POLICY IMPLEMENTATION: AN EMPIRICAL
INVESTIGATION OF THE HANDICAPPED EFFORT IN
HEAD START

Dissertation for the Degree of Ph. D.
MICHIGAN STATE UNIVERSITY
SELÇUK ÖZGEDIZ
1976

This is to certify that the

thesis entitled

POLICY IMPLEMENTATION: AN EMPIRICAL
INVESTIGATION OF THE HANDICAPPED
EFFORT IN HEAD START

presented by

Selçuk Özgediz

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Political Science

Major professor

Date May 6, 1976

O-7639

su

sp

in

re

of

to

Apr

Fra

per

The

ily

revi

tati

with

the e

ABSTRACT

POLICY IMPLEMENTATION: AN EMPIRICAL INVESTIGATION OF THE HANDICAPPED EFFORT IN HEAD START

By

Selçuk Özgediz

This study is on empirical explanation of the level of success achieved in implementing a new public policy. The specific policy examined is the so-called handicapped policy in Project Head Start which stems from a 1972 congressional requirement that at least ten per cent of the total number of enrollment opportunities in Head Start be made available to handicapped children. The fieldwork was conducted in April-May 1974. The sample consisted of thirty-five programs. The data were collected through on-site, person-to-person interviews with the directors of the sampled programs. The research design used for collecting the data was primarily a cross-sectional survey design .

The limited set of literature on implementation is reviewed, definitions for the terms "policy" and "implementation" are offered, and the place of policy implementation within the overall policy process is discussed.

The conceptual model of the study has its origins in the expectancy theory of organizational psychology. The

p
me
in
po
ab
fo
(E
bil
ind

FSI
ysi
for

is (
suc
inte
prog

the
data
incr
incr
CAPB
troli
The h
either

primary independent variable of the study is level of implementation success (LIS). The two variables used for explaining LIS are implementation potential (IPOT) and level of policy-specific effort (EFRT). The secondary dependent variable of the model is EFRT and the independent variables used for explaining it are environmental forces for implementation (ENVR), policy support (PSUP), and self-evaluation of capability (CAPB). All of these variables are measured through indices based on the survey data.

The effects of IPOT and EFRT on LIS and those of ENVR, PSUP, and CAPB on EFRT are studied through regression analysis. Two models, an additive and an interactive one, are formulated and tested for each relationship.

Implementation of the policy in the sampled programs is described and discussed. Observed differences between successful and unsuccessful programs are identified and interpreted and a profile of the implementation in a typical program is presented.

It is shown that there is strong empirical support for the overall conceptual model. Hypotheses confirmed by the data include the following: (1) controling for IPOT, LIS increases as EFRT increases, (2) controling for EFRT, LIS increases as IPOT increases, (3) controling for ENVR or CAPB or both, EFRT increases as PSUP increases, and (4) controling for ENVR, EFRT increases as PSUP x CAPB increases. The hypothesis: "controling for PSUP, EFRT increases as either ENVR or CAPB (or both) increase" is disconfirmed by

the data.

It is also found that there is strong empirical support for both the additive and the interactive models proposed for the two relationships. However, the criteria used in comparing the alternative models favor the two additive models over the interactive models.

Implications of the analytical findings for the handicapped policy in Head Start and for the study of policy implementation are discussed and a dynamic version of the model is recommended for future empirical testing.

POLICY IMPLEMENTATION: AN EMPIRICAL

INVESTIGATION OF THE HANDICAPPED

EFFORT IN HEAD START

By

Selçuk Özgediz

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Political Science

1976

ACKNOWLEDGEMENTS

I owe a great debt to a number of individuals for their support and assistance during different stages of the study. My committee chairman, Frank A. Pinner provided invaluable guidance and support throughout the study. The other members of the committee, Charles Cnudde and Paul Conn helped shape the dissertation with their constructive suggestions. I also owe them a great deal for their timely response to my inquiries.

The fieldwork portion of the study was conducted as a part of the Systems Research Incorporated-Syracuse University project entitled "Evaluation of the Handicapped Effort in Project Head Start" under a contract with the U.S. Department of Health, Education and Welfare. While co-managing this project I enjoyed the continuous guidance and support of Alan F. Bogatay, to whom I am also indebted personally and professionally. Several members of the Systems Research Incorporated project team provided valuable assistance during the conduct of the project. I am grateful to William Baucom, Gerard Fiala, Walter Sowles, Sandra Filion, Deborah Grether, Willard Walker, William Maas, and Sunni Harvey for their cooperation and expert assistance.

It was a distinct pleasure to work with Burton Blatt.

He

th

ke

an

te

ti

de

Gro

stu

ef

gra

ope

han

cla

lys

ful

cons

pers

grad

suppo

and e

full

the h

hers

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

This dissertation is about policy implementation.
More specifically, it is about how bureaucracies execute
legislative mandates.

During the last decade political scientists have paid
increasing attention to the question of how policies are
formulated. Indeed, the start of the policy analysis move-
ment in political science has marked the proliferation of
research on policy formulation. Yet, virtually no one has
paid attention to the question of what happens to a policy
after it has been formulated. This is quite surprising.
One would ordinarily expect a large body of literature in
an area that has such a direct and immediate effect on the
well-being of millions of persons. In the United States
alone billions of dollars are being spent annually on
hundreds of governmental programs; and each of these
programs is being implemented daily by one or more agencies.
The eventual success of these programs depends upon how well
these agencies can translate the original "intent" of the
program into "effect."

Policy implementation, in the sense we have used the
term here, refers to those tasks and activities that are

pe

al

to

ti

mu

aq

ac

to

al

ex

ti

th

in

po

He

re

sc

of

po

re

ti

in

cl

ap

performed after a policy has been formulated and operation-
alized. By operationalization of a policy we are referring
to formulation of rules, orders, and procedures and prepara-
tion of manuals and guidelines which indicate how the for-
mulated policy is to be executed by the implementation
agencies. Implementation, therefore, includes all tasks and
activities undertaken by the implementing agency in order
to execute the policy in the manner it has been operation-
alized. A successfully implemented policy is one which
exhibits a high positive correlation between the implementa-
tion objectives and the eventual outcomes.

Why study implementation? We can think of at least
three reasons: scientific, professional, and political.

From a *scientific* viewpoint we study implementation
in order to gain an understanding of how and why some
policies are implemented successfully and others are not.
Here the emphasis is on scientific explanation of the
results of the implementation. Knowledge gained through
scientific studies of implementation increases the relevance
of and contributes to the theory building efforts within
policy analysis and political science.

We can also study policy implementation for *professional*
reasons. Scientific and universally generalizable explana-
tions of implementation can lead to far-reaching improvements
in program execution and help bring the effects of a program
closer to its original intent. Here the emphasis is on
application of knowledge gained from scientific studies of

implementation to the solution of practical problems.

Finally, we can study policy implementation for *political* reasons. Improved understanding of what makes the implementation of some policies more successful can lead to improvements in policy formulation and policy operational-ization. Periodic feedback about the status of implementa-tion of a policy can generate information which the policy-makers can use to change the initial course of the policy or to devise new policies. Needless to say, the same informa-tion could also be used by the advocates and proponents of the policy for furthering their own interests.

In addition to its information generation aspects, studies of implementation can lead to a better understand-ing of the institutions and processes of implementation and the "capacity" of the bureaucracy to execute the specific policies formulated by the legislative branch of government. This, in turn, can lead to changes in bureaucratic and orga-nizational design for the betterment of the governmental delivery mechanism.

In this dissertation we come as close to formulation and empirical verification of generalizations about imple-mentation as our data permit. Although we are dealing here with a "marginal" policy within a single governmental pro-gram--the handicapped effort in Head Start--our findings have relevance to the implementation of other policies. Naturally, further verification of our findings by other researchers will increase the credibility of the

ge

tl

tl

se

ic

ge

vi

th

generalizations upon which these findings are based.

We have divided this chapter into two sections. In the first section we outline the purposes of the dissertation. The second section includes a description of the scope and limitations of the study.

Purpose and Objectives

We see as our principal task the formulation and empirical verification of at least one heretofore nonobvious generalization regarding policy implementation. In our view, fulfillment of this purpose entails achievement of the following study objectives:

1. To review the existing literature on policy implementation;

2. To define the central concepts, such as "policy", and "implementation", and illustrate the place of implementation within the overall policy process;

3. To formulate an analytical/conceptual scheme for explaining the variations in the primary dependent variable (level of implementation success);

4. To define and discuss the variables included in the conceptual scheme and their proposed relationships to one another;

5. To examine the theoretical support for the conceptual scheme and the hypothesized relationships;

6. To describe the setting within which the specific policy examined here (the handicapped policy in Project Head Start) was formulated, operationalized, and implemented;

7. To describe and discuss the research design and the methods and procedures used in collecting the data;

8. To describe and discuss measurement of the variables included in the conceptual scheme;

9. To analyze the data, test the alternative hypotheses, and summarize the empirical findings;

10. To discuss the implications of the findings, draw conclusions

regarding policy implementation, and formulate recommenda-
tions concerning future work in this area.

