⁴

Fox also makes the point, which is of critical importance to one of my central contentions, that "it is impossible to fully counter the argument that experimentation, in the sense in which the physical scientist uses the term, cannot be achieved in education, or for that matter, in any of the social disciplines."²⁵

It must be pointed out in deference to Fox's obviously broad-minded and eclectic approach, that his is one of the very few research texts which does not take the dogmatic view that parametric statistics are the supreme source of data for research paradigms. The vast majority of texts only mention descriptive, field study, and survey
methods in the most cursory, almost condescending, fashion. Fox devotes considerable space, almost half of a long text, to descriptive approaches which do not rely on statistical inference.

The issue of the appropriateness of the experimental model in education is tightly bound to the most profound issues in the science-humanism controversy and have very grave implications for the manner in which social science is conducted. Among these concerns are:

1. Experimental controls: a) Is it logical to assume that sufficient controls can be maintained in any complex social situation to generate data which meets reasonable scientific standards; and b) is it ethical to play the experimental game at all, unless all participants are adequately informed of the alternative conditions and expectations - and are permitted to freely choose their preferred treatment condition? How valid can the information derived from an experiment be unless the process itself has honored the most sacred democratic principles - which educators of all people should strive to uphold?

2. Covering laws: The fundamental premise of any research design employing a statistical hypothesis test is the assumption that propositions within the scope of the "covering law" are related in a manner which can be stated in the logic of calculus or mathematics - else why employ statistics in the first place? This assumption is
inextricably related to the principal assumptions of the hypothetico-deductive paradigm mentioned in Chapter Four: positivism, empiricism, reductionism, and determinism. The innate logic of the model mandates that cause and effect relationships can be determined in terms of statistical laws.

3. Determinism: A critical assumption – also related to the above assumptions – is that behavioral laws can be stated in terms of known causes. The foundation of the physical sciences is embodied in this principle. If antecedent conditions of any event or phenomena are known, then subsequent effects can be accurately predicted. Thus, it is necessary for scientists to carefully study cause and effect relationships under tightly controlled conditions, so that the essential covering laws can be determined with which to control and predict future events.

In criticizing the implications of this critically important argument, Krimerman takes the position that the social sciences must develop research paradigms which recognize the autonomous qualities of people, either individually, or acting in social groups. He contends that scientific determinism is basically at conflict with the democratic belief in free will: "Few social scientists see any need to demonstrate that science is inseparable from determinism and control much as they see no need to prove that science involves verifiable
hypotheses, or replicable and controlled experiments." Arguing against Skinner's position that "We cannot apply the methods of science to a subject matter that is assumed to move about capriciously," he builds a case for considering an approach to science which honors the principle of human autonomy.26

He defines the "antideterminist assumption of avoidability":

That people, under given sets of physical and sociocultural circumstances, can sometimes avoid doing in the present what they have previously done under those same circumstances, and that, without altering their circumstances, they could have in the past sometimes avoided what they in fact did. (author's emphasis)27

According to Krimerman, the science of autonomy rests on three major distinctions: "first, that between autonomous and non-autonomous acts; second, that between enabling and sufficient conditions; and last, that between deterministic and antideterministic laws or uniformities."28 These distinctions raise some fascinating questions. For example, he considers nonautonomous acts as biological functions, or instinctive reflexes - areas of study which certainly would be within the classical physical science paradigms. But this also suggests that controlled experiments which are based upon the assumption of determinism, may very logically provide valid data for nonautonomous functions, but completely irrelevant for autonomous functions.
These latter, or course, are precisely the areas of human behavior that seem most important to humanistically inclined educators. In fact, the goal of the learner as independent and autonomous has always been one of the most valued of educational objectives, at least in democratic societies. Provus claims that while research and evaluation must be conducted in as rigorous a manner as possible, educators must not permit themselves to become bogged down in the mire of inappropriate experimental or mechanistic measurement techniques. He stresses that meaningful evaluations must realistically reflect the concerns of "people - their problems; their programmatic efforts at solution and the evidence they can amass about their successes and failures . . . educational programs are enormously complex undertakings that must be evaluated in stages relative to the development and stability of the program being investigated. Values, not statistics, are the source of evaluations."^{29}

The Form of the Basic Argument

The critical question is whether or not the purposes, the methods, and the results of scientific-technical information and decision processes have enough utility for decision makers to justify the claims for primacy. My conclusion is, "No," and the remainder of this chapter and the balance of the dissertation will be devoted to the
rationale in support of this position.

The basic argument hinges on the quality and relevance of the information which can be obtained using quantitative, empirical methods. The sentiments of the researcher or evaluator may be exalted; their intents may be grandiose; but if their methodology employs one of the traditional forms for obtaining statistical significance, as through Pearsonian or Fisherian hypothesis tests, the actual information yield obtainable is very small, and of little utility to the decision maker concerned with real world educational problems.

The decision maker in social contexts craves information which enables him to render judgements on the basis of data which apply directly to the decision criteria. The scientist desires data which accurately predict the behavior of the variables of interest. The industrial manager wants to know alternatives: how much they cost, and what are the expected profits. The advertising executive wants to know expected sales resulting from one type of campaign as compared to another. Although there is always a certain amount of uncertainty within even the most tightly controlled parameters, it is commonplace today for market researchers and space scientists to predict the results of their efforts within very small percentage points. Scientific-technical measurement precision is so precise today that the probability estimates of social scientists are almost laughable in comparison.
The premises and the conventions of a scientific paradigm only serve in any area as long as they are able to provide satisfactory answers to important puzzles or issues. Social scientists have tried to develop similar paradigms based on the logic of quantitative laws and theories developed through the systematic testing of with statistical procedures. The physical sciences and certain of the biological sciences have achieved amazing and powerful results using these paradigmatic processes. But in order for statistical procedures to be honest numerous assumptions must be met. Most importantly, for the social sciences, these assumptions should not violate historic and cultural insights into the nature of man, however, many social scientists have seemingly disregarded these precepts. Many social scientists seem to have the belief that science is a law unto itself.

To return to the question of, "What type of information can be derived from statistical or quantitative models?" there are three categories: parametric statistics, non-parametric statistics, and correlations. Although statistical texts are frequently voluminous, and research instructors attempt to impress students with the vast amount of statistical knowledge which needs to be attained for competence in research, a careful, logical analysis reveals that the actual information obtainable from one of these models appears in very simple and concise form. What makes research statistics seem so abstruse
and complex is the amount of mind-boggling and unwarranted assumptions which must be assimilated by the student in order to apply the models.

Each of the three types of model categories yields a basic form of information, or quantitative "estimate". The parametric family yields an estimate related to one of the parameters of a sample. The non-parametric family yields an estimate of probability related to assumed differences or similarities between sample characteristics, and the correlation family yields an estimate of linear relationship between two or more variables.

A logical contingency analysis of the models in Figure 7, indicates that the combination of sample and data constraints will only logically yield a certain type of statistical or probability estimate, depending upon the nature of the sample and the type of data used. For example, any parametric statistic, in order to honestly meet the necessary assumptions, must: 1) be obtained from a randomly selected sample from an homogenous and homoscedastic population, and 2) must utilize data which meets either interval or ratio requirements. Research models utilizing parametric techniques usually employ some type of randomly selected and assigned experimental and control groups. The variables of interest must necessarily be in a measurement category which approaches interval data assumptions.
The interpretation of the philosophical and logical difficulties inherent in the parametric models appear to be essentially:

1. The model is based on the assumption that as samples are randomly selected from a population, as the samples approach infinity, the parameters of the sample become more and more similar to the parameters of the population. When applied in the physical sciences, and to a certain extent, in the biological sciences, these assumptions are logically and epistemologically sound, because a) populations can be assumed to have known, stable parameters; b) populations can be specified, categorized, or manipulated in a manner that their characteristics can be assumed to be genuinely homogenous, i.e., of the same class, and homoscedastic, i.e., having characteristics which vary similarly.

The argument: the psychosocial characteristics of human subjects are not known with certainty, but can only be assumed. If the manner in which the essential features of either homogeneity or homoscedasticity cannot be actually specified, the only logical position appears to be that of the central limit theorem - that repeated samples from the same population eventually tend to assume the shape of the normal distribution - but at this point, the logic becomes a mere tautology and cannot be framed into a practical test.
Logically and philosophically, if the essential psycho-social characteristics of a given human population cannot be specified, then it follows that random sampling does not completely meet the necessary assumptions. What is indicated in actual research results is that the variance within experimental and control groups is often as great as the variance between the groups. This seems to indicate that human populations vary so widely that the central limit theorem does not apply, at least in psychosocial areas of assessment. The "act" of random selection is sometimes possible, but most often in education, neither the real assumptions nor the practical operations of random sampling is possible.

2. Parametric data must be either interval or ratio. According to Stevens, interval data must meet the requirement of having the distance between scale points equi-distant. Ratio data must meet the assumption that what is being quantified can be divided into equal percentages or ratios. Although this issue is unpopular and usually overlooked, it has yet to be proven, either rationally or operationally, that the measurements commonly used in educational statistics meet even the interval standards.

Critical questions are: a) Is there a psychometric measure which has a measurement scale with equal intervals applicable to either a single student, or to a whole group of students? b) Is there
a measure which can be applied equally to all students? Psychologically, we know that students tend to vary widely in their attitudes, motivations, and perceptual abilities. A bushel of wheat is the same measure, whether taken in Chillicothe or Dubuque. Does an achievement test have similar veracity in these same respective communities?

The argument: In reality this model seems to use measures which only meet proper measurement assumptions on the most shaky grounds. It is never certain whether the variance within a group truly reflects the variability of interest, such as cognitive achievement, perceptual or motivational differences, or simply error due to the wide range of uncontrollable and unpredictable variability between individual students. The crucial argument may very well be that psychosocial measurement procedures, no matter how carefully developed and normed, may still be too unstable to adequately reflect the principles of the central limit theorem and laws of probability.

3. The information yield of any statistical model only reflects the estimate of population parameters, usually on only a single dependent variable, such as reading achievement, or I.Q. score. In some methods, several independent variables may be compared, but only against a single criterion. The statistic itself is in this case an esti-
mate, but it is an abstraction drawn from a number of students. It does not yield data on individual students, but is a form of summary of the performance of many students.

Further, because the statistic may only reflect the type of criterion variable which can be assumed to meet interval data requirements, it must of necessity be of the nature of an achievement test or an I.Q. type of test.

The argument: the sampling requirements, the issue of psychosocial measurement instability, and the very delimiting and controversial nature of achievement tests appear to weigh conclusively against the utility or the relevance of parametric measures in most educational problem contexts. The undesirable side effects of achievement testing alone, as mentioned in Chapter Two, would indicate that conscientious educators should forthrightly and diligently pursue alternative forms of assessment.

The issues related to non-parametric tests are much more simple, primarily because they are, by definition, not based upon the same types of sampling and measurement assumptions as are parametrics. Obviously, the forms of information which can be validly derived from non-parametrics also depend upon the type of data or statistic which particular model yields. Non-parametric models are sometimes used when the data are divided into categories, or where groups of indivi-
duals are classified according to some variable of interest.

The central issue with non-parametrics is whether or not the probabilistic hypothesis test is a valid procedure. For example, the most simple non-parametric test would be a chi square with only 2 cells, an instance would be if 100 subjects were asked if they were in favor of fluoridation. If 60 responded "yes" and 40 responded "no", a chi square statistic would be derived, based on the probability of 60 responses being significantly different for 40. This simple solution would be based upon the binomial distribution – where tables of probability have been calculated based upon the number of times that two randomly occurring factors might be distributed, according to the number of chances.

As illustrated in Figure 8, as more chances occur, the closer
the distribution approaches the ideal of the normal curve. In reference to the example above, with a large enough sample the law of the binomial distribution would very nearly approach 50 "yes" and 50 "no" responses.

However, the logical issue is the relationship of any sample of responses from humans to the binomial, or if more than two categories, multinomial distribution. Is the concept of chance at all a valid consideration in the study of human affairs? Granted, traffic statistics and actuarial tables might be interpreted as manifestations of the laws of probability, but they might just as well be termed the laws of effects. For, what educators are not interested in, effects due to chance occurrence, but rather, the reasonable cause and effect relationships which can be interpreted from any set of data. It must be admitted that there may be random factors which appear to be at work, but it is the concern of the educators to eliminate these insofar as possible, and to develop a rationally coherent framework for predicting or interpreting cause and effect results.

Linear correlations do not present a severe philosophical problem, except for the category of interval data assumptions. Some correlation may be applied with non-parametric statistics, such as rank order correlations. At issue is whether or not relationships in the psycho-social realm can best be interpolated by the use of indicators of linear
relationships. Many phenomena in the physical world interact linearly, such as temperature and a column of mercury, the expansion of metals due to heat, etc. The question must be answered as to whether or not interactions in the psycho-social realm might not be interpreted more realistically and appropriately through other descriptive procedures.

The issue of the laws of probability is critically important, because most non-parametric tests are based upon these concepts, and all parametric hypothesis tests rely on probability assumptions. As in the aforementioned examples of chi square and the binomial distribution, the principal assumption to be met is that the chances of any event occurring within the sample space must be equiprobable and independent.32

In any human or social context, can it be assumed that the range of events being sampled are truly random, equiprobable, and mutually independent? These conditions obviously can be met in relation to games of chance, tossing coins, and throwing dice, but are they relevant in any way to human behavior? Krimerman's position of the distinction between autonomous and non-autonomous types of behavior is obviously related to this issue. This is not a metaphysical question, but bears directly upon the relevance of all probability assumptions and statistical hypothesis tests as
they are used in the behavioral sciences.