These ten, then, are the specific objectives of the
dissertation. We address objective 1 in Chapter II and
objectives 2 through 5 in Chapter III. Objectives 6 through
8 are accomplished in Chapter IV and objectives 9 and 10 are
covered in Chapters V and VI of the dissertation, respec-
tively.

Scope

The term "scope" generally refers to those character-
istics of a piece of work that define and describe a set of
boundaries. In the case of a doctoral dissertation, these
include the boundaries of the problem being studied, the
limitations of the techniques and procedures employed, and
the generalizability of the empirical findings. We will be
discussing each of these issues in great detail in the com-
ing chapters. Our purpose here is to state our responses
to four major boundary-related questions. These four ques-
tions are:

1. Should the study of policy implementation be regarded as
   a proper branch of politics and political science?

2. What types of policies are the generalizations contained
   here most applicable to?

3. What type of a research design has been used in this study
   and what are the major dependent and independent variables?

4. What are the major limitations of the study?

The first question relates to the broad issue of whether
this dissertation should be accepted as a "proper" political

science dissertation. The other three questions all relate to important specific aspects of the work presented here.

Policy Implementation Versus
Political Science

Our thesis with regard to this issue is that so long as policy analysis, the study of causes, processes and consequences of governmental activity, is regarded as a branch of political science, policy implementation, the study of the execution of public policy, should also be accepted as one of its sub-branches.

We do not believe there is much opposition in the discipline to the view that policy analysis is a proper branch of political science. In fact, some claim that, following the traditionalist and the behavioralist movements, political science today is rapidly shifting towards policy analysis.[1] This is also evidenced by the increasing number of political scientists who are preoccupied with analytical-conceptual as well as empirical studies in policy analysis. Other indicators of the emergence of policy analysis as a proper branch of political science, if not as a discipline in itself, include publication of new journals in policy analysis,[2] the increase in the number of universities with

---

[1]See: Thomas R. Dye, Understanding Public Policy (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1972), Chapter 1.

[2]In particular: Policy Sciences (Amsterdam: Elsevier Scientific Publishing Company) and Policy Studies Journal (Urbana, Illinois: The Policy Studies Organization).

p

g

d

n

i

i

f

l

ca

cr

de

ex

ac

po

he

ce

eva

Eas
___

Pol

(Chi

(Chl

(Engl

Ph.D. programs in public policy, and the number of under-

graduate and graduate courses in public policy offered by

departments of political science.[1]

What about policy implementation? Wouldn't the dra-

matic changes in the status of policy analysis within polit-

ical science imply the occurence of similar changes in one of

its subdisciplines? The answer, as we will demonstrate more

fully in Chapter II, is "no." Most policy scientists have

limited the scope of policy analysis to the study of the

causes and effects of governmental activity. Implementation,

or the "activity" itself, has strangely been left out. Dye

defines the scope of policy analysis as "the description and

explanation of the causes and consequences of governmental

activity."[2] Sharkansky jumps from "determinants of public

policy" to "evaluation of public policy."[3] Ranney, although

he uses the term "implementation" in passing, views the

central concern of public policy analysis as formulation and

evaluation of public policies.[4]

Other examples could be given where, in perhaps a pure

Eastonian tradition,[5] emphasis is placed on the inputs,

---

[1]See: H. Hugo Heclo, "Policy Analysis", British Journal of Political Science, Volume 2, Part 1, January 1972, p. 83.

[2]Dye, Understanding Public Policy, p. 3

[3]Ira Sharkansky, ed., Policy Analysis in Political Science (Chicago: Markham Publishing Company, 1970).

[4]Austin Ranney, ed., Political Science and Public Policy (Chicago: Markham Publishing Company, 1968), pp.3-21.

[5]See: David Easton, A Framework for Political Analysis (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1965).

out

wit

bou

pai

th

fo

si

an

in

an

pha

cal

the

fal

of

as

pro

wit

hav

enc

bee

iza

Sur

Penn

outputs and the environment of a public policymaking system with little or no concern for what happens within the "black box." Needless to say, if one were to use the Eastonian paradigm, the place of policy implementation would be inside the black box.

One of the few policy scientists to recognize a role for policy implementation within the field of policy analysis is Dror. In Public Policymaking Reexamined he offers an "optimal model of public policymaking."[1] The Dror model includes three major stages--metapolicymaking, policymaking, and post-policymaking--and eighteen phases. Two of the phases under the post-policymaking stage are what Dror has called "motivating the executing of policy" and "executing the policy." Although Dror's discussion of these two phases falls far short of systematic analytic treatment, his is one of the few non-empirical studies recognizing implementation as a major branch of policy analysis.

If policy implementation were a straightforward process, i.e., if prior to implementation one could predict with great certainty what the outcomes would be, there might have been no need to study implementation. However, experience has taught us that implementation does not always proceed as specified during policy formulation or operationalization. Policies are sometimes modified to such a degree during implementation that one is not sure if it is the

[1]Yehezkel Dror, Public Policymaking Reexamined (Scranton, Pennsylvania: Chandler Publishing Company, 1968), Chapters 13 and 14.

initially formulated policy that is being implemented.

Knowledge of the consequences of a policy allows one to compare its results with the original intent. However, this comparison does not tell one why and how the eventual consequences were attained. Without studying its implementation all we can say about a policy is whether its intended effects were realized. Studying implementation helps us understand and explain why a particular policy succeeds or fails.

In conclusion, policy implementation should be viewed as an important branch of policy analysis and political science. It is a much talked about but seldom studied field. Analytic studies of implementation are required not only for making improvements in the formulation and execution of policies but also for explaining their observed consequences.

Universality of the Generalizations

Implementation of the handicapped policy in Head Start is the only policy studied in this dissertation. Therefore, from a purely statistical viewpoint, the empirical generalizations of this dissertation can be applied only to the Head Start program and its handicapped policy.

From another standpoint, however, our findings have applicability beyond Head Start. Leaving aside the issues relating to data and empirical verification, the theoretical-conceptual chapter of the dissertation was prepared somewhat independently of the handicapped policy in Head Start. In

pr

ti

fe

pr

ci

kn

pr

cr

re

be

vi

are

por

gre

gen

the

in

tho

exa

der

will

huma

node

Marco

preparing this chapter we were more concerned with formulation of a conceptual scheme that would be applicable to all federally-formulated but locally-implemented human service programs such as Head Start, Neighborhood Youth Corps, Model Cities and the Public Employment Program. Our first hand knowledge of the difficulties faced in implementing these programs at the community level had a lot to do with the crystallization of our thoughts and ideas. It is for this reason that we believe most of the arguments made here could be highly relevant to the implementation of other human service programs. Relevance of our generalizations to a policy area beyond human services, such as foreign policy or transportation policy is an issue we have not investigated.

Universality of a generalization determines, to a great extent, the credibility of the theory from which the generalization has been derived. As argued by Stinchcombe, the credibility of a theory increases as the number of cases in which different empirical consequences of the theory increases.[1] In this instance, Head Start provides only one example of an empirical consequence. If the generalizations derived here hold true in other instances--as we expect they will in the case of federally-formulated, locally-implemented human service programs--the credibility of our conceptual model will increase accordingly.

---

[1]Arthur L. Stinchcombe, Constructing Social Theories (New York: Harcourt, Brace and World, Inc., 1968), Chapter Two.

w

t

t

t

c

si

ab

is

de

le

is

abl

tal

eva

i.e

and

size

pote

(2)

poli

of e

h th

Research Design and Dependent and
Independent Variables

This study is based on a cross-sectional survey design
with a sample of thirty-five Head Start programs visited in
the spring of 1974 to collect data on the implementation of
the handicapped policy. The design, therefore, suffers from
the usual threats to internal and external validity that are
common to most non-experimental designs that are based on
small-sample survey research.

The analytic/conceptual scheme used includes six vari-
ables. Of these, one is the primary dependent variable, one
is an intervening variable, and the other four are indepen-
dent variables. The principal dependent variable is the
level of implementation success and the intervening variable
is the level of implementation effort. The independent vari-
ables of the model are implementation potential, environmen-
tal forces for implementation, policy support, and self-
evaluation of capability.

Four of the six variables are measured through indices,
i.e., through a set of "secondary" variables or indicators,
and the remaining two are measured directly. The hypothe-
sized relationships concern (1) the effects of implementation
potential and level of effort on implementation success and
(2) the effects of environmental forces for implementation,
policy support, and self-evaluation of capability on level
of effort. In the case of the first relationship three and
in the case of the second relationship two alternative

models are proposed. The explanatory power of each alternative model is assessed using the Head Start data.

Limitations of the Study

Perhaps the most important limitation of the study is the weakness of the theoretical support for the conceptual model proposed. Expectancy theory in organizational psychology provides the only major theoretical base for the hypothesized relationships. Other than this, the model stands mainly on intuitive and empirical grounds.

The second limitation of the study is the size of the sample upon which the empirical generalizations are based. Implementation of the handicapped policy in only thirty-five local Head Start programs is investigated. This alone limits the number of independent variables that can be considered in explaining the dependent variable. In addition, the size of the sample affects the magnitude of errors due to sampling.

The third limitation of the study has to do with the internal validity of the cross-sectional design. Although attempts were made to estimate some longitudinal data, the design lacks the desirable features of experimental and quasi-experimental research designs. These features include pre- and post-measurement, availability of a control group, and random assignment of programs to the two groups. These deficiencies of the design have been partially removed by adopting appropriate fieldwork procedures and by controling

for some variables through statistical means.

Finally, the fourth limitation of the study relates to the validity of the indices used for measuring four of the variables. Without the aid of reference measures or indicators the process of index construction becomes highly judgmental.

The four limitations discussed above, along with our earlier comments on generalizability, constitute the major limitations of the dissertation as we see them. Undoubtedly there are other aspects of the study which could be included here as limitations. However, we believe that they are less major than the ones discussed. In any event they will be reviewed in greater depth in the remainder of the dissertation.

This concludes our discussion of the purpose and the scope of the dissertation. In the next chapter we will review the literature pertinent to the study of policy implementation.

CHAPTER II

LITERATURE ON POLICY IMPLEMENTATION

There is not much literature on policy implementation. This was reported previously by Pressman and Wildavsky[1] and the supplemental literature search effort we undertook did not lead to conclusions different from theirs. In light of this negative finding, our purposes in this chapter are to describe this literature search effort, to review the most important of the analytic/conceptual and empirical works we examined, and to identify the common variables used or proposed by others for explaining implementation success.

The organization of the chapter is as follows. In the first section we summarize Pressman and Wildavsky's findings from a major bibliographic work on implementation and we follow this with a description of our additional search for literature relevant to the present work. In the second section we review what we consider to be the four most important empirical studies on implementation. The third and final section of the chapter is devoted to recapitulation of and conclusions from our literature review effort.

---

[1]Jeffrey L. Pressman and Aaron Wildavsky, Implementation (Berkeley, California: University of California Press, 1973).

Literature Search

The only major bibliographic work on implementation is
reported in Appendix 2 of Pressman and Wildavsky's
Implementation. Expecting a large body of literature, Press-
man and Wildavsky have conducted an extensive review of major
works in disciplines closely related to implementation. They
report the results of their investigation as follows:

> "Analytical study (as opposed to mere mention) of implementation
> seems so eminently reasonable that few can imagine it does not
> exist. . . . (E)xcept for the few pieces mentioned in the body
> of this book, we have been unable to find any significant
> analytic work dealing with implementation.
>
> "How shall we persuade others that it is fruitless to look for
> a literature that does not exist, or that we cannot connect
> our work with nonexistent analyses by others? It is a little
> like searching for the little man who isn't there. It is
> always more difficult to establish a negative rather than posi-
> tive relationship. While we cannot hope to provide conclusive
> evidence (after all, there can always be something we have over-
> looked) we do hope to be persuasive beyond the mere assertion
> that there is no analytic literature on implementation."[1]

Pressman and Wildavsky present their findings from the
extensive literature search under five headings: periodical
indexes, poverty, civil rights, operations research, and
public administration. In all instances they conclude that
nothing of importance can be found in the works examined.
In the case of public administration, the texts reviewed by
Pressman and Wildavsky generally use the term "implementa-
tion" in the introductory sections of the work for defining
other terms. But implementation itself is never defined or
explained. It is mostly used in the sense of "the end

---

[1]Ibid., pp. 166-167.

product of administrative efforts." In conclusion, the search conducted by Pressman and Wildavsky reveals that implementation is a much-talked-about but rarely-studied topic. Except for a few empirical works, which we review in the second section of this chapter, they have found no analytic studies of importance on policy implementation.

We supplemented the Pressman-Wildavsky search reported above with a search of our own. Our purpose in conducting this investigation was to confirm Pressman and Wildavsky's findings and to carefully scan literature not covered by Pressman and Wildavsky.

In the first instance we undertook a library search using the subject catalogues of the Main Library at Michigan State University. We searched for titles under topics such as policy implementation, implementation, and execution. This search did not reveal anything different from the findings of Pressman and Wildavsky.

Second, we conducted an in-depth review of some of the leading texts in public policy analysis. This review included, principally, the works by Dye, Dror, Rourke, Ranney, Sharkansky, Havemen and Margolis, Jones, Lindblom, Braybrooke and Lindblom, Laswell and Kaplan, and Easton.[1] Although none

---

[1]Thomas R. Dye, _Understanding Public Policy_ (Englewood Cliffs, New Jersey; Prentice-Hall, Inc., 1972); Yehezkel Dror, _Public Policymaking Reexamined_ (Scranton, Pennsylvania: Chandler Publishing Company, 1968); Francis E. Rourke, _Bureaucracy, Politics and Public Policy_ (Boston, Massachusetts: Little, Brown and Company, 1969); Austin Ranney, ed., _Political Science and Public Policy_ (Chicago, Illinois; Markham Publishing Company, 1970); Robert H. Haveman and Julius Margolis, ed., _Public Expenditures and Policy Analysis_ (Chicago, Illinois: Markham Publishing

of these major works used the term "implementation", some
referred to it as the "application" and others as the "exe-
cution" of a policy. Except for Dror's inclusion of imple-
lementation as a major stage in his "optimal model of policy-
making" we could not find an analytic-conceptual discussion
of policy implementation in these works.

Third, we turned our attention to policy evaluation,
which, in the recent years has been emerging as an important
discipline within policy analysis. The literature we
reviewed in this area, such as works by Suchman, Caro, Weiss,
Wholey, et. al., and Rossi and Williams,[1] seemed to be more
concerned with questions of definition, measurement and
design than the relationship of implementation to program
success.

Fourth, and finally, we conducted an extensive,
article-by-article review of six major periodicals. This
included all issues of four journals: Policy Sciences,

Company, 1970); Charles O. Jones, An Introduction to the Study of Public
Policy (Belmont, California: Wadsworth, 1970); Charles E. Lindblom,
The Policy-Making Process (Englewood Cliffs, New Jersey: Prentice-Hall,
Inc., 1968); David Braybrooke and Charles E. Lindblom, A Strategy of Deci-
sion (New York, The Free Press, 1963); Harold Laswell and Abraham Kaplan,
Power and Society (New Haven, Connecticut; Yale University Press, 1950);
David Easton, A Framework for Political Analysis (Englewood Cliffs, New
Jersey: Prentice-Hall, Inc., 1965).

[1]Edward A. Suchman, Evaluative Research (New York, Russell Sage
Foundation, 1967); Francis G. Caro, ed., Readings in Evaluation Research
(New York, Russell Sage Foundation, 1971); Carol H. Weiss, Evaluation
Research (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1972);
Joseph S. Wholey, et al., Federal Evaluation Policy (Washington, D.C.:
The Urban Institute, 1971); Peter H. Rossi and Walter Williams, eds.,
Evaluating Social Programs (New York: Seminar Press, 1972).

Policy Studies Journal, British Journal of Political Science, and Public Choice.[1] In addition, we reviewed the issues of The American Political Science Review published after 1968 and Public Administration Review published after 1969. This investigation revealed a few articles of relevance. We devote the remainder of this section to a review of four of these articles.

First, in a recent article in Policy Sciences Thomas B. Smith presents a discussion and a conceptual model of the policy implementation process.[2] Smith argues that the assumptions about implementation made in most policy studies (e.g., "once a policy is formulated, its implementation is automatic") is a major reason for the neglect of policy imple- mentation in the models constructed for studying policy- making. In Smith's model policies are viewed as serving a tension generating force in society. Accordingly, the four components of the policy implementation process--the ideal- ized policy, the implementing organization, the target group, and environmental factors--interact to produce "discrepancies and tensions." The tensions result in transactions, trans- actions sometimes generate institutionalization; and, "feed- back in the form of relieved tensions or increased tension is introduced back into the tension generating matrix from

---

[1]The first few issues of Public Choice were published under the title : Papers on Non-Market Decision Making.