The Probability Issue

Although the Pearsonian-Fisherian tradition of statistics, upon which most social science statistical applications are based, uses the calculus of probabilities based on the binomial distribution, there are several other concepts of probability. The problem is a very torturous one, in which all of the historic, philosophical, and pragmatic issues are considered. For the purposes here, I will attempt to limit the discussion to three differing viewpoints: 1) classical theory, 2) empirical, and 3) subjective.

1. Classical Theory: As mentioned earlier, this theory was derived from gaming theory and observations of games of chance, such as dice throwing. The principal assumption upon which this theory is based is that of the binomial distribution. The binomial concept is based on the shape of the symmetrical distribution of frequencies of equiprobable events such as the tossing of coins or dice. The shape formed as the number of chances approaches infinity is the famous "normal curve." A system of calculus has been formed, based on the binomial distribution, which permits the comparison of distributions which are assumed heretofore to
be related to this concept: Guilford states that, "The normal distribution is the limit of the binomial distribution where \( p \) (probability) = .5 and \( n \) (number of chances) becomes indefinitely large." Guilford also gives the equation "that describes mathematically the normal-distribution curve."^{33}

It is critically important in regard to the argument presented here, that the principal concepts of the binomial distribution be reiterated, because both the equation, and the probability assumptions inherent in the normal distribution are derived from these initial concepts. In describing these initial concepts, Edgington states: "The intent of the classical definition (of probability) is
to prescribe a computational procedure for exact determination of
a numerical probability value . . . In order to compute probabilities by following the classical definition, however, it is necessary to determine which events are equally possible. **How this is to be determined is usually left unanswered.**^{34} (Emphasis added). Thus, the following conditions must be met in order to meet the inherent assumptions of the binomial or multinomial probability model:

a. Events, sample space, and chances for occurrence must be discrete and specifiable.

b. All events must have an equal opportunity to randomly occur within the specific sample points.

c. Events must be within the same logical category so that continuous data assumptions must be met.
The obvious difficulty with the classical model is whether or not any type of human or biological phenomena meet these assumptions. Are we to assume that students (events) within a classroom each have an equal opportunity to fall within any one of the points on a test score continuum?

2. Empirical probability: The likelihood of occurrence of any class of events may be calculated by dividing the number of chances by the number of trials. This type of probability is called necessary by Hogben, and statistical elevance by Salmon. However, this approach to probability is limited to settings where it is practical to take repeated measures, such as industrial or laboratory settings. It is basically an empirical approach which can be derived from the averages of numerous observations. Thus, concepts such as measurement error, or sampling error are not necessarily relevant.

3. Subjective probability: This is a concept advanced by such theoreticians as Bayes and Savage. Savage advances the position that most types of events in the social realm are subjective, i.e., no one is certain just what will occur in the future, nor are there necessary laws or rules of inference which provide a basis for predicting the future based upon a calculus of past experience.

The issue might be summarized by posing the question: "Are
researchers more interested in testing the utility of statistics as they are commonly employed, or in the degree to which statistical models fit reality, or in the converse—forcing reality to fit the model?"

Hogben, Salmon, Savage, and Stevens, among many others, not only question the relevance of classical probability to concepts in the psycho-social world, but also argue that there has never been conclusive empirical evidence which indicates that there is any necessary logical connection between social events and probability models derived from the gaming tables. Hogben asserts that Fisherian models make the same assumption as Quetelet (the promulgator of the constancy of large numbers theorem), that errors of observation are subject to the binomial law.38

Stevens takes the position that statistical-probability models are in themselves formal and tautologous, i.e., they may be entirely consistent within themselves, but have little or no relevance to a particular aspect of reality:

Mathematical theories of probability inhabit the formal realm as analytic, tautologous, schematic systems, and they say nothing at all about dice, roulette, or lotteries. On the empirical level, however, we count and tabulate events at the gaming table or in the laboratory and note their relative frequencies. Sometimes the relative frequencies stand in isomorphic relation to some property of a mathematical model of probability; at other times the observed frequencies exhibit scant accord with 'expectations.' . . . Those features of statistics
that invoke a probabilistic schema provide a further instance of a formal-empirical dichotomy: the distinction between the probability model and the statistical data.\textsuperscript{39}

The problem is well summarized in the statement of Savage:

The concept of probability almost unanimously adopted by statisticians throughout the first half of the century, and the one that still seems to be regarded as fundamentally correct by the majority of statisticians today, is the frequency concept of probability, in which a probability is the relative frequency of some kind of event in a certain type of sequence of events, or, according to some category in a set of events...\textsuperscript{40}

In contrast to this position, as stated by Savage, "is the personalistic view, but the concept of personal probability has not until recently been ripe for acceptance by statisticians. In the personalistic concept, probability is an index — in an operational sense — of a person's opinion about an event."\textsuperscript{41} Savage does not, of course, state that it is impossible to develop statistical estimates of the likelihood of certain classes of events occurring. He states that it is possible to develop statistical estimates without resorting to the classical concepts of the central limit theorem, the multinomial distribution, or the assumption of randomness of occurrence.

According to Savage, "The subjectivistic view of probability represents a relation between a statement and a body of evidence, but it is not a purely logical relation." Thus, the test of sufficiency of any statistical method employing probability assumptions
should be its utility. As Stevens relates:

A statistician, like a computer, may perhaps feign indifference to the origin of the numbers that enter into a statistical computation, but the indifference is not likely to be shared by the scientist. The man in the laboratory may rather suspect that, if something empirically useful is to emerge in the printout, something empirically meaningful must be programmed for the input.42

Although social scientists have generally seemed to accept the premises of classical probability model uncritically, the logic is inescapable that these models must be critically investigated in regard to their utility for yielding valid, empirical data for decision making in real social contexts. Savage raises the issue by stating:

The important point is that a probability statement is taken as making an assertion about the world, it may be right or wrong - and it is generally held that we never really know with certainty which it is - but it is a statement, like a statement about lengths or weights, which is either true or false, and for which the evidence is chiefly observations. In order to find out whether or not a probability statement is true, we must make an empirical investigation.43

Significance Tests: A Problem of Utility

Research in the social sciences has almost universally revolved around the premise that the statistical hypothesis test is the primary source for deciding the relative significance of experimental findings. What lay educators and many social scientists fail to realize is that the principal probability model upon which most hypothesis tests are based -
the Pearsonian-Fisherian tradition - depends upon a very tenuous set of assumptions which do not yield direct estimates or comparisons. The model does not purport to give a direct estimate that group A performed 10 percent better than group B, nor does the logic of the model explicitly provide a basis for making the direct statement that "There is a 90 percent probability that group A is 10 percent better than group B."

What the model does say is that based upon a calculus of probabilities (derived from the multinomial distribution and the central limit theorem) there is a specified estimate or likelihood that estimates of sample parameters, such as mean scores or variance, will tend to be related according to the ratio calculated between sample size and error of measurement.

The mechanism of the hypothesis test therefore only provides an estimate of significance based upon the calculus of probabilities, rather than a specification of a real difference between sets of data. The Pearsonian-Fisherian model makes no pretense of providing either direct measurement of differences between groups, or even of specific statements of probability. Any competent statistician or text carefully delimit the assumptions and the potential applications of these statistical models. In general, the logical steps in the models are:

1. Random samples are drawn from a population with specified characteristics.
2. Measures are taken on the variables of interest after each sample has received control or treatment.

3. Statistics are derived from the standard error/sample size ratio for each sample and compared.

4. The statistical comparison ratio is referred to an appropriate table to determine probability estimate.

The tables have been constructed according to a calculus of probabilities which varies in significance according to sample size and size of significance ratio. The specific statistical estimates that are derived from this method do not reflect direct measures or specify percentage or magnitude of differences between groups. Rather the estimates are considered to reflect differences only according to the appropriate calculus of probabilities.

This argument is not presented as a criticism of the statistical logic of Pearson, Fisher, or their followers. The critical issue is the potential utility of this hypothesis testing process for the social sciences, and whether or not measurement and observations validly apply to the necessary statistical assumptions inherent in the model.

In order for these assumptions to be met, certain conditions are necessary:

1. The normal distribution principle - Population characteristics are assumed to be distributed in a manner analogous to the bell shaped curve concept.

2. Principle of random sampling - It is assumed that the
parameters of random samples from the same proportions as the population.

3. Measurement - It must be assumed that the measurements themselves are valid and reliable, and meet the necessary assumptions of the particular statistic.

4. Validity - It must be assumed that the context of the experimental conditions meet as many of the conventions for internal and external validity as possible. The validity matrix developed by Campbell and Stanley represents a summary of these standards as developed to this point in time.44

5. Probability - Although this issue is defined so poorly in both social science research and general statistical texts that it seems almost impossible to define the meaning of classical probability models in operational terms, the logic seems inescapable that behavioral phenomena which are purportedly being measured must meet the same assumptions as the behaviors upon which the probability laws were established, i.e., all points within the sample space must have an equal likelihood of occurring.

Collectively, these assumptions comprise the essential features of the epistemology of statistical hypothesis tests. In my view, which I take to be supported by many others as cited throughout this paper, is that these assumptions have yet to meet the practical test of utility in the social sciences. Further, there seems to be no practical way to test the assumptions through empirical processes, as is possible in other areas of science. In the hard sciences, the true test of utility can be ascertained through replicability, or irrefutable logic.

Stevens calls this process the "schemapiric principle" - where the practical test of a model is in its degree of empiric isomorphism
Stevens emphasizes that although mathematical models began in the empirical mode where there was always a one to one isomorphism between the model and reality, models "in modern times are debated in the formal schematic, syntactical mode, where (they) bristle with symbols". He further defines the problem:

... mathematics now enjoys full freedom to 'play upon symbols' as Bauss phrased it, with no constraints imposed by the demands of empirical measurement ... So also with other formal or schematic systems. The propositions of a formal logic express tautologies that say nothing about the world of tangible stuff. They are analytic statements, so-called, and they stand apart from the synthetic statements that express facts and relations among empirical objects. There is a useful distinction to be made between the analytic, formal, syntactical propositions of logic and the synthetic, empirical statements of substantive discourse.46

Stevens, of course, is not merely debating the philosophy of mathematics and statistics in the abstract, but is urging that users in all areas of research carefully consider the dangers inherent in applying formal models without being constantly alert to the problem of schemapiric utility.
REFERENCES


4. Ibid.

5. Ibid., p. 400.


7. Provus, op. cit., p. iii.


9. Ibid.


12. Ibid.

13. Ibid.


15. Ibid., p. 18.

17. Ibid., pp. 376-377.


19. Ibid., p. 67.

20. Ibid., pp. 68-69.


23. Ibid., p. 459.

24. Ibid., p. 457.

25. Ibid., p. 457.


27. Ibid., p. 328.

28. Ibid., p. 332.


32. Ibid., p. 116.

33. Ibid., pp. 124 and 175


42. Stevens, *op. cit.*, p. 44.


44. Campbell and Stanley, *op. cit.*


There is a well-known demonstration of the variability with which different people tend to interpret the same stimulus. Mooney has frequently used the figure above to demonstrate that in the absence of an orienting frame of reference (a ground, or Gestalt) people vary widely in the way in which they interpret the three parallel lines. Some people tend to perceive the lines as parallel on the same plane, others tend to perceive the lines in terms of depth, i.e., telephone poles or fence posts disappearing into the distance.¹

One of the basic premises of Gestalt psychology is that people tend to perceive stimuli or events differently, depending upon their previous experience, cognitive orientation, motivation, and mind set. The purpose of this chapter is to develop a rationale which demonstrates
that the positivist-empiricist view of science has developed a know-
ledge system - an orienting frame of reference - which treats problems
as though they were all closed systems. This view insists that any
set of phenomena or class events can be analyzed and described in
terms of cause and effects which are mathematical, or empirically
rational. That there is any teleology of values or purposes which
cannot be explained with the rationality of science, is denied.

The classical science problem-solving paradigms are viewed
here as excluding much of the traditional knowledge systems, cultu-
ral transactions, and even the importance of autonomous thought
itself. Classical science paradigms are seen as dealing with problems
exclusively in terms of closed system methods, and further, as inten-
tionally excluding any teleological or value considerations related to
open system concepts.

Thus individuals with varying backgrounds and value systems
tend to look at the same issue from a very different perspective.
Persons with training in the methods of classical science will tend to
view a social problem in terms of simple cause and effect relationships,
as though closed system values and assumptions were sufficiently
explanatory. Others with a humanistic or open-system orientation will
take the position that linear cause and effect or closed-loop feedback
assumptions are insufficient and irrelevant as suitable explanations
This chapter will focus on the philosophical and logical differences between open and closed systems, emphasizing that information processes within closed systems are not adequate to reflect the complexity of processes within or between open systems.

**Definitions of Closed and Open Systems**

Because open system concepts are subject to philosophical debate, there seem to be no easy definitions which may be neatly applied across many situations. In my view "open" is a relative concept, and is dependent upon the degree that a system is capable of transacting with its environment; conversely, a system is "closed" to environmental transactions that it is unable to process.

Hall and Fagan state that a system is open if there is an exchange of materials, energies, or information with environments, and that a system is closed if there is no import of energies in any of its forms such as information, heat, physical materials, etc. However, this is an untenable logical position because there is a sense, both real and philosophical, in which all systems can be considered open to transactions with the environment: we all know that even an inert appearing lump of metal will be affected
by changes in temperature, or by forces such as corrosion or electro-magnetic energy.

Whereas, mechanistic information functions can be described in terms of cause and effect relationships, or by exchanges in energy or mechanical movements, living functions are infinitely complex, and cannot be adequately defined in similar terms. There are four main distinctions concerning the mechanistic or non-living systems which can be stated here:

1. Non-living systems are open only to high limited types of inputs which have been mechanically constructed specifically for the operation of the system.