[2]Thomas B. Smith, "The Policy Implementation Process," Policy Sciences, Volume 4, 1973, pp. 197-209.

transactions and institutions."[1]

The variables proposed by Smith for studying the four components of the policy implementation process include, among others, type of the policy, level of commitment of the government to the implementation of the policy, level of demand or support for the policy, characteristics of the target group, characteristics of the implementing organization, and intensity and care taken by the implementing agency to organize for the implementation. These and other variables, according to Smith, can be used to examine the within- and inter-component tensions and their consequences. Smith's argument falls far short of systematic formulation and none of his implied hypotheses have been verified empirically.

Second, in another study Thomas Smith reports his findings from an empirical analysis of individuals in policy formulation and policy implementation roles in New Zealand.[2] Based on a sample of 119 individuals in 27 government departments, the study compares policy formulators and implementors in terms of their social background, career patterns, education, job satisfaction, career ambition, work load, hierarchial relations, and political saliency. Smith's principal findings are stated as follows:

---

[1]Ibid., p. 203.

[2]Thomas B. Smith, "Policy Roles: An Analysis of Policy Formulators and Policy Implementors," Policy Sciences, Volume 4, 1973, pp. 297-307.

"From the evidence of this study, the New Zealand administrative system does not allocate policy formulating roles to individuals different from those who implement policies. Differences between formulators and implementors in the New Zealand system appear to stem from the nature of the work of the two policy roles."[1]

The theoretical significance of these findings are questionable. The specific finding for New Zealand is perhaps more significant for elite theorists than for policy scientists.

Third, in a recent article in Policy Sciences, Douglas Bunker, arguing for the view that implementation is an integral part of the policymaking process proposes a heuristic model for studying different implementation strategies.[2] In Bunker's model, which is based on Gergen's leverage-point concept,[3] implementation is viewed as a problem of social persuasion. A three dimensional scheme is presented for "locating actors and selecting directions of preferred change to optimize probabilities of implementation."[4] The three dimensions of the scheme refer to variables representing

(1) the centrality of the issue for a particular actor,

(2) the number and potency of political resources he has

---

[1]Ibid., p. 297.

[2]Douglas R. Bunker, "Policy Sciences Perspectives on Implementation Processes," Policy Sciences, Volume 3, 1972.

[3]This concept is based on the argument that the implementation process can be influenced at a finite number of leverage points. Groups and individuals within an organization may also be conceived of as leverage points. See: Ibid., p. 75. Also see: K. Gergen, "Assessing the Leverage Points in the Process of Policy Formulation" in R. A. Bauer and K. H. Gergen, eds., The Study of Policy Formation (New York: The Free Press, 1968).

[4]Bunker, "Perspectives on Implementation Processes", p. 76.

available, and (3) the degree to which he is in agreement with the policy. Bunker also argues that policy implementation in a federal framework requires multi-level participation of four groups of participants: administrative-bureaucratic, political leadership, rational-analytic, and constituent elites. Successful implementation of a federal policy, according to Bunker, demands "both vertical (central office-field) coordination and the integration of multiple functional inputs."[1] Bunker's analysis, as he himself recognizes, has more heuristic than practical value.

The *fourth* article we will cover here is a brief essay by Marshall pointing the need for research on implementation of federal poverty and welfare programs.[2] "Political scientists have given more attention to the formulation of federal policy than to its implementation. Thus we know much less about what happens after the legislative drama fades and the bureaucrats go to work to implement the legislation."[3]

According to Marshall there are three major issues in the study of implementation. The first issue concerns the results of implementing social policies; the second issue is about the characteristics of the implementation process itself; and the third issue relates to the factors

---

[1]*Ibid.*, p. 78.

[2]Dale Rogers Marshall, "Implementation of Federal Poverty and Welfare Policy: A Review Essay," *Policy Studies Journal*, Volume 2, No. 3, Spring 1974.

[3]*Ibid.*, p. 152.

explaining variations in the way different cities respond
to federal policy. After reviewing the extent of research
conducted in these three areas, Marshall concludes his essay
as follows:

> "Just as the pattern of policy implementation becomes familiar,
> so does the pattern of review essays with their concluding
> (self-serving?) pleas for more research. . . . But, alas, as
> is blatantly obvious from the amount of verbiage expended, it
> is much easier to call for studies than to do studies and stand
> the gaff of the methodologists attacking lack of rigor and the
> practitioners bemoaning documentation of the insignificant."[1]

Perhaps Marshall's remarks are as good a point as any for
concluding this section of the chapter. In the next section
we will review some of the major empirical studies of imple-
mentation.

Major Empirical Studies of Implementation

During the literature search stage of the preparation
of this dissertation we were able to identify only four
major empirical studies on policy implementation. These
four studies, in the order they will be reviewed in this
section, are: the Pressman-Wildavsky study of the implemen-
tation of an experimental Economic Development Administra-
tion project in Oakland, California; the Bailey and Mosher
study of the implementation of The Elementary and Secondary
Education Act of 1965; the Sudquist and Davis study of pro-
gram coordination at the community level; and Martha
Derthick's investigation of the implementation of a housing
program for the poor.

---

[1] Ibid., p. 156.

There is a fifth study which we could have included in this section along with the above four, one which could perhaps be regarded as the "first" empirical study on policy implementation. This is Selznick's renowned work on the Tennessee Valley Authority (TVA).[1] Although the Selznick study and the concepts it has introduced, such as cooptation, grass-roots administration, voluntary participation, and commitment, have not lost their relevance, they have been restated by others in forms more appropriate to today's problems. Indeed, some of the studies we review below use a line of argument similar to Selznick's in explaining the variations in successful policy implementation.

A Major Study of Implementation

The only major work we have come across that carries the title "implementation" is the Pressman and Wildavsky study of the Economic Development Administration's (EDA) experimental project in Oakland, California.[2] The Oakland Project, started in 1966, was the first project funded within EDA's major rebuilding program. "To help solve the problems of unemployment and racial unrest, the new program (was designed to) finance public works and business loans that would lead to the creation of jobs for the unemployed, primarily black."[3] Prior to implementation of the project

_____

[1]Philip Selznick, TVA and the Grass Roots (Berkeley, California: University of California Press, 1949).

[2]Pressman and Wildavsky, Implementation.

[3]Ibid., p. 1.

it was announced by EDA that more than twenty-three million dollars in federal funds would be offered for public works grants and business loans in Oakland. EDA estimated at that time that the direct employment effect (without spinoffs) of these grants and loans would be about 3,000 additional jobs. Yet, in 1969, three years after the project began, only three million dollars had been spent and the total employment effect of the project was negligible.

Pressman and Wildavsky present an insightful analysis of the apparent failure of the Oakland Project. Here we shall summarize the four reasons given by Pressman and Wildavsky.

First, implementation of the Oakland Project involved participation of a large number of agencies and individuals and each of the participants had a "distinctive perspective from which it viewed the EDA operation."[1] Although none of the participants disagreed strongly with the major goals of the project, they differed in terms of the means of attaining them. They each had their own interests and preferences about the sequence of events to be followed. This introduced lengthy delays in clearances and agreements necessary to implement the program.

Second, the number of "decision points" that had to be passed for the program to continue, the number of participants involved at each major decision point, and the total

[1] Ibid., p. 98.

number of agreements required presented major roadblocks to successful implementation.

Third, the two major goals of the program, construction of public works and creation of jobs for the hard-to-employ, required two separate decision paths and series of action. Delays in one decision path stalled the other.

Fourth, the project experienced, what Pressman and Wildavsky call, "the emergence of unexpected decisions." As they state: "Not only was it necessary to complete two decision paths in order to achieve the EDA program goals; but within each path, the agency found that the number of decisions and clearances required was constantly growing. . . . The paths of required decisions . . . were soon characterized by more unexpected elements than expected ones: they were anything but straight lines leading directly to goals."[1]

In concluding the study Pressman and Wildavsky summarize their major finding as follows:

> "What had looked like a relatively simple, urgent, and direct program - involving one federal agency, one city, and a substantial and immediate funding commitment - eventually involved numerous diverse participants and a much longer series of decisions than was planned. None of the participants actually disagreed with the goal of providing jobs for the minority unemployed, but their differing perspectives and senses of urgency made it difficult to translate broad substantive agreement into effective policy implementation."[2]

The significance of Pressman and Wildavsky's work does

---

[1] Ibid., p. 112.

[2] Ibid., p. 113.

not lie in the methodological soundness of the explanation
offered in the study. Instead, it lies in the fact that it
is the first major empirical study carrying the title "imple-
mentation." The study demonstrates the usefulness of study-
ing policy implementation. From a theoretical standpoint,
it introduces a few variables which could explain variations
in implementation success in other instances. The study
also illustrates the potential power of the case-study
method in the formulation of hypotheses which could be fur-
ther tested in other circumstances.