2. Non-living systems use information exclusively to control or maintain the system processes in strictly a kinetic or cybernetic fashion.

3. Information modes in mechanical systems are in the form of mechanistic inputs in a specific form for the particular system, such as electric current, chemical catalyst, etc.

4. Information about the non-living types of systems, the static, the kinetic, and the cybernetic is highly specific, accurate, and directly related to the particular phenomena.

It can be noted, therefore, that the types of information which can be developed for non-living systems fit the criteria of the "pure" sciences very well, and have in fact played a central role in the development of the sciences. The information criteria of specificity, reliability, accuracy, and predictability, serve
these applications well because from the standpoint of system theory, the input characteristics can be described, weighed or measured with scientific accuracy.

In order to further clarify the roles of information in the hierarchy of system models, Figure 1 sets forth the types of systems in both the non-living and living areas. Each type of system is illustrated by at least one example and followed by the information functions, types of interaction, and information modes.

It may also be helpful to briefly mention the types of systems and give definitions of each, with the exceptions of the biological, social and personal types which are self-explanatory.

Static Systems: Static is a relative term, depending upon one's scale of observation or frame of reference. For example, a bar of metal or a table appear to be static, as they don't have moving parts, or appear to receive or give off energy. However, with the advent of modern science, we now know that there is subatomic activity in all types of matter, and that even non-living systems contain myriad processes at this level, in addition to being subject to subtle but real changes in the environment. However, compared to more complex systems, and depending upon the scale of observations, a static system does not contain mechanical parts which move, nor incur dynamic energy transactions with the environ-
<table>
<thead>
<tr>
<th>TYPE OF SYSTEM</th>
<th>EXAMPLES</th>
<th>INFORMATION FUNCTION</th>
<th>TYPE OF INTERACTION</th>
<th>INFORMATION MODE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static</td>
<td>chemical</td>
<td>external description or control</td>
<td>none, other than control of system</td>
<td>limited to description of static state</td>
</tr>
<tr>
<td></td>
<td>table</td>
<td>same as above</td>
<td>same as above</td>
<td>same as above</td>
</tr>
<tr>
<td>Kinetic</td>
<td>can opener</td>
<td>inputs to control system</td>
<td>simple cause and effect</td>
<td>description of static state, or measurement of kinetic process mechanical, or electrical</td>
</tr>
<tr>
<td>Cybernetic</td>
<td>machine control</td>
<td>control of function</td>
<td>feedback loop</td>
<td>dependent on machine design - electrical or kinetic or both</td>
</tr>
<tr>
<td>Closed Loop</td>
<td>computer</td>
<td>control and direct processes</td>
<td>cues, feedback, feedback loops</td>
<td>electrical systems</td>
</tr>
<tr>
<td>Biological</td>
<td>plants</td>
<td>growth and maintenance</td>
<td>personal interaction (informal or formal) formal group exchange formal &amp; informal</td>
<td>verbal and non-verbal print, pictorial and electronic media</td>
</tr>
<tr>
<td>Personal</td>
<td>person</td>
<td>growth, maintenance, physical or metaphysical values</td>
<td>capable of dynamic interaction at many levels including biological &amp; conscious</td>
<td>similar to biological systems, plus potential of various aspects of complete human experience</td>
</tr>
</tbody>
</table>

Figure 9
Kinetic Systems: Von Bertalanffy terms kinetic systems as "prototypes of undirected physical process, describable in terms of linear cause and effect relationships". Simple machines or chemical processes which interact with other systems or transmit some type of energy to the environment are examples. The kinetic system may be a component within a complex arrangement of other systems, but is the receiver or the transmitter of relatively simple types of energy or "information" transactions.

Cybernetic Systems: These are relatively complex organizations of static and kinetic system components and incorporate a degree of self-regulation, as in a thermostat, or control valve. Heating systems and automated industrial machinery are examples of systems which are designed with self controlling or regulating features. Computers and guided missiles are examples of very sophisticated self regulating devices.

Von Bertalanffy distinguishes between the cybernetic principles as designed into machinery, and as shown by living systems. In machines, the cybernetic principle has been designed into the machinery for self-regulation and therefore information transactions or environmental exchanges are limited strictly to the design characteristics. The cybernetic principle as applied to living systems is
subject to the purposes, drives, or types of homeostatic levels inherent in the organism.

The principle differences between biological systems and non-living systems, according to GST, is in the nature of transaction with the systems environment. Living systems are capable of dynamic transactions with the environment, purposefulness, negentropy, i.e., tending toward increased order and complexity, and have the potential to adapt or evolve to meet changing conditions. Von Bertalanffy further clarifies this:

Living Systems . . . are maintained in a state of fantastic improbability, in spite of innumerable irreversible processes continually going on. Even more, organisms in individual entogeny as well as in phylogenetic evolution develop toward more improbable states, towards increase of differentiation and higher order of matter.4

In contrast to Von Bertalanffy, however, I do not include living systems under the cybernetic heading. Although the concepts of feedback and cybernetic principles are frequently used as though they were analogous to both living and machine systems, the functions in both types of systems are similar only in a general sense.

It is my position that since classical science model is essentially a process for determing the cause and effect relationship between inputs and outputs of physical or biological systems, the model is both philosophically and operationally a closed system;
i.e., it will only accept inputs which contribute to the function of the system as a mechanical or logical process. Logical systems and mathematical systems could be considered as exceptions to this interpretation, but because of their philosophical and historical assumptions, they are analogous to the mechanical model, at least as I interpret the issues. Logical systems are nearly analogous to mechanical models in so far as the theories, axioms, and propositions of a particular system must be in a mode or linguistic form acceptable to the system. Plane geometry is an example of modal constraints, for obviously, Markovian or Boolean algebraic systems don't fit the logical input requirements of plane geometry.

Mathematical systems are perfectly isomorphic with mechanical models because all of the system functions are logically related and can be expressed in symbolic terms. Further, these systems can be broken down and analyzed, and reassembled without distorting the basic system purposes or logical relationships. This is an essential difference between living and non-living systems, for as Ross Ashby states, the way not to approach a complex system is by analysis because:

... this process gives us only a vast number of separate parts or items of information, the results of whose interactions no one can predict. If we take such a system to pieces, we find we cannot reassemble it.
Rationale

The basic purpose of the model illustrated in Figure 1, is to depict as graphically as possible a number of false analogies between the mechanistic processes of non-living systems and the living processes of living systems. Although I feel that the basic source of the controversy concerning the interpreting of scientific terminology in phrases which result in either mechanomorphic or anthromorphic concepts lies in the deep philosophical differences between the two philosophies of science and humanism, many times these complex and perplexing arguments are the result of semantic confusion.

I will try to support the rationale by discussing the information roles under the columns in the model: information function, type of interaction and information mode.

Information Function

This term can be used to describe similar systems principles in both non-living and living systems. Examples are 1) information exchange provides communication between system components, 2) efficient information exchange makes it possible for the system to serve its purpose, and 3) information exchange serves to control or sustain the system functions. However, there is a distinct difference between the types of function in living and non-living
systems which makes it impossible to use the term interchangably between the two systems. In non-living systems, the information roles are closed, i.e., limited entirely to the mechanical controls of the system functions.

Even in the largest integrated industry, although the information functions are voluminous, they can be still defined in the relatively simple terms of predicting resources, measuring inputs, controlling and monitoring processes, and measuring outputs. Since the information tasks are related to static, kinetic or cybernetic systems, they can be performed with relative efficiency and accuracy. Again, the space program and the larger industries are vivid examples of complex processes where the role and function of precise and efficient information functions are performed with amazing accuracy.

In living systems, the words which describe maintenance of the system or homeostatic states of the system are the same as for mechanistic systems, but obviously the nature of the functions are vastly different. For in addition to maintenance and homeostasis, biological systems must provide for growth, reproduction and adaptation to changing environments.

A human being is also a part of the biological process, but in addition to what was mentioned above, this kind of "living
system" deals with inter-personal and social interactions, and must interact with environments which are extremely variable. As an example of an extremely variable environment, we could consider the school, which includes thousands, perhaps millions, of variables and of which could be included in a system analysis of a particular educational problem.

In summary, the basic differences in information functions between non-living and living systems are: 1) functions in non-living systems are for the purpose of controlling the system; 2) non-living information functions are closed and therefore limited to the mechanical or electro-magnetic constraints of the system; and 3) functions in living systems, in contrast, are open to literally thousands of interactional variables with the environment and are only closed at the levels of actual physical transactions at the biological level of interaction. Although behaviorists might disagree with this point, I believe that at the personal level the individual is open to a wide variety of choices, and at the social level has the potential for a diversity of social interactions.

Type of Interaction

Type of interaction refers to the actual information process occurring in the functioning of a system. Again, there are general principles which can be used to describe the nature of interaction
between system components, such as cueing, stimulus-response and feedback loops. However, types of interaction in non-living systems are limited to mechanical, electrical, electro-magnetic, or chemical interactions. When any of these types are incorporated into a living system, they do not function in a mechanistic fashion but are organized into the amazing complexity of the life process, with its myriad of involved transactions between the individual and his environment. A need therefore exists for different principles and concepts to describe and understand the interactions of living systems as contrasted with the non-living.

Information Mode

The information mode provides the most vivid contrast between types of systems because input models for non-living systems are the basic processes of physical, electrical or chemical energy. And, of course, with the exception of the computer, most non-living systems will only accept one or a very few input modes. Here man is clearly superior to the machine, for in a single day an individual actually processes millions of different information inputs, and in the case of the educated tourist in a metropolitan center, can sample hundreds of modes of human expression as found in architecture, museums, concerts, personal and social interactions, and from the mass media. He is capable
of interacting then with a tremendous range of information modes in the psycho-social realm.

In comparison to the weak information exchange in closed systems and communication processes, communication at the psycho-social levels of human interaction is extremely rich and complex, involving myriads of transactions which have meanings at the level of cognitive reasoning - not comparable in any meaningful way to closed system information processes.

The Experimental Model as a Closed System

The contention here is that the hypothetico-deductive system, while highly appropriate for the physical sciences, is not adaptable to the needs of the social sciences because of the inherent constraints imposed by closed system epistemologies. The following model, (Figure 11)\(^6\) is from Hill and Kerber, who state that educational research models are analogous to those developed in other disciplines:

The General Analog Model of educational research is developed on the basis of its being a representamen of 'common' elements of inquiry drawn from selected philosophies. The 'model' is a set of elements which is isomorphic to a defined set of elements found throughout methods of inquiry . . . The analog shows clearly the relationships between the elements which are: a) the problem; b) the hypothesis; c) the universe; d) the sample; e) data collection processes; f) data processing, and g) research decision. \(^7\)
General Model of Educational Research

Figure 10
Note the heavy emphasis of the model upon concepts such as sampling from an assumed universe, data, and hypothesis testing.

My position is at opposite poles from that of Hill and Kerber, for I contend that the necessary information and communication elements needed in educational research are decidedly not "isomorphic to a defined set of elements found throughout methods of inquiry" associated with scientific disciplines. Nor are the scientific philosophies of positivism, empiricism, reductionism, and determinism compatible with those of either the humanistic or General Systems Theory orientation.

Figure 2 below, is cited as another example of the unified concept of the hypothetico-deductive model in the behavioral sciences. 8

---

*Steps in Hypothesis Testing*

1. Theory
2. Research hypothesis
3. Experimental design
4. Experimental treatments
5. Statistics
6. Parameter estimates
7. Test of statistical hypothesis
8. Revision of theory

*Figure 11*
This model was taken from the text of Keppel who relates in his description the usefulness of this approach:

The most important feature of the experimental method is that it is possible to infer a cause-effect relationship. That is, we can conclude that the difference we observe in the performance of the subjects in our two conditions was caused by the experimental treatment. 9

Again, note the emphasis on the assumed logic of being able to derive useful information through statistical inference of observed differences between experimental and control groups.

The concluding rationale is that the hypothetico-deductive model can be considered as a system of logic with a sequence of antecedents (inputs), logical transactions (processes), and outcomes (products), as depicted in the following figure:

<table>
<thead>
<tr>
<th>Inputs</th>
<th>Processes</th>
<th>Products</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theories</td>
<td>Sampling</td>
<td>Statistics</td>
</tr>
<tr>
<td>Propositions</td>
<td>Measurement</td>
<td>Comparisons</td>
</tr>
<tr>
<td>Hypotheses</td>
<td>Data Analysis</td>
<td>Hypotheses</td>
</tr>
<tr>
<td>Decision Rules</td>
<td>Hypothesis</td>
<td>Tests</td>
</tr>
<tr>
<td>Rules of Inference</td>
<td>Testing</td>
<td>Inference</td>
</tr>
</tbody>
</table>

Figure 12

The assumptions inherent in all of these system components have been derived from the physical sciences, and are based upon closed system concepts. Further, as an information model, the process by
definition posits that the only valid inferences which can be drawn from the process, are those based upon a strict interpretation of the hypothesis test(s).

The only communication which the model is logically capable of transmitting is therefore in the form of the inferred differences between the data samples.

**Systems Concepts in Education**

It may be useful at this juncture to pause and attempt to clarify some of the ways that the term "system" has been used in education. Educators frequently tend to be sloppy in the way that they adapt new terminologies which are really windowdressing for old processes under new names. For example, educators tend to use PERT (Program Evaluation and Review Technique), flow charting, and system analysis almost interchangably. If a superintendent draws a flow diagram of his administrative structure, or if a doctoral student analyzes the components of an instructional problem, both processes are sometimes referred to as system analysis.

I will briefly discuss and define three usages of systems concepts which have developed distinct conventions over the past several years: system analysis, system approach, and general systems theory. It is very important that the applications and the particular philosophical
history of each of these concepts be carefully analyzed, because the logical premises of the models have important implications.

The System Approach

This term may seem somewhat curious in relation to the term, system analysis, yet there is a distinctive set of conventions associated with it which are considerably broader in scope and more comprehensive. The system approach is commonly used as a method for planning and designing the characteristics of total systems.

The comprehensiveness of the system approach may be seen in the four phases of the approach used by Silvern:

1. **Analysis** is performed on existing information to identify the problem, identify existing elements and identify the interrelation.

2. **Synthesis** is performed to combine unrelated elements and relationships into a new whole.

3. **Models** are constructed which can predict effectiveness without the actual implementation of the system.

4. **Simulation** is performed which reveals alternative solutions.\(^1\)

Silvern has used this method extensively in the design of training models for educational and industrial organizations. It has proven to be a very functional approach in 1) analyzing the nature of instructional problems, 2) studying alternative solutions to the problem, and 3) designing systems based on the most feasible solutions.
Kaufman stresses that the system approach provides an useful approach to rational planning and decision making. He states that, "a system approach is both a process tool for more effectively and efficiently achieving required educational outcomes and a mode of thinking that emphasizes problem identification and problem resolution."[1]

He defines the system approach as a 6 stage strategy with the following steps: 1) needs are identified, 2) problems selected, 3) requirements for problem solutions are identified, 4) solutions are selected, 5) means are obtained and implemented, and 6) results are evaluated.

Kaufman emphasizes that careful planning is necessary to properly identify needs and goals. He suggests that, "By identifying required outcomes first, and then deciding about the most effective and efficient 'process', we forestall the likelihood of having solutions that do not meet the actual needs." He stresses that, "If we attempt to solve problems poorly defined, we are faced with 1) an infinite number of possible solutions, and/or 2) a situation in which we treat only the symptoms and never really solve the problems."[2] He describes the benefits of a needs analysis as providing a frame of reference which delineates the discrepancy between the goals of a particular educational context, and the actual conditions which exist in
regard to those goals.

There are many - perhaps most - situations in educational decision making where the problems involved in accurately identifying problems and deriving workable alternative solutions are almost insurmountable. The busing issue, head start programs, reforming the curriculum, humanizing institutional structures; these are only a few of the many examples where the need is great, but where the methods simply do not exist for providing facile, efficient problem solutions.

System analysis and designers frequently talk about the importance of such constructs as needs analysis, or program evaluation. In fact, in industrial and scientific settings, it is the accurate quantification of these important functions which are the most crucial ingredients for the success of a system. The precise analysis of needs, the accurate and controlled feedback and evaluative processes, the measurement of program output - these are the very criteria which make the success of industrial and scientific processes - but which are the downfall of educational and social organizations.

The point that I am making, of course, is that the systems approach will not be a panacea for educational problem solving until the appropriate information processes are developed to enable the methodologies to be realistically applied. For every system objective, for every rational planning and decision making criteria, there are a
necessary set of contingencies which can only be accomplished through adequate, valid, and reliable information.

**General System Theory**

The General System Theory (GST) is not necessarily a method for analysis or planning, but rather, is a conceptual tool for deriving theories and understanding of the principles of systems organization, the interrelatedness of ecological and organic systems, and the wholistic nature of man.

GST has primarily evolved from the thought of men such as Bertalanffy, Wiener, Boulding, and Buckley, as a result of what I see as 3 principal reasons:

1. To provide a conceptual framework for organizing and integrating complex organic and social systems.

2. To provide a wholistic image of man to replace the disintegration that has resulted from the proliferations of modern technology.

3. To develop a rational basis for understanding organized complexity as a replacement for the inadequate theories of scientific reductionism.

As Bertalanffy has stated:

So far, the unification of science has been seen in the reduction of all sciences to physics, the final resolution of all phenomena into physical events. From our point of view, unity of science gains a more realistic aspect. A unitary conception of the world may be based, not upon the possibly futile and certainly farfetched hope finally to reduce all levels of reality to the level of physics, but rather on the isomorphy of laws in different fields.
Bertalanffy does not imply that there are laws which apply uniformly across all fields. He is stating that there are principles of organization which apply to all levels of reality, though perhaps at different stages of complexity. GST theorists seem to believe that there are organizing principles which apply to all of life which are not merely teleology or witchcraft, but which are based in the very nature of organic revolution.

**Implications of the General Systems Theory in Education**

Within the framework of General Systems Theory, it isn't useful to isolate problems into narrow contexts and study only the variables which can supposedly be controlled. The systems approach suggests that problems be studied in their "wholeness" with as many of the inputs and influences from environmental settings included in the study as can possibly be identified or conceptualized. This viewpoint precludes the usefulness of the simple, mechanistic cause and effect models which have been so commonly used in research and evaluation studies.

In a recent volume on systems theory, Laszle describes the shortcomings of the traditional scientific approaches to viewing the relationships between man and the universe. He states:

> Early scientific thinking was holistic but speculative; the modern scientific temper reacted by being
empirical but atomistic. Neither is free from error, the former because it replaced inquiry with faith and insight, and the latter because it sacrifices coherence at the alter of facticity. We witness today another shift in ways of thinking: the shift towards rigorous but holistic theories. This means thinking in terms of facts and events in the context of wholes, forming integrated sets with their own properties and relationships . . . The systems view gives us a perspective for viewing man and nature. It is a mode of organizing existing findings in reference to the concept of systems, and systemic properties and relationships. 

This perspective presents a difficult challenge to educators who are used to thinking of problems in terms of linear relationships or in terms of two dimensional cause and effect factors. The ability to describe and measure "events in the context of wholes", and "sets with their own properties and relationships", will require a revolution in the way that research and evaluation problems are conceptualized and measured. The qualities of social systems, the nature of the effects of machine systems on human values, and the wide range of individual differences between human beings, are all extremely important factors in educational development which must be adequately described in order to make meaningful decisions about schooling effects. These complex qualities of educational systems are the very qualities which must be translated into valid and reliable information, for as Trzebiatowski states, "information is the glue that holds a system together". 

15
However, when applying systems concepts to organizational objectives and information processes, educators must be careful interpreting technical methods and procedures. For the information needs of a system depend upon the particular objectives and context within which the system exists. Trzebiatowski suggests that educational systems should not be conceptually described in the same manner as static, engineering-oriented systems where the "elements are discrete and manipulable and therefore subject to operational prediction with engineering precision."16 There is a great temptation on the part of educational technologists to use terms such as "feedback", "cybernetic", or even the term "system", in a mechanomorphic sense. There is a confusion prevalent in modern society as to whether men are actually machines or not - as though humans and social institutions should function with the same engineering precision as machinery in a factory or instruments in a laboratory. Educational systems seem to lie much closer to the types of organizations found in social or political realms than to the tightly woven, highly controlled systems found in scientific organizations or industry.

Some very difficult problems arise when attempting to apply systems concepts to educational processes. Systems boundaries are hard to delineate, and it is hard to decide which system factors seem to influence or control inputs and outputs. In fact, because of the many
variable political, social, and community factors, it is not easy to say just who does control the process of education, even within a single classroom. From the standpoint of discussing the role of information flow in systems processes, the definition proposed by Trzebiatowski appears to be very useful. Quoting Ryan, he states:

First, the interaction of the system elements is dependent upon a common information network. Second, the function of the system is dependent upon the control of the flow of transmission of information. Third, the system can be characterized by the way it processes the available information.17

Using these three criteria, it becomes obvious that educational processes meet the principles of systems operations in only a general, rather imprecise manner. Under the first criteria, a common information network, it appears that systematic information exchange in education only applies to a few activities such as report cards and fiscal accounting procedures. The most important goals of education, the moral and patriotic concerns, or the effects of educational quality on the lives of the students, are simply too complex to be developed into systematic sources of information, at least up to this point. Obviously, important qualities such as teacher ability, program effectiveness, or the attitude of the students, are extremely difficult to quantify in a manner suitable for a systematic information system. It appears that in education, only the barest skeleton of a framework for using information in the sense of industrial cybernetic feedback is now
possible.

Under the second criteria, "the function of the system is dependent upon the control of the flow of transmission of information", there are undoubtedly many controls in effect in educational processes, but they are likely to be either haphazard or spontaneous, as in the snap judgements of teachers and administrators, or from the influences of various pressure groups. Many educational controls lie in the ponderous system of checks and balances between community groups, school boards, and state or national legislatures. Controls in education hardly exist with the brisk precision possible in business or industry.

Under the third criteria, "the system can be characterized by the way it processes the available information", the only truly systematic examples seem to be in the basic record keeping and the grade reporting departments. Otherwise, information exchange in education is a veritable free-for-all between the journals, texts, and publications of the professional organizations. There are some isolated examples of newsletters and curriculum research exchanges in some of the professional groups and regional laboratories, but these appear to be exceptions rather than the rule.

It is my belief that a new world view with the potential of combining the analytical power of science with the insight of humanism
may be possible through General Systems Theory. In contrast to the mechanistic approach to problem solving which prescribes that phenomena be reduced to their most simple terms and that events be interpreted by analysis of cause and effect relationships, General Systems Theory attempts to study problems from the perspective of wholism, interrelatedness of parts, and principles of development.

Unfortunately, General Systems Theory is no panacea which can be applied by humanists like a magic wand to transform the older mechanistic world views of science. As in so many scientific discoveries, the benefits of systems analysis and GST can be much more easily applied to the forces of profit and destructive power than to humanistic goals.

At this stage it must be said in favor of the humanists that the leading proponents of GST such as von Bertalanffy, Wiener, Buckley, and Laszle, are very much inclined toward the humanistic frame of reference. The fact that leading philosopher-scientists within the social science area such as Mooney, Rogers, and Maslow have espoused theories and views which fit so well into the organismic, man-in-the-universe view of GST premises a great potential that a new comprehensive model may be developed for understanding human development within a humanistic frame of reference.

Von Bertalanffy has stated that "Every science is a model in the
broad sense of the word, that is, a conceptual structure intended to reflect certain aspects of reality". He stresses that what is needed now are new conceptual models to provide a new perspective of principles of living systems such as purposiveness, organization, and directiveness. Von Bertalanffy stresses that instead of the traditional reductionism of science we need a new approach to understanding the unifying principles of life at all levels, which he terms "perspectivism".

In reference to this problem, he states:

The mechanistic world view, taking the play of physical particles as ultimate reality, found its expression in a civilization which glorifies physical technology that has led eventually to the catastrophies of our time. Possibly the model of the world as a great organization can help to reinforce the sense of reverence for the living which we have almost lost in the last sanguinary decades of human history.19

Essentially, GST provides a frame of reference for the study of living systems in terms of organized complexity, the interactions between a system and its environment, and the concept of a living system as a purposive organism with the potential increasing differentiation and adaptivity to the environment. In contrast to reductionism, GST recognizes the importance of studying a system in its wholeness, and of recognizing that the organizational complexity of a system cannot be broken down and analyzed into parts without serious danger of destroying the integrity of the system as a whole.
REFERENCES

1. Mooney, op.cit.,


4. Ibid., p. 76.


7. Ibid., p. 20.


9. Ibid.


12. Ibid., p.33.


15. Trzebiatowski, op.cit., p. 36.

16. Ibid., p. 37.

17. Ibid.

CHAPTER 7

Toward a Humanistic Frame of Reference

Introduction

The educational decision maker must operate within a complex matrix of activities which require communication skills of the highest level. The stereotype of the decision maker as one who issues edicts, or tightly controls a management system are passe', if they ever were viable. Although the decision maker may have the ultimate responsibility as an executive or manager, in a social frame of reference, his success depends upon how well he is able to stimulate and catalyze communication linkages. For, in a complex system, it is the quality of communication processes which integrates the system components.

As has been discussed in previous chapters, information needs within various types of systems differ widely. In a mechanical or industrial context, information processes are usually synonymous with control. In social systems, however, the most desirable type of information processes are those framed in the most humanistic context possible. Although many modern businesses are taking a much more
humane and democratic approach to organization, educators perhaps more than any other profession must be aware of the wide varieties and the many subtleties of communication.

There are varying views regarding the types of bonds which are most desirable for organizing societies. Some views hold that the most critical links of society should be inflexibly welded with the bands or coercion and the force of law, while other views maintain that it is much more desirable to interweave the fabric of society with many-hued threads, each representing a particular interest or cultural group. It is certainly true that a complex cybernetic machine, or an industrial complex, can only be integrated through tightly-controlled and efficient information systems. However, the manner in which a democratic society is integrated depends upon the quality of information transactions which are enabled to occur. As the American founding fathers so astutely foresaw, a democracy can only be as vital as the quality and effectiveness of free and open communications permit.

Conflicts and dysfunctions within societies are usually caused by some type of information barrier: either misunderstanding, or simple absence of communication channels. As has been discussed, Von Bertalanffy's definition of open and closed systems is analogous to the problem of information barriers, for a system is open to the degree that it is able to transact with its environment. Havelock compares static
and dynamic systems by the type of information exchange which occurs. A static system exchanges no information; a dynamic system exchanges information.