Implementation of a Major Act: The Case of ESEA

Another major empirical work on implementation is found
in Bailey and Mosher's account of the U.S. Office of Educa-
tion's administration of The Elementary and Secondary Educa-
tion Act of 1965 (ESEA).[1] In contrast to the conceptual-
analytic nature of the Pressman-Wildavsky study the Bailey
and Mosher piece is more of a descriptive-historical under-
taking. Although a lengthy chapter is devoted to the pro-
cess of implementation,[2] this section goes little beyond a
description of the tasks and activities carried out during
the implementation of the six titles of the act. Lack of
data on successfulness of implementation seems to be the
apparent reason for this emphasis on description of tasks

---

[1]Stephen K. Bailey and Edith K. Mosher, ESEA: The Office of
Education Administers a Law (Syracuse, N.Y.: Syracuse University
Press, 1968).

[2]Ibid., pp. 98-159.

and activities rather than explanation of variations in success.

The conclusions reached by Bailey and Mosher about implementation are based on their assessment of the barriers to successful implementation. "Although each one of the Titles of ESEA produced its own particular difficulties and problems, the major issues that developed in the partnership of implementation cut across most, if not all of the titles."[1] These major issues were:

- difficulties in interpreting the legislative intent of the mandate;

- lateness of funding;

- difficulty of finding specialized staff;

- paper work, red tape, and "federal control";

- difficulties created by interagency clearance requirements; and,

- administrative and constitutional issues related to application of ESEA provisions to non-public schools.

Bailey and Mosher present considerable qualitative data in describing the problems faced during implementation. Because of lack of data on successfulness, however, Bailey and Mosher were not able to test their hypotheses empirically.

A Study on Coordination

A recent study undertaken by the Brookings Institution concentrates on an important aspect of implementation:

---

[1]Ibid., p. 188.

program coordination at the community level.[1] The Brookings study does not explicitly deal with the problems of implementing a specific program. Instead, it addresses the question of coordinating the simultaneous implementation of several federal programs at the local level. The study is based on about 700 interviews conducted during 1967-69 with federal, state, and local officials administering federally sponsored programs, citizen leaders and informed observers in eight states.

Sundquist and Davis zero in on problems of coordination in "urban America" and compare them with coordination difficulties faced in implementing programs in "rural America." In the case of urban America, they study the failure of the Community Action Agencies as local coordinating structures and contrast it with the relatively greater success level achieved by Model Cities. For resolving the coordination problems in rural America Sundquist and Davis suggest that the model to be adopted by federal, state, and local governments should be a "universal system of multi-county agencies with responsibility both for planning and for facilitating action programs covering the entire range of community activities relating to economic and community development."[2]

In the Sundquist-Davis study coordination is viewed

---

[1]James L. Sundquist and David W. Davis, Making Federalism Work (Washington, D.C.: The Brookings Institution, 1969).

[2]Ibid., p. 231.

as

ce

fe

lo

nat

whe

imp

arg

dan

whi

of

are

comp

Dert

dept

John

of b

owne

hous

Altho

start

of th

Instit

as a prerequisite to successful implementation. Indeed, the central premise of the study is that effective execution of federal programs depends heavily upon the competence of local institutions to plan, initiate and coordinate. Coordination, according to Sundquist and Davis, is best achieved when it is assigned to a "neutral" agency instead of one with implementation responsibilities. The implication of this argument for policy implementation is that, given the abundance of federally-sponsored, locally-implemented programs which are dependent on one another, successful implementation of these programs will require each urban community and rural area to establish a series of "neutral" coordination centers.

A Comparative Study of Policy Implementation

The final work we will cover in this section is the comparative study on implementation reported in Martha Derthick's New Towns In-Town.[1] Derthick conducted an in-depth study of a small federal program initiated during the Johnson Administration. The program was based on the idea of building model new communities on surplus, federally-owned land in metropolitan areas. Its aims were to provide housing for the poor and revitalize the nation's cities. Although the program seemed to have a quick and ambitious start, in three years it faced complete failure. In three of the seven local areas selected for program implementation

---

[1]Martha Derthick, New Towns In-Town (Washington, D.C.: The Urban Institute, 1972).

th

pr

de

dr

pr

at

je

gr

ex

th

on

su

of

the projects died outright. In the other four areas the projects faced eventual failure. Three years after President Johnson's announcement of the program only three hundred housing units were under construction.[1]

Derthick's study mainly concentrates on the individual projects. The historical-analytic account of each project attempts to explain the reasons for the failure of the projects and the program. Since the projects, taken as a group, do not show much variation in terms of successfulness, explanation of success in implementation is not possible through this study.[2] By studying cases of failure, however, one gains insights into what might have made the program more successful.

Derthick's analysis yields four reasons for the failure of the New Towns In-Town program:[3]

- Failure, on the part of federal officials, to anticipate and counteract resistance to the projects from some local groups;

- Lack of federal incentives to induce local governments to implement the program;

- Inability of federal and local officials to reinforce one another and act effectively for generating public support,

---

[1]President Johnson announced the program on August 30, 1967. In his announcement he declared that the program would generate modern housing and related services for about 25,000 persons, in Washington, D.C., alone: See: Derthick, p. 7.

[2]This should not be viewed as a criticism of the Derthick study. Implementation of the program involved only seven projects; and Derthick studied all of them, i.e., the "total population" of projects. Needless to say, the researcher cannot be blamed for lack of variation in the dependent variable.

[3]Derthick, pp. 83-102.

In

st

wi

th

p

de

be

s

so

th

i

/

overcoming opposition and building an administrative frame-
work;

● Statement of objectives that were too ambitious for attain-
ment at the local level.

In the author's own words:

"In summary, the surplus lands program failed both because the
federal government had limited influence at the local level and
because it set impossibly high objectives. Its goals exceeded
by far its capacity to achieve them. . . . Fundamentally, the
program failed because of characteristics of the federal govern-
ment that are associated with . . . its central position in the
governmental system. These characteristics are the scale of its
jurisdiction and its separation from the actual execution of
domestic programs. Separation, in turn, results from the divi-
sion of authority among governments in a federal system and the
distance from the 'top' and the 'bottom' level of government
hierarchy in a large, complex society."[1]

This concludes our review of the four major empirical

studies on policy implementation. In the next section we

will recapitulate what was presented earlier and analyze

their implications for the present study.

Conclusions from Past Work on Implementation

Let us now summarize what we have covered in the two

preceding sections and assess where we stand.

First of all, we documented, rather laboriously, and

demonstrated the fact that policy implementation has not

been studied extensively. There are less than a dozen

studies that can be claimed as policy implementation studies,

some of those by stretching the definition and boundaries of

the field and others by lowering the standards for scientific

inquiry.

---

[1]Ibid., p. 93.

w

s

t

h

o

a

e

"

r

ɔ

l

t

t

r

ɔ

ɔ

n

ɬ

c

d

ɛ

ɛ

ɛ

Secondly, we presented summaries/reviews of a few key works on implementation. In terms of their theoretical soundness perhaps none of the studies reviewed can pass a tough methodological test. A few of these were purely heuristic attempts at conceptualization. The others, the ones we have called "empirical," were historical/analytical accounts of the reasons for the failure of one or more federal programs. None of the studies examined included a "grand theory" of implementation; each followed a line of reasoning that seemed to be most suitable to the specifics of the case under investigation.

Finally, in our review of most of the studies we listed the key explanatory variables proposed by the investigators. We did not use any rigid criteria in choosing the variables to be included in these lists. Instead, we relied heavily on the authors' assessment of the importance of a potentially explanatory factor.

In the remainder of this chapter we will attempt to integrate the principal explanatory variables of the eight major works reviewed in the first two sections of this chapter. Our purpose in carrying out this exercise is to find out what commonalities exist between these eight, rather diverse studies.

The procedure we followed is quite straightforward. First, we listed all the explanatory variables proposed in each study. This yielded a list of about twenty-five variables. Some of these were "composite variables", i.e.,

they consisted of a combination of more than one "simple variable."[1] In such instances we divided each composite variable into its simple variables. This increased the total number of variables to about thirty. We then examined this last set of variables in order to group them into broader categories. After a few iterations of this, what might be called "conceptual factor analysis," we ended up with five new variables. Despite the trial-and-error nature of the procedure used, we believe that these five categories adequately represent the host of explanatory variables pro-posed by the eight studies. The following is a list of the five new variables.