This leads to the central analogy of the dissertation: "What types of communication can logically occur as the result of research and evaluation methods derived from the epistemologies of science?" In describing one of the principal qualities of a static system, Havelock states, "in the context of knowledge flow the term 'barrier' is probably most expressive. Its defining property is that it stops messages."^{2}

If the context of education were considered to be a complex system, but one whose boundaries and interfaces were tenuous and difficult to delineate, we would need to ask: What would be the types of messages and communication channels which would serve to make the system more viable, more dynamic, more integrated? Would these messages logically be derived from the experimental methods of the physical sciences? From the statistical probability models of the gaming tables of 19th century Europe? From the psychometric measurement methods of the animal and behaviorist psychologists of the late 19th and early 20th centuries? Or, are we willing to admit that these borrowed problem solving models actually erect communication barriers between the designers and the clients of educational systems?
Setting Aside the Old Myths

As I have tried to show in preceding chapters, there is much that is admirable in many of the basic principles of the scientific method. In fact, scientists have long recognized the fundamental importance of such concepts as accurate observation, measurement, and reporting; perceptive inquiry, honest analysis, and the open-minded acceptance of new information which tends to disprove old hypotheses. It is my basic contention that no matter how admirable these premises have proven to be in the physical and biological sciences (essentially in closed system applications), their veracity in the social arena has not been born out. It is not that the principles themselves are wrong, for they are composed of the basic stuff of living human inquiry as pointed out by Bargar and Duncan. The problem is that the social sciences 1) have been so blinded by scientific empiricism that they have neglected to build their own viable and unique constructs; 2) have neglected to develop principles and procedures for inquiry which have honored the distinctive and unique characteristics which distinguish human psycho-social behavior from the inanimate and biological universe; and 3) have failed to develop both philosophical and operational constructs which justify the nature of social science inquiry.

Further, the social sciences will not progress so long as their
gaze is riveted on the past, or upon the reconstructed logic borrowed from other sciences. Obviously, new methods will not be developed as long as we fail to give attentions to paths that lie before us. For this reason, misconceptions about the nature of scientific activity must be revised, and misconstructions of the scientific method must be revitalized in the image of the unique needs of human problem solving. Following is a delineation of the myths that need to be shed in order to begin the much needed reconstruction of social inquiry.

1. **The Context of Social Decision Making is Rational.**

   For the most part, the sciences have developed in an atmosphere of scholarly isolation, apart from the conflicts of society. By definition science is a rational activity, where problems are studied in a controlled environment. If experimental variables cannot either be controlled or accounted for in a rational manner, deterministic laws of cause and effect cannot be applied.

   Although the issue can be debated, it is extremely difficult to conceive of a social situation where the rational conditions of determinable cause and effect relationships prevail. It does little good for researchers and evaluators to use scientific methodologies as if social conditions were rational, when indeed they are not. This is not only an irrational act in itself, but the uncontrollable and unexplainable phenomena within the social context create conditions which negate
conditions which negate the logical deductions required in the scientific method.

Further, decision makers in the social context do not profit by being reminded that their problems appear illogical and unsolvable. What they really need is insight, understanding, and meaningful information to help them better cope with the conflicts and complexities of which they are already well aware.

2. Social Constructs are Discrete and Manipulable.

The foundations of science are the control and manipulation of variables, or at the least, sciences must be able to specify, classify, and categorize the types of phenomena with which they are dealing. Again, this issue is highly debatable, but it appears that the only discrete and manipulable variables that the social sciences have been able to construct are actually artifacts such as occur within laboratory experiments, or the "measurement" of attitudes by various scales.

If variables are not discrete and manipulable, than a nomothetic science is not possible. This is obviously the condition which prevails in clinical psychology and anthropology, for example, and it is certainly the condition which applies to the broad spectrum of psychosocial behavior. It is one thing to speak in terms of descriptions and principles of behavior; it is quite another to speak in terms of nomothetic nets involving specifiable cause and effects, as the behavioral
sciences have attempted to do in relation to the study of humans.

3. **Social System Parameters and Interfaces are Palpable.**

In order to describe and understand social systems, it is important to be able to specify the types of system interrelationships and interconnections. Yet, in a democracy, many of the system interconnections are purely in the form of information exchange. These linkages are extremely difficult to specify and measure. The prevailing models for identifying system components in terms of information exchange are derived from biological, mechanical, and cybernetic systems where information flow can either be experimentally controlled, or at least, observed and quantified.

Information exchange in human social interaction is very difficult to identify, let alone measure with precision. Yet these very nebulous and tenuous interactions are the foundations for all human communications, and for the democratic concept of education itself. How then are the information exchanges and measurement constructs of science and technology to be applied? The communication transactions of the psycho-social world are best exemplified in the arts, humanities, and the many types of personal and social interaction. These communication qualities have yet to be codified into the statistical language of science.
4. The Dynamics of Human Interaction are Linear or Cybernetic.

Human personal and social transactions are more closely related to the concept of field theory, than to the types of interaction which occur within closed or techno-cybernetic systems. Human personal and social interactions appear to be explainable in difficult to define terms of flux within a field, but the concepts of linear or circular cybernetic explanations do not seem to be sufficiently descriptive. For example, what linear or statistical constructs sufficiently explain the effects of a symphony or a drama? These events are not metaphysical - they are real - yet what concept of scientific explanation begins to describe such an event? More simply, how is one to describe the affect of conversing with a friend - is it something that can be described or measured in terms of statistical cause and effect? Or is the basic nature of human social interaction more closely allied to that of being immersed in a field of energy - musical energy, dramatic energy, the energy of interpersonal dynamics?

Although the sciences have historically scoffed at esthetic experiences and subjective feelings as either metaphysical or illusory, the advent of modern field theory suggests these types of experiences are the real stuff of life. It is the artifactual nature of scientific measurement that is illusory.
5. Science is the Paragon for Problem Solving.

The social sciences have placed the scientific method on such a high pedestal that they have helped strengthen the dogma that not only is the scientific approach the only acceptable means for social problem solving, but that scientific information constructs are also the only objectively valid source for providing communication linkages. Numerous authoritative sources cite the basis for any dissemination or diffusion paradigm as being scientific research.

In the scientific areas of study, this is certainly true. However, as I have indicated in other places in the dissertation, there are numerous types of transactions in the psycho-social world which simply are not amenable to either the problem solving techniques or the language of science. The social sciences would be much better advised to devote their efforts toward the creation of constructs which describe and reflect the nature of human reality in descriptive terms, rather than trying to force human behavior into the molds of science.

Further, this dogma denies that either human wisdom or important social inventions are possible without the scientific method. Yet, the history of human culture emphatically denies this conception. Many, many important social developments have taken place without the scientific method, such as agriculture, domestication of animals, and many forms of technology. The traditions of the arts and the humanities have
have been developed largely without the aid of science. Art, music, drama, and architecture are important examples of human creativity and social invention which have evolved independently of science. The concept of human liberty – the power of human reason itself are tributes to the wisdom and insights of the ages which occurred before the advent of formal science.

This is not to deny the benefits of science. Modern medicine especially has utilized scientific methods to great benefit. The issue, of course, is that human freedom, creative social inventiveness and science should not be antithetical, but complementary.

6. Science as an Empirialistic Mode of Knowing.

Science is certainly a rational form of thought. However, the process of evolution has created a variety of reasoning faculties and perceptual modes, which culminate in the reasoning power of the intellect. True, science is an important forum for fathering evidence and debating the nature of certain forms of reality. But in the hierarchical levels of nature, progressing through the physical, vegetable, animal, and human, science is limited in its scope for the explanation and understanding of evolution. As is evident in the traditional rules of positivism, empiricism, determinism, and reductionism, the scope of scientific inquiry and debate are strictly limited to the calculus of probabilities and to empirical observations of behaviors.
During the thousands of years of human culture, men have developed an astounding capacity to create and reason utilizing a variety of modes of thought. As Polanyi and Bronowski have articulated so thoroughly, statistical reasoning and laws of cause and effect compose only a part of man's mental capacities.4

Research and Evaluation as Communication

Hopefully, the foregoing rationale will help to set the stage for the possibility of viewing educational research primarily as means for communicating descriptions and reflections of the most meaningful types of educational constructs. As has been pointed out, the only types of educational constructs which may logically be portrayed by traditional methodologies are either abstract statistical estimates of population parameters, or probability estimates derived from non-parametric statistics. These constructs are additionally limited by the nature of the assumptions which must be met in order to logically apply the quantitative methodologies, which are almost always cognitive achievement test averages, or events which can be discretely quantified.

The difficulty is that researchers and evaluators, in their efforts to be scientific, have attempted to utilize the constructs and the precision inherent in the "sign" language of statistics and calculus rather than in the symbols and ideographs of the arts and humanities.
The following chart (Figure 14) depicts the most common measurement statistics utilized in education. The information output of each one of these techniques is a statistical term, which is in the language of science, a sign, denoting some specific quality or estimate of behavior. It is important to keep in mind that the nomothetic nets and the calculus systems of science have been developed so that a sign, though an abstraction, is isomorphic with reality; i.e., the sign can be translated into a phenomena or event with fidelity.

The difficulty in using this "sign" language in educational research and evaluation is two-fold. First, the sign itself is an abstraction with very dubious isomorphic qualities. Second, and even more importantly, these commonly used statistical abstractions are incapable of representing the educational values of most concern to contemporary educators. Descriptions of cognitive processes, pupil self-image, individual creativity, personal development - these are all very important qualities which cannot be described in abstract physical language. Educators need constructs which provide a basis for describing these goals in terms which are truly descriptive, rather than being statistical abstractions of averages or hypothetical population parameters.

My position is that the research or evaluation study should be considered as a component of a systematic information process. The
<table>
<thead>
<tr>
<th>METHOD</th>
<th>TYPE OF RELATIONSHIP</th>
<th>APPLICATIONS</th>
<th>INFORMATION OUTPUT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson &quot;r&quot;</td>
<td>2 sets of related scores</td>
<td>same group compared on 2 measures</td>
<td>correlation statistic</td>
</tr>
<tr>
<td></td>
<td></td>
<td>reliability coefficients</td>
<td></td>
</tr>
<tr>
<td>Spearman &quot;Rho&quot;</td>
<td>2 sets of ranks</td>
<td>comparing similarity of ranks</td>
<td>rank order</td>
</tr>
<tr>
<td>&quot;z&quot; statistic</td>
<td>between sample and population with known parameters</td>
<td>compare similarity of sample to population</td>
<td>Z statistic</td>
</tr>
<tr>
<td>significance of difference between correlations</td>
<td>relationship between 2 correlations</td>
<td>test of significance for 2 Pearson &quot;r's&quot; (see Guilford p. 190)</td>
<td>Z statistic</td>
</tr>
<tr>
<td>&quot;&quot; difference between proportions</td>
<td>relation between 2 proportions</td>
<td>&quot;&quot; for 2 proportions (see Guilford p186-187)</td>
<td>Z statistic</td>
</tr>
<tr>
<td>&quot;t&quot; test</td>
<td>relationship between 2 randomly sampled groups from same pop.</td>
<td>comparison of means of 2 groups</td>
<td>T statistic</td>
</tr>
<tr>
<td>independent samples</td>
<td></td>
<td>groups which are matched, or pre-post test groups means</td>
<td>T statistic</td>
</tr>
<tr>
<td>&quot;&quot; for correlated or related groups</td>
<td>comparison of 2 related groups</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;F&quot; ratio</td>
<td>comparison of variance of two groups</td>
<td>similar application to &quot;t&quot; test, but is considered slightly more powerful</td>
<td>F statistic</td>
</tr>
<tr>
<td>Analysis of variance</td>
<td>same as F, but for 2, of many groups</td>
<td>2 or more groups may be compared with several classifications, such as 2x3, or 3x4, etc.</td>
<td>F statistic</td>
</tr>
<tr>
<td>ANCOVA</td>
<td>same as ANOVA, but groups are &quot;equated&quot;</td>
<td>groups compared are &quot;equated&quot; on variable such as pre-test</td>
<td>F statistic</td>
</tr>
</tbody>
</table>
process could be developed from the standpoint of a system analysis of a particular event, and the logical information contingencies which should be developed to adequately describe the event. In its most basic form, the procedure would follow the premises of communication theory. The following model developed by Berio could be used as a frame of reference. He describes the communication process as a sequence of events leading from the message source, to the message content, to the message form, to the receiver, as follows:

Source --> Message --> Medium --> Receiver

(Figure 14)

The following is an example based on the illustrations in figure 15, using a Pearson "r" correlation:

Source --> Message --> Medium --> Receiver

Comparison of reading achievement with social studies test scores correlation written decision quantity report maker?

(Figure 15)

The point that I am trying to illustrate here is that when the research or evaluation report is considered as a component of a systematic information process, the output of the particular study becomes the
input into some type of decision process, whether it is the researcher interested in proving a theory, or simply a curriculum committee attempting to evaluate the worth of a particular method or program. The form of the output has two aspects, the actual statistic or quantitative component, and the form in which it appears, usually a written report.

If a formal procedure is followed, the only information output which can be legitimately interpreted is the statistic itself in combination with an hypothesis test. According to the hypotheticodeductive system which is usually followed in research and formal evaluations, only the objectively verifiable data which is reported in the study becomes part of the logical system for interpreting results.

My contention is that this procedure, no matter how well-based in scientific theory, and no matter how well-intentioned, is very limiting in its usefulness, and is actually sometimes dysfunctional. Following is my rationale for taking this position:

1. The relationship between the message and the source is very tenuous, because so few real life educational problems can be adequately described with these statistical concepts.

2. The degree of isomorphism between the actual source of events, and the quality of description that is possible with this model, is again, very low. This is simply because a statistic is not adequate to describe the complexity of educational contexts.

3. The necessary linkages between each of the four components of the model are so weak that the model is susceptible to a complete breakdown, which frequently occurs.
4. The methodology itself severely limits the type of problems which may be studied, primarily because of the limited applications of the statistical concepts.

5. The nature of the message and the form that it is reported severely limit the type of audience to receive information. The more complex statistical models are all but uninterpretable except to the most highly trained statisticians.

It would be foolhardy to state that educational research is a waste, or that evaluation is a useless activity, because of the basic weaknesses in present models. There are important studies; there are useful evaluations. But the studies which have been the most useful have also been the most descriptive, and have managed to convey important information concepts in spite of the inadequacies of statistics and research models as communicative devices. Research and evaluation must be useful, not because they are "scientific" or "empirical", but because they provide meaningful information which is truly descriptive of the type of real life, complex issues that most concern educators and the lay public.

Critical Differences Between Research and Evaluation

Up to this point, research and evaluation have been discussed as though both concepts were capable of playing an equal role in providing information for decision makers. However, before research can be equated in potential utility with evaluation, some critical distinc-
tions must either be clarified or reconciled. Research has traditionally been associated with the scientific epistemologies of establishing theories and laws, and of determining cause and effect relationships. The definition of evaluation is a much more simple matter: Evaluation is the process of determining what people value.

The difficulty with the traditional applications of research methodology is that they really do not yield information upon which people can attribute value judgements. In fact, value judgements are a contradiction in research, because both the purpose and the methodology of research are supposed to be value-free, in keeping with the traditions of science.

As a practical example, I am reminded of a student with whom I worked on a thesis project. She was a county curriculum supervisor at the elementary level and was interested in comparing two different reading programs. As she wanted to do a "quick" evaluation, she collected data from two different schools over a period of three years. Each of the schools had been using one of the two reading programs. The best she could do in the way of random assignment was to randomly select three classes from each of the schools. Although the methodology was questionable, she computed an analysis of covariance with the data. An I.Q. measure was used as the covariate, and the reading test scores served as the criterion measures.
 Needless to say, the analysis of covariance yielded no signifi-
cant difference between the sets of reading scores in the two dif-
ferent schools. She was quite disappointed, both because she had
worked so hard gathering and processing the data, with no results;
and because there was no way that the data could provide her with in-
formation to attach a relative value to either one of the reading pro-
grams.

But the point is, neither the methodology inherent in educational
research techniques, nor the statistical techniques used, can logically
be expected to yield evaluative data. Even if the statistical hypothesis
test does indicate significant differences between comparison groups,
the data cannot be interpreted in value terms. The statistic itself is
nothing more than an estimate of difference based upon probability.

Critical value questions would more likely be:

1. Which program seemed to motivate the children best?

2. What were the comparative indicators of interest ex-
   pressed by the children?

3. What did the parents think of the programs?

4. What did the teachers think of the programs?

5. What were cost factors, efficiency, ease of use?

These are the type of questions for which the research paradigm
is incapable of providing answers. Even worse, there is an almost
insurmountable logical dilemma inherent in the comparative methodology
itself: How can you ask students which program they like the best, when they only experience one program? The experimental model does try to circumvent this problem by randomly assigning subjects to treatment. However, in using this approach two assumptions are made: 1) that the process of randomization does serve to equalize the two groups from the standpoint of ability; and 2) that any differential effects will be due to superior performance on the criterion variable of interest, not because of innate differences in the groups, or treatment which might contribute spuriously. Both of these assumptions in the comparative method are questionable at best.

There seems to be no logical method or statistical procedure which can circumvent the necessity for expert judgement in providing evaluative data in the educational context. The variables are too complex, the experimental conditions too tenuous for control, to assume that all of the essential assumptions of measurement and logic can be met.

Another logical difficulty in the research approach is that the model makes no pretense of accommodating either preconditions or process variables which highly concern educators. The research model seeks to either prove the validity of a construct, or to compare results due to differential treatment effects, all within a restricted time frame and a rigidly applied treatment method. There can be no room for flexibility, else the logic of the cause and effect is destroyed.
search model, because it is essentially comparative and value-free, does not seek to help the educator determine where his true values lie, explain important process effects, nor does it provide a frame of reference for determining the congruence between expectations and achieved results.

However, there is a new principle of evaluation which does attempt to provide just this type of information. It is based on the fundamental principle that human judgements are based not upon abstract notions of comparative data, or probability estimates, but upon the degree of congruence between what is expected or hoped for, and what actually transpires in the course of events. This movement in evaluation undoubtedly dates to the early 1930's when Ralph Tyler advanced the notion that educators could not systematically evaluate their program results until they had first specified their objectives.

As mentioned in Chapter 1, Stake has developed an evaluation procedure which gives the decision maker a framework for applying evaluative data at all of the critical stages of program development, each evaluative stage based on the principle of the degree of observed congruence between expected outcomes, and what is achieved (see Figure 3). Stake frankly states that often the only available data upon which to base these decisions is judgemental data.

Provus has also based an evaluation model upon the principle of
observed discrepancy or congruence between expectations and results. He states that evaluation should occur at five stages of program development, 1) design, 2) installation, 3) process, 4) product, and 5) cost:

At each of these stages a comparison is made between reality and some standard or standards. The comparison often shows differences between standard and reality; this difference is called discrepancy. On the basis of the comparisons made at each stage, discrepancy information is provided to the program staff, giving them a rational basis on which to make adjustments in their program.  

Guba and Stufflebeam have also proposed a model which provides a comprehensive framework for analyzing the logical components of educational programs. However, in spite of the fact that these authors have discovered very important principles for assessing programs, which appear to be obviously superior to traditional research models, the basic problem of developing and reporting accurate data upon which to base evaluative judgments, remains.

Although the logic of evaluation appears to have greater potential for providing decision makers with useful information than do the traditional research models, the research tradition has made valuable contributions to the principles of valid and reliable data. The problem of reporting accurate and relative data which adequately describe the issues of concern to educator must be given serious attention if the various approaches to evaluation are to prove viable.
REFERENCES


2. Ibid., p. 2-4.


7. Guba and Stufflebeam, op. cit.
CHAPTER 8

TOWARD A NEW PARADIGM

Introduction

Information processes in education must be generated from within a frame of reference that reflects the cultural and psycho-social complexity of human experience. Rather than ignoring or subverting manifestations of individual differences and values, these processes must honor individual dignity and serve to enhance humanistic goals. This requires that the information be of a character and fidelity that does not alter or distort the nature of human social processes, nor constrain the richness and diversity of cultural phenomenena.

The essential qualities of psycho-social experience are frequently distorted or obliterated by being described with inappropriate data forms, thus rendering the information not only invalid, but irrelevant. These vital attributes of human values, purposes, and culture simply cannot be viably transcribed into the closed system problem solving technology currently in vogue in science and industry. We now need to combine the rigor of science and the reason of humanism into a new
symbiosis which will permit the articulation and solution of difficult social problems which have heretofore remained enigmatic. It may be that educators and social scientists have been following a research and measurement tradition similar to the "normal science" paradigm described by Kuhn:

No part of the aim of normal science is to call forth new sorts of phenomena; indeed those that will not fit the box are often not seen at all. Nor do scientists normally aim to invent new theories, and they are often intolerant of those invented by others. Instead, normal-scientific research is directed to the articulation of those phenomena and theories that the paradigm already supplies.

Kuhn further suggests that normal science activities have both strengths and weaknesses: "A strength because the concerted efforts of scientists permit considerable expansion of a paradigm, and contribute to the abilities of phenomena defined by the paradigm to be measured." However, the sharply defined nature of the problems which are acceptable for study within a paradigm may also cause a certain narrowness of vision. Kuhn states in regard to the narrow problems defined by a particular paradigm:

Perhaps these are defects. The areas investigated by normal science are of course, minuscule; the enterprise under discussion has drastically restricted vision . . .

There are numerous indications that educational research and evaluation may be undergoing a transformation into a new paradigm—
away from a tradition which has heavily emphasized experimental models and quantitative assessment techniques--toward approaches which recognize the need for more naturalistic and wholistic views of educational contexts. As Kuhn describes the transition from older to newer ways of conceptualizing:

Led by a new paradigm, scientists adopt new instruments and look to new places. Even more important, during revolutions scientists see new and different things when looking with familiar instruments in places they have looked before . . . paradigm changes do cause scientists to see the world of their research differently. 3

Kuhn emphasizes that it is only in the world of textbooks and popular journalism that science is pictured as a unified, monolithic structure which coolly and objectively attacks problems with universally accepted procedures. This is actually a facade, according to Kuhn, which disguises the foment and the debates as various fields from time to time, drastically change their perceptions and their methodologies. In discussing textbooks "with both the popularizations and the philosophical works modeled on them," Kuhn states that they "address themselves to an already articulated body of problems, data, and theory, most often to the particular set of paradigms to which the scientific community is studying at the time they are written."
At times of revolution, there are frequently acrid debates, as the scope of problems and the acceptable methodologies gradually change. There is frequently a generation gap, with the younger scientists failing to understand why there elders cling to older, outmoded traditions. During these periods, as Kuhn states:

... at times of revolution, when the normal-scientific tradition changes, the scientists' perception of his environment must be re-educated—in some familiar situations he must learn to see a new gestalt. After he has done so the world of his research will seem, here and there, incommensurable with the one he had inhabited before.*

Kuhn is cited here, not because I believe that education is in a position to violently overthrow all of the past research and measurement models in favor of newer, more humanistically oriented approaches, but because so many educators and social scientists have slavishly followed some mythical notions that there are "scientific methods" which are the most respectable and authoritative ways to identify and solve educational problems.

Educators have tended to follow certain research and measurement techniques simply because they were quantitative, or somehow met some vague criteria for being scientifically acceptable. Yet, as Kuhn so articulately points out, there is no single uniform standard for what is either validly scientific, or for defining the types of
problems and assessment techniques which permit a problem to be studied within a particular paradigm. The history of science, according to Kuhn, is therefore not an unbroken chain of triumphs, each of which has been based upon the uniform application of fundamental principles of discovery. Rather, science has consisted of numerous revolutions and counterrevolutions as paradigmatic practices have tended to change with new discoveries, new measurement techniques, and new conceptualizations of the nature of the universe.

The misconception of scientific method in the social sciences has resulted in the nearly universal adoption of only one small part of the history of scientific logic - the hypothetico-deductive system. As von Bertlanffy has clearly illustrated, the positivist tradition presupposes an orderly universe whose components fit together with the logic of a cross-word puzzle, and whose phenomena are describable in the primary terms of statistics or calculus. Because this world view is only concerned with the analysis and solution of closed system problems, open system issues involving complex personal or social interactions are either denied, or simply dismissed as metaphysical.

Educators are now in the position of either persuading the social sciences to broaden the scope of their methodologies, or of striving to create new research and evaluation procedures which honor the
distinctive, autonomous qualities of psycho-social experience. We cannot pretend that this will be simple, nor that solutions to our problems will be easily found. Indeed, consensus may not be possible or desirable, for one of the critical differences between science and the humanities is that scientific problem solutions demand definitive answers, whereas in the humanities, diverse solutions and a variety of viewpoints are considered a virtue.

The balance of this chapter will be devoted to suggestions for resolving the great issues between science and the humanities as they bear on educational research and evaluation, and for presenting what is termed "A Rational Framework for Analysis and Evaluation of Problems in Education".

The Great Issues

Before the profession can embrace a new paradigm for educational research and evaluation, many issues must be resolved related to measurement and methodology. No matter how open minded the researcher, or how practical the evaluator, the great issues of valid and reliable reporting of events remain. The main reason why science has played such an important role in the history of Western man is through the sheer power of communications. Scientists have developed effective processes for describing, evaluating, and reporting information
so that the results of discoveries and experiments may be shared or replicated by anyone able to interpret the appropriate signs and symbols which comprise scientific reporting systems.

Within the parameters of closed system epistemologies, the hard sciences have solved the problems inherent in the great issues with tremendous success: 1) accurate description, 2) valid and reliable measurement, 3) attribution of cause and effect, 4) prediction, and 5) communication. The overarching rule which embodies all of these principles is the empirical mode - wherein statements or observations must be objectively verifiable.

Social scientists have correctly interpreted that these great issues are analogous, whether physical phenomena or human events are being studied. However, the social sciences have been remiss in not applying the supreme principle - the empirical test of utility - to the results of science in social applications. Seemingly, the lesser principles have been so busily applied out of sheer dogma, that the critical importance of simple tests of relevance and utility have been disregarded.

Thus, it is my conclusion that generally, the social sciences have seriously misconstrued the intent of scientific methods, and in so doing, have not been scientific at all. Rather, they have attempted to apply borrowed methodologies out of sheer faith - simply because
they have performed so powerfully in other applications. For, without the application of the empirical principle to test the utility of results, the social sciences have only partially applied scientific principles. Without the ultimate test of practical usefulness the logical completeness of the scientific method is not brought to bear.

Nonetheless, the powerful mystique which has been attributed to the methods of science by social scientists must be dissipated before many practitioners will seriously entertain research and evaluation methods which do not rely on inferential statistics or formal hypothesis tests.

Further, the powerful logic with which the physical sciences have been able to solve the problems inherent in the great issues quite naturally suggests that many of these same solutions might be applicable within the social domains. However, apparent analogies and isomorphisms must be analyzed very carefully. It is my own belief, which I take to be shared by numerous other writers, many of whom have already been cited, that scientific methods and mechanistic measurement models are generally inappropriate and dysfunctional when applied in social domains.

Note that the purposes and the general principles of science are specifically not being impugned. It is the application and the utility of specific methodologies and philosophical assumptions which are
being castigated.

Scientists deserve ample credit for developing the powers of perpectivity, inquiry, reasoning, and logic, which have such useful results within certain domains. However, for every important principle which has proven to be valid within the parameters of closed systems, there are valid epistemological reasons for stating the principle is inappropriate when applied within open system contexts. Human behaviors within the psycho-social domains are here stated to comprise these particular open system contexts.