- Level of governmental commitment

- Straightforwardness of the implementation

- Level of inter- and intra-agency agreement

- Capability of the implementing agency

- Local support for the policy

The first variable, *level of governmental commitment*, refers to commitment to the implementation of the policy by govern-mental agencies, including the legislature, at levels higher than the implementing agency. In most instances this vari-able reflects the commitment of the federal government. It encompasses Derthick's "federal incentives," Smith's "fed-eral commitment," and Bailey and Mosher's "funding level."

---

[1]We are using the term "simple variable" to refer to a variable with which a simple hypothesis can be formulated. A simple hypothesis is of the form "if A then B".

In addition, the commitment concept covers variables such as "establishment of realistic (attainable) objectives," "clarification of policy intent," and "design and distribution of clear guidelines."

Straightforwardness of the implementation covers most of the variables proposed by Pressman and Wildavsky. These include "number of decision points to be passed," "number of participants involved at each decision point," and "total number of agreements required." This variable also covers Bailey and Mosher's "number of interagency clearance requirements."

Level of inter- and intra-agency agreement is a variable encompassing a host of factors. In the first instance it refers to the stand taken by the leadership of the implementing organization vis-a-vis a positive advocate of the policy (Bunker). In the second instance, it refers to consensus within the implementing organization, i.e., agreement of the subordinates with the leadership. Finally, it refers to "differences in perspectives of the agencies involved in implementation" (Pressman and Wildavsky), "communication between and working together of federal and local agencies" (Derthick), and "existence of 'neutral' coordinating structures" (Sundquist and Davis).

The fourth variable, *capability of the implementing agency*, covers three of Smith's factors: "stability and structure of the implementing organization," "qualifications of the implementing organization's staff," and "intensity and care taken by the implementing agency to organize for the

implementation of the policy." It also encompasses Bunker's notion of "number and potency of the resources available to the implementing agency" and Bailey and Mosher's "staff capability."

The final variable, *local support for the policy*, refers to the level of support the policy receives from the members of the target group and related local agencies. Most of the studies reviewed proposed at least one support-related variable for explaining implementation success.

Are these five variables independent of one another? We do not think so. Even though we attempted to make them as independent as possible, it is not difficult to construct conceptual schemes and a rationale to illustrate the partial dependence of the five variables and the concepts they represent. Nevertheless, the important point is that during the reconstruction exercise we consciously tried to generate variable categories that are at least as independent of one another as in any other categorization.

The significance of this exercise, in our opinion, lies in its heuristic value. By generating a set of variables that covers the full range of concepts used by others to explain policy success we have a clearer idea of what has worked in other instances and what might work in the present undertaking.

This concludes our review of the literature related to policy implementation. In the next chapter of the dissertation we present a conceptual/theoretical framework which is

partly based on the variables summarized above. In addition,
we outline the theoretical support for the model and list
the specific hypotheses which are later tested in Chapter V.

CHAPTER III

CONCEPTS AND THEORY

In this chapter we introduce, describe, and discuss
the major conceptual and theoretical issues which underlie
our major purpose: explanation of variations in the level
of success achieved in implementing a new public policy.

We have divided the chapter into two broad sections.
In the first section, entitled "Concepts," we discuss the
definitions offered previously for the term "policy" and
introduce a definition of our own. We follow this by a
discussion of the role of policy implementation within the
overall policy process.

In the second section of the chapter, entitled "Theory,"
we concentrate on methodological and theoretical issues.
First, we introduce a conceptual model which summarizes our
hypotheses regarding explanation of implementation success.
Second, we review the theoretical support for the conceptual
model. In the third part of the section we define, describe
and discuss the six variables of the model. In the fourth
part we examine the postulated relationships between these
six variables. In the final part of the section we review
a set of methodological issues which concern the conceptual
model and its theoretical bases.

37

Concepts

The Term "Policy"

The central concept of this dissertation is implementation. However, implementation, in and of itself, has very little meaning. To assign it a meaning, "a verb like 'implement' must be associated with an object like 'policy'."[1] Therefore, it is essential that we arrive at a working definition of "policy" before we can talk about implementation.

Policy, as it is commonly used, refers to a strategy or a means adopted for achieving some prespecified goals. This coincides with the definition of policy given by Webster--"any governing principle, plan, or course of action"[2]--and the synonyms cited by Roget: "polity, principles; program, procedure, course, line, plan of action; platform."[3] In everyday language, however, the term policy is used quite inconsistently. Sometimes we talk of a policy in the sense of a goal or a desirable outcome ("our policy is to minimize the tension in the Middle East"). In other instances policies refer to generalizations from our past behavior ("our policy is to train all caseworkers," meaning that we already did train them).

---

[1]Pressman and Wildavsky, Implementation, p. xv.

[2]See: Webster's New World Dictionary, College Edition (1966). Webster includes two other definitions under "policy": "1. political wisdom or cunning; diplomacy; procedure; artfulness. 2. wise, expedient, or crafty conduct of management." The one we have quoted in the text, and the one we are interested in, is the third definition given by Webster.

[3]See: Roget's International Thesaurus, Third Edition (1962).

There are also a wide variety of definitions of policy offered by political and policy scientists. These range from simple definitions such as Sharkansky's ("policies are actions taken by governments")[1] to complex and sometimes ambiguous formulations such as the one by Reynolds: "A policy is a move or a series of moves (acts, speeches, etc.) made from a position in a practice which, other things being equal, could be reasonably expected to have at least one of the following results: (1) a modification of that practice or some other practice or (2) a lack of modification of that practice or some other practice when a modification could be reasonably expected to occur were the move not made or (3) the creation of another practice."[2] In between these two extremes lie a variety of definitions. Laswell and Kaplan refer to it as "a projected program of goal values and practices."[3] Braybrooke and Lindblom view it as a term encompassing "both decisionmaking and the course that policies take as a result of intercorrelations among decisions and/or in which (policy-making) incorporates certain political processes, in addition to analytic processes, into the determination of courses of action."[4] According to James Robinson a policy "refers to

---

[1] Ira Sharkansky, "The Political Scientist and Policy Analysis," in Ira Sharkansky, ed., Policy Analysis in Political Science (Chicago: Markham Publishing Company, 1970), p. 1.

[2] James F. Reynolds, "Policy Science: A Conceptual and Methodological Analysis," Policy Sciences, Volume 6, Number 1, March 1975, p.7.

[3] Harold Laswell and Abraham Kaplan, Power and Society (New Haven, Connecticut: Yale University Press, 1950), p. 71.

[4] David Braybrooke and Charles E. Lindblom, A Strategy of Decision (New York: The Free Press, 1963), footnote 1, p. 249.

goals (objectives, ends) of any social system, the means

chosen to effectuate these goals, and the consequences of

the means, i.e., the actual distribution of values."[1] Dror

views policies as guidelines: "Public policymaking is a

very complex, dynamic process whose various components make

different contributions to it. It decides major guidelines

for action directed at the future, mainly by governmental

organs. These guidelines (policies) formally aim at achiev-

ing what is in the public interest by the best possible means."[2]

To Easton policies are "decision rules adopted by authorities

as a guide to behavior. . . . In this sense, policies would

be just a term for a kind of authoritative verbal output."[3]

Finally, Austin Ranney defines a policy in terms of its five

principal components (elements): (1) a particular object or

set of objects, (2) a desired course of events, (3) a selected

line of action, (4) a declaration of intent, and (5) an imple-

mentation of intent.[4]

The eight definitions of policy quoted above are not all

---

[1]James Robinson, Congress and Foreign Policy-Making (Homewood , Illinois: Dorsey Press, 1962), p. 3.

[2]Yehezkel Dror, Public Policymaking Reexamined (Scranton, Pennsylvania: Chandler Publishing Company, 1968), p. 12.

[3]David Easton, A Systems Analysis of Political Life (New York: John Wiley and Sons, 1965), p. 358, quoted in Robert H. Salisbury, "The Analysis of Public Policy: A Search for Theories and Roles," in Austin Ranney, ed., Political Science and Public Policy (Chicago: Markham Publishing Company, 1968), p. 152.

[4]Austin Ranney, "The Study of Policy Content: A Framework for Choice," in Austin Ranney, ed., Political Science and Public Policy (Chicago: Markham Publishing Company, 1968), pp. 6-7.

contradictory.   Perhaps the most common element among these
definitions is that policies refer to courses of action,
which is precisely how Webster and Roget define the term.
The major drawback of the set of definitions considered here
is that they are not consistent with each other, i.e., a
course of action which is a policy according to one defini-
tion may fail to meet the specifications of another.   In
view of this inconsistency--and in light of the centrality
of the concept in our work--we need to formulate a single
definition for use throughout the dissertation.