Accurate Description

A scientific model is isomorphic with reality to the degree that it accurately reflects the essential qualities it has been designed to represent. To a considerable extent, modern civilization has developed in all of its complexity and diversity because science and technology have been able to communicate with signs and symbols which though abstract, may be transformed into accurate representations. In the physical sciences, events and phenomena may be described in terms of discrete, highly specific attributes.

All of the great principles of science depend upon the assumption that specified attributes will remain consistent over time, within certain known or predictable limits. Description in the psycho-social realms is no less important than in the physical sciences, but the
events described are idiographic rather than nomothetic; dynamic rather than stable; heavily value-laden rather than value-free.

Therefore, the same logical and philosophical assumptions cannot be easily applied across disciplines. As Barzun has so succinctly stated:

The social sciences today have yet to show one universal element or controlling 'law', one unit of measurement, one exactly plotted universal variable, or one invariant relation. The ologies in their latest avatar are not so much young as unborn.

Valid and Reliable Measurement

Without the power and precision of measurement the world of science and technology as we know it could not exist. The ancients were adept at some forms of measurement, such as astronomy and civil engineering. In the modern world, the technology of power, electronics, transportation, and machine tools simply could not take place at present levels without precision measurement.

In science and technology, the concepts of validity and reliability are much simpler than in the social sciences, because tests of utility, sufficiency, and accuracy are much easier to conduct. A scientist or technician rarely bothers to report information unless he is reasonably certain that his instruments are functioning reliably and that they accurately portray the phenomena under observation.
The issue of measurement is undoubtedly the most thorny of all questions in relation to the application of the scientific method in the social sciences. As in the question of the meaning of probability, the social sciences are rarely explicit in defining the meaning of measurement. The purpose is very clear - to quantify events and phenomena so that the paraphernalia of inferential statistics and formal hypothesis tests may be utilized. However, it doesn't seem that measurement in the social sciences has a meaning similar to the physical sciences, as defined by Caws:

Measurement is the assignment of particular mathematical characteristics to conceptual entities in such a way as to permit (1) an unambiguous mathematical description of every situation involving the entity and (2) the arrangement of all occurrences of it in a quasi-serial order.6

Measurement in the social sciences clearly is for the intent of ascribing certain attributes to phenomena so that magnitudes or relative values may be assigned. The critical puzzle to be solved is whether or not the measurement of psycho-social events is sufficiently analogous to the epistemology of physical science measurement to warrant the use of inferential statistics and quantitative or probabilistic hypothesis test procedures.

Handy has devoted a complete volume to the topic of, The Measurement of Values, without quite defining the empirical usefulness
or validity of assigning metrics to value constructs. Handy bases his case more on the need for measurement to satisfy the purposes of science than on questions of usefulness or relevance. He states that the importance of measurement "in scientific inquiry stems from the facilitation of predictions; i.e., successful quantifications allows more accurate predictions than otherwise would be possible."  

As appears to be the case in most social science measurement assumptions, Handy's purpose is closely allied to the general problem of creating an epistemology which justifies the use of metrics and scales for methodological, rather than empirical reasons:

In general, then, much depends on whether scientific inquiry is viewed as a human effort designed to facilitate prediction and control, or whether it is judged in terms of the formal perfection of the 'logic' of inquiry.  

Although Handy insists that the former purpose is related more to his own approach than the latter, his volume does not convince. His approach is similar to that which seems to prevail in most of the social sciences - a confusion as to whether measurement is intended to ascribe discrete, specific metric values to phenomena, in which case statistical inference and laws of probability would be appropriate - or whether interval and ratio data assumptions cannot be met, in which case, measurement would only be appropriate as a general descriptive tool in the nominal and ordinal data categories.

However, the question still remains as to whether psycho-social
events validly apply to interval or ratio data assumptions. The conclusion to be drawn by the viewpoints presented throughout this paper is that measurement is only as useful as it is isomorphic with the reality being described. If psycho-social events cannot be adequately described by the use of statistical data, then other communicative modes must be utilized.

**Attribution of Cause and Effect**

The purposes of description and measurement in the nomothetic sciences have been primarily linked to the concept of cause and effect. By precisely describing inputs, controlling processes, and measuring outputs, scientists have developed marvelously sensitive and complex procedures for discovering cause and effect relationships. As has been discussed extensively, this tradition has culminated in the hypothetico-deductive system.

Educators and social scientists alike are concerned about analyzing cause and effect relationships. In order to understand and improve social processes such as education, it is critically important to be able to identify variables which influence behavior. The scientist and the business executive are able to predict and control events through the many technologies of experimentation, prediction, and statistical methods. It is only natural for educators and civil servants to yearn for similar methods.
Unfortunately, the social sciences have interpreted the power of these technologies to lie almost exclusively in the domain of experimentation and statistical hypothesis tests. Thus, the all-important issues of analyzing the humanistic aspects of problem solving and decision making have been neglected. Educational researchers and psychologists in particular have overlooked the important implications to be derived from fields such as law, politics, social debate, and the humanities for problem solving models.

If measurement and experimental control assumptions are invalid when applied to psycho-social phenomena, then the hypothetico-deductive system itself is invalid as a formal model in the social sciences.

**Prediction**

Prediction is the foundation upon which science and technology have been built. Through the historic processes of identifying, describing, classifying and measuring most of the known physical elements of the world, scientists and engineers have been able to develop a great reservoir of knowledge with which to predict cause and effect relationships. The hypothetico-deductive system has played a large role in this historic development.

Again, the social sciences have attempted to adopt the hypothetico-deductive system as the principle model for determining cause and
effect relationships with either the logic or power to predict future transactions.

Numerous reasons have already been cited concerning the lack of utility of this particular approach. These may be summarized as:

1. Because the model was evolved to serve physical science needs and assumptions, where events may be controlled and measured with relative precision, it has proven to be awkward and of questionable utility in the social arena, where similar conditions are difficult to apply.

2. Because the model only frames questions which may be phrased in terms of specific statistical or mathematical formulae, it is irrelevant for issues which cannot be similarly presented.

3. Perhaps most important of all, the logic of the model is usually based upon testing hypotheses in terms of statistical probability. The model does not lend itself as presently constituted, to providing evaluative information upon which qualitative judgements may be formulated.

Communication

Because the subject matter of the scientific-technical fields have dealt primarily with closed system concepts, both the forms and means for communication are logically and philosophically unrelated to open system contexts within the social arena. The essence of the psycho-social world simply cannot be reduced to the analytical terms and statistical dimensions of the physical universe. From my perspective, it is a logical fallacy to assume that many psycho-social qualities such as cognitive knowing, esthetic exper-
ience, and interpersonal transactions, can in any meaningful sense be converted into the language of the physical sciences.

The task of the social sciences is not to reduce these qualities to physical science terms, but to understand that these are the stuff which comprise the psycho-social world. Scientists have learned to isomorphically transform information concerning the physical universe into codes and formulae which may be translated from the concrete to the abstract. Similar transformations do not seem possible for the various types of cognitive, esthetic, and evaluative behaviors which form the world of the mind.

Therefore, the challenge facing the social sciences is to develop communication systems and information processes which are based upon the valid constructs and viable descriptions which honestly reflect the essential attributes of psycho-social processes. These processes are undoubtedly much more closely related to the forms of communication which compose the fields of drama, literature, music, and the visual arts, than the statistical, probabilistic modes of the physical sciences.

A Rational Framework for Analysis and Evaluation of Problems in Education

The approach that is proposed here is more closely allied to information systems and communication models in business, the
humanities, operations research, and the political process, rather than to the formal models of science. The basic principle being proposed is that people have the potential of solving their problems through logical analysis and reasoned debate. The pivotal issue is that of communications, for people must learn to describe their concerns to each other and develop processes for resolving conflicts and deriving problem solutions.

This is not an easy task, one of the difficulties being that the fields of science and technology have themselves tended to create thought patterns and monolithic structures which are derogating the fundamental principles of Jeffersonian democracy. As Matson and Montagu point out:

... the field of communication is today more than ever a battleground contested by two opposing conceptual forces - those of monologue and dialogue. The 'Monological' approach, which defines communication as essentially the transmission of symbolic stimuli (messages or commands), finds its classical formulation in the art and science of rhetoric and its characteristic modern expressions in cybernetics, combative game theory, and the repertoires of mass persuasion. The 'dialogical' approach, which regards communication as the path to communion and the ground of self-discovery, found its original champion in Socrates and has its spokesmen today in such diverse currents of thought as religious existentialism, post-Freudian psychotherapy, and sociological interactionism.9

A second important consideration is that of the empirical principle itself - applied to the utility of methodological approach,
rather than as a nebulous adjunct to the subject of inquiry. Educators and social scientists must learn to apply the test of utility to both purposes and methods. For example, the critical test to be applied to any problem solving methodology might be: "How well does this method serve the purpose for which it was intended?" In this same manner, the test of utility should be periodically applied to every phase of educational systems in order to assure that the most important needs of clientele are being met satisfactorily.

To reiterate, the information and communication needs proposed for this approach do not necessarily reflect the logical dictates of classical science models, which have primarily focused upon issues of probability, prediction of cause and effect, and statistical hypothesis tests. These models are presumed to have failed the tests of utility which have been applied through the experience of the past 70 years in the various social sciences. The evidence for this presumption is:

1. The critical statements of numerous eminent educators and philosophers: many have been cited in this paper.

2. The obvious lack of utility of the models for practitioners, who generally seem to ignore or abhor research.

3. The inconsistency of results within paradigmatic practices. Indeed, with the exception of the graduate school curricula around the country, it is highly questionable whether any social science research paradigms have attained the stature of paradigms within other fields, such as medicine, law, or architecture.
The Proposed Model and Its Rationale

The model proposed is not new, but is based upon a synthesis of approaches borrowed from principles of inquiry and analysis which have been developed in system analysis, operations research, educational planning, and management information systems design. Some of the leading writers in these fields are Cook and Kaufman in educational planning, Nadler and Silvern in system analysis, and Murdick and Ross in management information systems.

Following is a general overview of the stages which comprise the model.

Figure 16

System Analysis and Evaluation Model

<table>
<thead>
<tr>
<th>Phase</th>
<th>Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Logical Analysis</td>
<td>1.1 Define the nature of the decision context.</td>
</tr>
<tr>
<td></td>
<td>1.2 Identify decision making needs.</td>
</tr>
<tr>
<td></td>
<td>1.3 Identify system clients and the range of their needs.</td>
</tr>
<tr>
<td></td>
<td>1.4 Project idealized scope of services to provide client's needs.</td>
</tr>
</tbody>
</table>
### FIGURE - Cont.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Strategic Analysis</td>
<td>2.1 Identify system resources and constraints</td>
</tr>
<tr>
<td></td>
<td>2.2 Analyze range of client services feasible to provide with available resources.</td>
</tr>
<tr>
<td></td>
<td>2.3 Analyze alternatives for providing client services in terms of cost-effectiveness.</td>
</tr>
<tr>
<td></td>
<td>2.4 Evaluate Alternatives</td>
</tr>
<tr>
<td>3. Test and Development</td>
<td>3.1 If feasible, construct trial program on pilot test scale.</td>
</tr>
<tr>
<td></td>
<td>3.2 If full scale test is not feasible, test program for logic and completeness with critical decision makers.</td>
</tr>
<tr>
<td></td>
<td>3.3 Implement or revise, according to test results.</td>
</tr>
<tr>
<td>4. Evaluation and Feedback</td>
<td>4.1 Evaluate program impact according to criteria.</td>
</tr>
<tr>
<td></td>
<td>4.2 Compare program effectiveness with original objectives.</td>
</tr>
<tr>
<td></td>
<td>4.3 If revisions appear necessary, re-evaluate objectives in terms of clients' responses and needs.</td>
</tr>
</tbody>
</table>
The basic rationale for this approach lies in the logic of empirical testing - to empirically analyze the purposes and goals of a system, to analyze the reality of the decision making context, to determine the roles of decision makers and their actual information needs, and to evaluate the performance and utility of the programs which are intended to satisfy client needs. The critical problem in this approach is to develop and provide only that information which is needed to meet the system objectives.

This rationale is almost completely at odds with traditional approaches to providing research and evaluative information. Social scientists - educational researchers and psychologists in particular - have taken the stance that the goals of education and the manner in which decisions are made should be decided from the standpoint of what is scientifically rational. Unfortunately, this stance completely ignores the political reality of educational decision making.

At least three critical assumptions are at issue. First, it is assumed that decision makers who are the real gatekeepers of power are interested and capable of weighing decision alternatives on the basis of scientific information. Second, it is assumed that traditional research and evaluation models are capable of providing information of a scope and quality which would actually improve education. Third, and most critical because it embraces the total logic of the scientific
method, it is assumed that the rules and procedures developed from the epistemology of science actually have the potential for advancing societal goals for education.

These statements are not intended to deprecate the abilities of either educational researchers, decision makers, or scientists. The painful truth is that educational decision making usually occurs within a matrix of many politically opposing forces, so that decisions cannot be carefully deliberated within the alleged coolness of scientific rationality. Following is a brief discussion of the rational basis for the analysis and evaluation model.

Logical Analysis

Researchers and evaluators simply cannot assume that they know in advance what the information needs are for their primary clients—decision makers, or for their secondary clients—the particular public receiving educational services. In higher education especially, it is usually assumed that professors, theoreticians, and other sundry experts know more about the needs of students than administrators, school boards, and legislative bodies. This may or may not be the case. The logical analysis of the real decision making context, however, seems to indicate that it is these last mentioned groups who actually have direct access to students. Further, these decision makers will exercise the most direct control over content and
methods for the foreseeable future.

The primary intent of this phase is to study rationally, and without preconceptions, the dimensions of the actual context within which decisions are made for a particular problem. If the needs and operations of decision making cannot be operationally analyzed in terms which can be met by the researcher or evaluator, then there is little point in trying to develop any kind of rational framework. This analysis can only be performed through empirical inquiry.

There are procedures which have been developed specifically for this purpose in management information system science, in the field of business. Beginnings have also been made in this direction in the developing field of educational needs analysis. Coffing and Hutchinson, for example, have developed an approach which they term "client demand methodology".

**Strategic Analysis**

The purpose of this phase is to develop and analyze alternative strategies for providing client services. In an educational program, this might entail alternative approaches to instruction. In providing research or evaluation methodologies, this would entail alternative ways of providing needed information for decision making. It is critically important that the information users and providers work together to analyze the dimensions, costs, and potential effectiveness
of the projected alternatives for providing client services.

At this point, the researcher or evaluator should work with his primary clients to devise a test for completeness, which might be applied either from the standpoint of logical consistency, or empirical validity.

Test and Development

How many doctoral studies, research designs, or evaluation reports have been performed without testing either the logical or empirical validity of the measuring and reporting instruments? It is very typical for a researcher or evaluator to design some type of measuring or reporting procedure, carry it out, and report the results, with very little empirical basis for the utility or validity of the process. A scientist or engineer wouldn't seriously consider such a foolhardy step for fear his professional reputation would be damaged for life.

The same problem applies, of course, to the need for testing instructional programs before they are implemented. Millions of dollars have been spent on programs such as foreign language laboratories, with little empirical basis for assuming that the particular method is even worth the investment. This is not to say that empirical testing is ever easy. The typical tests for reliability and validity of achievement measures are usually circular and
tautologous.

**Evaluation and Feedback**

Researchers and evaluators have become very adept at reporting information concerning experimental results, or product performance, for example. The problem has been to report information which is relevant for real decision making contexts and which has utility for implementing programs. Social planning which includes the rational elements of 1) empirical identification of needs, 2) systematic weighing of alternatives, and 3) evaluation and revision of program results according to observed effects, is a problem of immense complexity.

It is probably safe to say that no local school board, state department of education, or federal bureau has yet succeeded in implementing such a rational approach. Frymier and Hawn take the position, relative to this problem that:

> Every effective social system reflects three phases of operation which accomplish separate functions that enable the system to maintain itself in a dynamic way (planning, effecting, and evaluating) . . . Any careful study of social systems other than education suggests that these three functions are relatively discreet, and that they are accomplished by different groups, each one of which has a realm of power.\(^\text{16}\)

These authors stress that, "**evaluation plays the critical role of providing corrective feedback to the other parts of the system so that the entire operation can be improved** (author's emphasis).\(^\text{17}\)
They emphasize that, "Looked at in terms of such a social systems model, education obviously has a conceptual flaw", because, "No aspect of the educational system regularly generates evaluative-type data; nothing in the concept requires that those within the system pay attention to the feedback data if they should appear."

They go on to summarize the problem:

Education cannot improve as presently conceptualized, and as it presently functions. This description of the reality, though, suggests exactly where we need to apply our energies if we really want curricula to be improved: program evaluation.

These authors have summarized a very complex social issue: for who is to decide evaluative criteria; who will conduct the evaluations, and who will analyze and effect constructive feedback?

The problem has immense social dimensions and must eventually be debated and decided by the democratic process itself, rather than making the false assumption that groups of expert researchers and evaluators should decide the issues by executive fiat.

The Problem of Linkage

Frymier and Hawn have proposed a set of logical steps which, if followed, should comprise a rational system approach for educational programs. As they suggest, however, present approaches to the design of educational programs are flawed, because provisions are not made
for applying evaluative feedback data for either summative judgements, program monitoring, or corrective action. These logical elements are crucial to the design of a rational system, of course. However, in order for these rational system purposes to be effectively realized, appropriate communication linkages must be developed which both promote the system functions, and integrate the system elements.

Viable and effective system communication processes are just as important in the functioning of a program as other resources such as manpower, money, and materials. In fact, system resources will likely not be effectively utilized without the appropriate communication systems for planning, monitoring, and evaluating performance. Because phenomena and events within scientific and industrial processes are usually so much more discrete and specific, information systems are naturally easier to develop and implement than in social areas.

This is the central issue in developing valid and reliable information, for unless the message being communicated accurately describes the event of interest, effective communication does not occur. Researchers and evaluators have typically overlooked this problem in their reporting procedures, accepting the assumption that statistics and experimental constructs should somehow meet the needs of educational decision makers, without actually taking the empirical steps
necessary to analyze the substance and attributes of the information as specified by decision makers themselves.

Following is a simple model which might depict this process, where the event of interest is described in terms of 1) basic information units, which become the 2) message, that is transmitted via a 3) channel to the 4) audience.

```
Information Unit → Message → Channel → Audience
```

**Figure 17**

The usefulness of the information reporting process must be tested both for utility and completeness by first, determining the type of information desired by the decision maker, second, analyzing the appropriate message and channel to report the information, and third, testing the effectiveness of the communication process. In the final analysis, the usefulness of the information is determined by the manner in which decisions are benefitted.

Within the scope of an educational system or program, the rational approach to planning, implementation, and program evaluation cannot be fully realized, therefore, unless, valid and reliable information processes are implemented as components of a comprehensive management information system. The purpose of the information should
be to facilitate and integrate the various program functions and elements, but this cannot occur unless the necessary resources are developed to gather, report, and utilize data. Figure 19 portrays an approach to such a model. The phases in the left hand column refer to the components of the analysis and evaluation model mentioned earlier, but would also apply to any function or system element. The object being that there must be a process for coordinating and integrating the elements and functions within a comprehensive framework of management. The communications elements are discussed as follows.

**Data Criteria** - The purposes for the data are specified in operational terms to insure that the whole communication process is firmly based upon the actual data processing needs demanded by the decision making context. For example, the decision maker might want to know about the cost, efficiency, and performance attributes of instructional materials. In which case, measures would have to be developed for each of these criteria.

**Data Source** - The source of the data is specified. This obvious step is necessary for the purposes of planning the communication system so that necessary steps may be taken to identify sources and organize necessary contingencies.

**Data Form** - The information provider and the user must agree
### MODEL FOR PROGRAM INFORMATION SYSTEM DESIGN

#### Program Phases 

<table>
<thead>
<tr>
<th>Program Phases</th>
<th>Communication Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Criteria</td>
<td>Data Source</td>
</tr>
<tr>
<td>1. Logical Analysis</td>
<td></td>
</tr>
<tr>
<td>2. Strategic Analysis</td>
<td></td>
</tr>
<tr>
<td>3. Test and Development</td>
<td></td>
</tr>
<tr>
<td>4. Evaluation and Feedback</td>
<td></td>
</tr>
</tbody>
</table>

**FIGURE 18**
in advance concerning the purposes and uses for the information to
insure that the data forms will be maximally useful.

**Linkage** - Resources must be developed to provide appropriate
and effective channels for the information. If these linkages are not
created, the system will simply not be able to utilize information to
full advantage.

**Data Review** - Depending upon the size and complexity of the
system or program, steps may need to be taken to organize procedures
for the review and evaluation of the data before it is presented to
decision makers formally. In many large organizations, data are
reviewed at lower levels of management and summarized - or in
many instances decisions are made programmatically according to
agreed upon criteria.

**Decision Cycle** - At this point, monitoring and performance data
are reviewed for decision making. Decisions might be made to revise,
accept, or discontinue the programs on the basis of evaluative data.

It must be emphasized again that comprehensive and well inte-
grated management information systems are quite feasible in scientific-
industrial settings where problem parameters may be stated with some
precision and where input and output relationships are discretely
quantifiable. Processes and events in education cannot be similarly
specified are quantified. This question of what constitutes valid and
reliable measurement or description upon which to base substantive
decisions is undoubtedly the most critical problem facing all those
concerned with educational program development.

SUMMARY

The central theme of the dissertation has revolved around the
issue of information and communication. The argument has been
that educational decision makers need information which is obtained
within a structure that provides useful evaluative data, bearing on
the deepest values which society holds for its educational goals.
These value questions reflect concerns related to the total context
of human development, not merely academic achievement.

Anyone who has spent any degree of time with elementary school
teachers is well aware that the main concerns relate to the wholistic
nature of children - their personal and social development - their
attitude toward school and their friends - and their level of motivation
and aspiration. These important concerns are not easily transformed
into the traditional repertoire of scientific-technical problem solving
models. Professionals in real world problem situations are not bothered
about questions of probability or statistical significance, yet they
are daily faced with decisions which require evaluative data. If
researchers and evaluators are actually to help practitioners solve
their most worrisome problems, new decision making paradigms must be developed which utilize information that is descriptive in the most humanistic sense.

Humanistic data comprises the visual arts, music, drama, audio-visual media, and the widest possible range of cultural experiences. Documentary films and educational television, for example, are among the most powerful communications media ever known. Are they any less valid as evaluative data than a statistical hypothesis test? If the social sciences are to participate actively in the dynamic growth of cultures, they must learn to forego many of their preconceptions concerning the purposes of science, and appreciate more fully the richness and diversity of human consciousness.

CONCLUSION

Many branches of the social sciences, including educational research have been guilty of uncritically accepting the assumptions of the positivist-empiricist tradition of science: that there is an extant body of empirical and quantitative research methods which may be universally applied over broad and diverse fields. However, this has been essentially an unwarranted extrapolation, for which their is little practical or empirical justification.

Somehow, the social sciences have refused to admit publicly
that they need a truly distinctive philosophy - an ordered knowledge system which recognizes the unique constructs and descriptive processes which must be developed that honestly reflect the complex values and interactions of the social world. Thus, research and evaluation methodologies are typically presented to graduate students and the lay public as empirically validated components of an allegedly unified epistemology of science. Experimental methods and psychometric measurement constructs are portrayed as the foundation stones upon which a science of education must be constructed.

Yet, as indicated in the previous chapters, a critical analysis of both the practical performance and the logical assumptions of the scientific-technical models commonly used in the social sciences, reveals these assumptions to be flawed. Although many eminent practitioners openly criticize these methods, there is still a problem in determining the directions to take in correcting these flaws. Stanley, for example, insists that research can only be improved by developing more powerful statistical analyses.\textsuperscript{19} Even someone of the statesman-like stature of Ebel has emphasized that the way to improved assessment lies in the direction of more refined measurement constructs, in the tradition of Thorndike.\textsuperscript{20}

I cannot agree that improving the quality of information processes in education is simply a matter of more and better of the same tradi-
tional assumptions. The answer is surely not in the direction of more complex statistical models and experiments. If the social sciences are to actively participate in the solution of pressing social ills, new problem solving methodologies based upon a humanistic frame of reference must be developed.

Educators in particular must reaffirm the primal principle that it is the growth of enlightenment and the power of reason which are the central reasons for educational development, rather than the narrow growth of special fields and the proliferation of scientific paradigms. Although science is one of the great paths on the road to reason, it is not the only one, and in fact, without being firmly rooted in the soils of humanistic goals, is likely to threaten, rather than contribute to development.

Many are beginning to question the exaggerated emphasis placed upon science and technology. The writings of intellectual leaders such as Maslow, Mooney, Matson, Mumford, Reich, and others attest to this movement. As explained by Huxley:

This new idea-system, whose birth we of the mid-twentieth century are witnessing, I shall simply call Humanism, because it can only be based on our understanding of man and his relations with the rest of his environment. It must be focused on man as an organism, though one with unique properties. It must be organized round the facts and ideas of evolution, taking account of the discovery that man is part of a comprehensive evolutionary process, and cannot avoid playing a decisive role in it.
REFERENCES

2. Ibid.
3. Ibid., p. 112.
4. Ibid.
8. Ibid., p. 41.
9. Matson and Montagu, op.cit., p.viii


17. Ibid.

18. Ibid.


20. Ebel, *op. cit.*

### Addendix A

**INFORMATION YIELD OF QUANTITATIVE MODELS**

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Constraints</th>
<th>Information Output</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parametric Group:</strong></td>
<td>All parametrics must be used with samples randomly selected from populations which are homogeneous and homoscedastic. Data must be interval or ratio.</td>
<td></td>
</tr>
<tr>
<td>a. central tendency</td>
<td></td>
<td>mean or average score.</td>
</tr>
<tr>
<td>b. comparison of means</td>
<td></td>
<td>estimate of difference based on ration of error to sample size</td>
</tr>
<tr>
<td>c. comparison of variance</td>
<td></td>
<td>similar to mean comparisons, but group variance is basis of comparison</td>
</tr>
<tr>
<td><strong>Non-Parametric Group:</strong></td>
<td>Same sampling assumptions as above are preferable, but not necessary. Data may be nominal or ordinal.</td>
<td>probability estimate of distribution of frequencies</td>
</tr>
<tr>
<td><strong>Correlation Group:</strong></td>
<td>Data must be interval or ratio. Power of hypothesis tests is related to sampling procedures.</td>
<td></td>
</tr>
<tr>
<td>a. two variables</td>
<td></td>
<td>estimate of linear relationship</td>
</tr>
<tr>
<td>b. multiple variables</td>
<td></td>
<td>estimate of linear relationship for pairs or multiples</td>
</tr>
<tr>
<td>c. factor analysis</td>
<td></td>
<td>patterns of inter correlations</td>
</tr>
</tbody>
</table>

**FIGURE 7**