The definition we chose to use here reflects our notion
of policy.   Of the eight definitions quoted earlier it is
closest to Ranney's view of policy.   Yet there are some
major differences between his definition and ours.   These
differences will become clear in the course of our discus-
sion of the definition which follows its formal statement.

> *A policy is a statement, usually written, which has been (or*
> *could potentially be) acted upon by an authoritative decision-*
> *making body and which at least includes (1) a target object*
> *or group of objects, (2) a set of goals or objectives regard-*
> *ing one or more aspects or characteristics of these objects,*
> *and (3) a well-defined strategy or course of action.*

We will now elaborate on this definition by parti-
tioning it into its components (concepts and phrases) and
discussing their significance and implications.

One, policies are statements and a great majority of
them are in the form of written statements, such as those
included in legislative acts or resolutions of decision-
making bodies (a city council, board of commissioners of a

county, board of trustees of a university, executive board
of a corporation, etc.). Unwritten policies are not uncom-
mon and they may relate to significant as well as insig-
nificant matters. Thus, a small corporation may have an
unwritten "marketing policy" or a social welfare department
may have an unwritten "outreach and recruitment policy."
However, most of the policies studied by political and
policy scientists are in the form of formal, written state-
ments.

Two, policies are not always the final outcomes of
decision-making processes. Some policies are "considered"
during the decision-making process but not adopted at the
end. Only one of these alternative policies is eventually
implemented. The alternatives considered, if they meet the
other specifications of the definition, are also regarded
as "policies." The policy selected during the decision-
making process is what we call "the adopted policy."

Three, the decision-making body acting on the policy
question has to have an authority relationship to the
target group of objects defined by the policy. If there is
no authority relationship, i.e., if the group of objects do
not fall within the legitimate span of control of the decision-
making body, the statement adopted by the group is not consid-
ered a policy, it only reflects the opinions and desires of
the decision-making body in question.

Four, if the decision-making body is one generally
accepted as a "public decision-making body," the adopted

policy--and the alternative policies considered--are referred
to as "public policies."

Five, policies are designed to affect, i.e., induce
changes in one or more characteristics of, a well-defined
target object or group of objects. "Well-defined" here
means that there are clear rules with which one can deter-
mine if a particular object falls within or outside the
target group. The target group may consist of the whole
nation or any of its subgroups (elderly, youth, unemployed,
handicapped, etc.), institutions or organizations (small
business organizations, universities, non-profit hospitals,
etc.), physical or geographic entities (regions, river
basins, highways, urban areas, etc.), or, decision-making
rules (the filibuster rule in the Senate, equal rights for
women, the age of majority, etc.).

Six, policies are purposive, or goal-directed state-
ments. The goals or objectives relate to aspects (communi-
cation structure, coordination, resource allocation, etc.)
or characteristics (employability, health, education, etc.)
of the target group of objects. The goals and objectives
define a desired state of affairs which is to be attained
through implementation of the policy. Without goals it
becomes quite impossible to judge policy success.

Seven, a policy prescribes a particular strategy or
course of action which is to be followed during its implemen-
tation. "Course of action" refers to a sequence of acts to
be performed by a prespecified group of actors in a

particular manner and within a given time frame. Well-defined policies indicate <u>what</u> (course of action), <u>who</u> (<u>what</u> actors), <u>when</u> (time-frame), and <u>how</u> (sequence of acts). Most policies also specificy (directly or indirectly) <u>why</u> the particular course of action was chosen from among the alternatives considered.

Strategies are the "backbones" of policies. They contain the principal "if X, then Y"-type cause-effect relationships implicit within the policy statement. In a sense strategies form a bridge between scientific knowledge and its application to the resolution of real-life problems. Without a scientific base, strategies would be little more than "random stabs in the dark."

What we have called "strategy" above is what is generally referred to in the policy analysis literature as "policy content"[1] and most of the typologies Froman calls "abstract policy categories" are in effect typologies of this construct.[2] The so called "traditional policy

---

[1]See the essays by Ranney, Van Dyke, Froman, and Salisbury in Austin Ranney, ed., <u>Political Science and Public Policy</u>.

[2]Included among these are: (1) the distinction made by Berelson, Lazarsfeld, and McPhee between "style" and "position" issues; (2) the "material" and "symbolic" satisfaction dichotomy proposed by Edelman; (3) Huntington's distinction between "structural" and "strategic" policy issues; (4) the classification of policy issues as "distributive," "regulatory," and "redistributive" suggested by Lowi; (5) Salisbury's reclassification of the Lowi trichotomy (distributive, regulatory, redistributive, and self-regulatory); (6) Froman's distinction between "areal" and "segmental" city policies; and (7) the distinction made by game theorists between "zero-sum" and "non-zero-sum" policies. For sources of these classifications see the essays by Lewis A. Froman and Robert H. Salisbury in Austin Ranney, ed., <u>Political Science and Public Policy</u>.

categories"[1] in fact refer to the "boundaries" of the policy as opposed to its "content."

Policy Implementation and the Policy Process

Implementation is but one link in a temporal chain commonly referred to as the "policy process." Our purpose here is to describe and discuss the place of policy implementation within this process.

We are using the term "process" here in the same sense it is used in industry. When we talk of the production process we generally refer to a series of stages the manufacturing of a good goes through. Such processes usually start with a "design" stage, followed by a sequence of stages which reflect the conversion of inputs into outputs through the application of a given technology. The process usually ends with a stage such as "quality control" and/or "packaging." At this point another process, marketing, takes over and carries the product until it reaches the final consumer.

---

[1]Included among these are the following: (1) categories defined by substantive policy areas (health, manpower, education, etc.); (2) categories defined by the specific decision-making authority involved with the policy (presidential, congressional, city-council, etc.); (3) categories defined by the level of the decision-making authority within the governmental hierarchy (federal, state, regional, local, etc.); (4) categories defined by the type of target group (human resource, institutional, environmental, etc.); (5) categories defined by the subgroups within a particular type of target group (aged, handicapped, poor, youth, etc.); and (6) categories defined by time periods (Nixon Administration years, post-Vietnam era, depression period, space age, etc.). See Froman's essay in Austin Ranney, ed., Political Science and Public Policy.

The term process, therefore, usually refers to a set of sequentially-related stages. The stages are also time-related and the outputs of one constitute part of the inputs of another. The process is a transformation process. Inputs, such as managerial, labor, material, and technology are transformed into desirable (and some undesirable) outputs.

The policy process, at least in concept, is not significantly different from the production process of industry. Much like the production process, the policy process refers to a transformation, conversion of inputs into desirable (and some undesirable) outputs through the machinery of the government. It is the policies that guide this transformation. They specify the desirable outcomes (through the goals) and how these outcomes are to be achieved (through the strategies). However, policies, in and of themselves, are no more than designs for action, much like the designs of production processes. Policies as designs (and as statements on pieces of paper) do not produce the desired changes in people's lives. It is the implementation of policies that produce these changes.

Where does the policy process start and where does it end? This is not an easy question to answer. In most instances specification of the starting point of the process and the stages that follow it are up to the investigator and the purposes of his investigation. In our case we view the policy process as consisting of the following four major

components:

- formulation

- operationalization

- implementation

- evaluation

Briefly, *policy formulation* refers to emergence of a policy issue and development of a strategy which addresses that issue. It covers the period up to the adoption of a particular policy and, in the case of most federal policies, it ends after the legislative branch of government passes an act on the policy issue in question. *Policy operationalization* bridges the gap between formulation and implementation. It refers to preparation of rules, orders, procedures, plans, and manuals which refine the adopted policy and guide its implementation. *Policy implementation,* in most general terms, refers to execution of operationalized policies. Finally, *policy evaluation* involves assessment of the effectiveness of the policy in question.

How do we define implementation. Better yet, can we define it any other way than in Webster or Roget? Perhaps not. Synonyms provided by Webster ("to carry into effect; fulfill; accomplish")[1] and Roget (carry out, execute, produce, accomplish, perform)[2] fully convey the meaning of the verb "to implement." The thing that is executed,

---

[1] See: <u>Webster's New World Dictionary, College Edition</u>. (1966).

[2] These are selected synonyms from <u>Roget's International Thesaurus, Third Edition</u>. (1962).

carried out, or accomplished, of course, is a policy as defined and described during its operationalization.

When operationalization is extremely detailed, the degree of flexibility given the implementing agency is usually limited. In such instances an implementor is frequently judged in terms of whether he has "complied with" the instructions outlined in the guidelines. In a sense, therefore, implementation is like carrying out an order. The more specific the order, the less is up to the implementor to decide what actions he should take. In military or pseudo-military settings both the incentives for successfully implementing an order and the sanctions against not complying are well specified. In these settings the implementor knows, prior to implementation, how he will be judged on his performance. Complying or not complying with the order becomes a matter of individual decision making with known outcomes.

In most human service organizations, however, the incentives for compliance and the sanctions against noncompliance are not known to the implementors. In fact, in most instances, it is not even known what constitutes successful implementation. This usually reflects a deficiency in the components of the policy process which precede implementation.

This brings us to the question of what constitutes successful implementation. Before answering this we must make a distinction between policy evaluation and evaluation of policy implementation. This distinction is necessary

because it is quite feasible for an implementor to be extremely successful in implementing a policy which, over-all, does not generate any of the desired effects. This is the case of "bad policy-good implementation". The implementor is not at fault. Failure of the policy is either due to poor policy formulation, improper operationalization, or both.

Thus, in evaluating the implementation of a policy success/failure is judged by comparing the immediate outcomes of the implementation with the original implementation objectives. In evaluating the policy overall, however, we compare the effects--usually called "impacts"--of the policy with the overall goals the policy was designed to attain in the first place.

Theory

The principal aim of this dissertation is explanation of the differences between successful and unsuccessful implementation of a public policy. We want to find out why some implementors are more successful than others and what factors account for the observed variations in their levels of success.

In Chapter II we reviewed some of the previous work on implementation and summarized the key factors proposed by others as potential explanator variables. It became apparent during the discussion in Chapter II that most of the explanations attempted were nonsystematic and/or nonempirical. The

major empirical studies covered concentrated on the implemen-
tation of a policy in one or a few cases, and, therefore,
did not allow formulation of strongly supported empirical
generalizations. Single program case studies or comparative
analysis of policy implementation in a few cities, despite
their value in helping us understand some of the factors
accounting for poor implementation, have not provided us
with broad, systematic, and sound conceptual structures
which can be used for generating a set of testable hypoth-
eses about implementation. Our aim in this section is to
attempt to fill this gap and pave the road for empirical
confirmation or disconfirmation of some hypotheses about
the specific relationships between implementation success
and some explanatory variables.

We have divided this section into five parts. In the
first part we introduce a conceptual model of policy imple-
mentation. In the second part we review the main elements
of expectancy theory in organizational psychology which
provides strong theoretical support for the conceptual model.
In the third part of the section we define, describe, and
discuss the six variables of the model. In the fourth part
we examine each of the two functional relationships and for-
mulate the alternative hypotheses that are tested in Chapter
V. The final part of the section is devoted to clarifica-
tion of a set of methodological issues concerning the con-
ceptual model and its theoretical bases.

The Conceptual Model

The primary concern of policy scientists is explanation of policy success. Understanding the ingredients which make a policy more successful is essential for increasing the responsiveness and the effectiveness of governmental activity.

Policy success, when simply defined, refers to the difference between goals and accomplishments. The closer the eventual effects of the policy are to its original goals, the more successful is the policy. The goals and objectives referred to here are the ones identified and defined during policy formulation. They reflect the overall direction of the policy and the specific targets to be attained at the end of its implementation.

These goals may be, and sometimes are, different from the goals used in assessing the successfulness of policy implementation. The overall goals of the policy may be stated over a longer time span than the implementation goals. Or, the overall goals may reflect national targets, whereas the implementation goals may vary between the implementing agencies. And, sometimes, the policy goals may be modified during policy operationalization, resulting in the generation of a different set of goals for implementation. The term policy success, when defined as above, does not specifically refer to successful implementation and a sucessfully implemented policy may turn out to be a failure when evaluated in terms of the overall policy goals. Thus, when we refer to level of policy success we talk about the overall

successfulness of the policy as measured against the policy goals identified during policy formulation. Alternatively, we use the term level of implementation success (LIS) when referring to the successfulness of policy implementation as measured against the implementation goals. Our primary emphasis in the dissertation, of course, is on explanation of the variations in LIS.

The overall success of the policy depends to a large extent upon what takes place between statement of the goals and the measurement of the results. If the policy formulated during the first stage of the policy process is deemed to be effective, i.e., if there is strong theoretical, causal, and other rationale strongly supporting the point that the formulated policy will lead to the desired effects, the chances for the success of the policy will be large, other things being equal. Similarly, if the formulated policy is interpreted, operationalized, and otherwise put into action with minimum deviation from its original specifications and with full understanding of the circumstances which may potentially affect its implementation, the chances for policy success will similarly be large, other things held constant. Finally, if the policy is implemented exactly as it is operationalized and if the implementation objectives are attained or surpassed, the policy's chances for success will be substantially enhanced . . . again, other things being equal. Successful accomplishment of the goals assigned to each stage, therefore, will

ensure the attainment of the policy's overall goals. And, because of the strong linkages between the three stages, a failure in any one stage may increase the chances for a sizable discrepancy between the desired goals and the actual results. If the failure takes place during the formulation stage, in all likelihood, an ineffective policy will be implemented. If operationalization is done unsuccessfully, what is eventually implemented is likely to be different from the original policy. If however, the first two stages are completed successfully but implementation turns out to be a failure, the original goals of the policy will still remain far ahead of the actual results. Thus, the overall policy success is largely affected by the level of success achieved at each of the three stages of the process.

As the preceding discussion indicates, our conception of the policy process--its stages, interrelationships between the stages, and the role played by each stage in the achievement of policy success--necessitates, at the very least, three different explanations: explanation of the variations in the levels of success reached during policy formulation, policy operationalization, and policy implementation. Here, we are principally concerned with the last of these three explanations. We recognize the importance of policy formulation and operationalization in the achievement of policy success. However, our study of these two aspects of the policy process does not go beyond this recognition.

How does one go about explaining the overall policy

success and/or level of success at each stage empirically?
Here we will discuss three alternatives.

The first one is what we will call the policy-specific
approach. Here the emphasis is on explaining the success-
fulness of a single policy. Explanation of success during
policy formulation usually takes the form of a case study,
since there is not any variation as such to study. Opera-
tionalization, if done at a central level, is also studied
using the case study method. Of course, decentralized
operationalization can be studied by collecting and analyz-
ing data on each operationalizing unit. The heaviest
emphasis while using the policy-specific approach is placed
on implementation since there are usually quite a few imple-
menting agencies executing the policy.

We have entitled the second approach the policy area
approach. Here explanation centers on several individual
policies within a broad policy area, such as housing, educa-
tion, welfare, etc. Because of the presence of several
policies, formulation, operationalization and implementation
can all be studied using empirical methods that capture the
variation between policies and between agencies.

The third approach entails empirical investigation
of all policies at a given level of government, regardless
of the specific policy area, and can be entitled the non-
policy-specific approach. The emphasis here is to reach
generalizations concerning policy formulation, operational-
ization, or implementation at the federal, state, or local

level. As in the previous approach, variations among policies, policy areas, and agencies allow the analyst to employ a wide variety of empirical methods.

In this dissertation our interest lies with the first of these, the policy-specific approach. We are involved with empirical investigation of the implementation of a single policy across several implementing agencies. Formulation and operationalization were done centrally, by the Congress and the Office of Child Development (HEW), respectively. The fact that formulation and operationalization were done centrally implies that *the effects of policy formulation and policy operationalization variables on LIS can be assumed to have remained as constants.* Each implementing agency can be viewed as having been subjected to the same effects of formulation and operationalization, and, thus, these two variables can be assumed to have been controlled for. This implies that in attempting to explain LIS we need not be concerned, except for descriptive purposes, with what has taken place during the early stages of the policy process. The types of variables we should consider for the explanation should reflect salient aspects of the implementation stage alone. The conceptual model we introduce below satisfies this requirement.

We contend that level of implementation success, our primary dependent variable, is a function of two key explanatory variables. These are <u>implementation potential</u> (IPOT) and <u>level of effort</u> (EFRT). Briefly, implementation potential refers to the overall capacity of the agency to implement

the policy in the same manner as it is operationalized. It indicates the presence/absence of a set of capacity-related factors which would be required for effective and efficient execution of the policy. Level of effort refers to the amount of policy-specific effort expended during the implementation.

Our secondary dependent variable is level of effort and our contention is that it can be explained in terms of three independent variables: environmental forces for implementation (ENVR), policy support (PSUP), and self-evaluation of capability (CAPB). Environmental forces for implementation refers to the net balance of the forces for and against the implementation of the policy. It reflects the net effects of the pressures and influences directed towards the implementing agency by other agencies and interested groups of individuals. Policy support refers to the position taken by the leadership of the implementing agency vis-a-vis a positive advocate of the policy. It reflects the extent to which the leadership of the implementing agency agrees with and supports the policy. Self-evaluation of capability refers to the perception of the leadership of the implementing agency regarding the chances that the agency can implement the policy successfully.

The conceptual model which connects the four independent variables and one intervening variable to LIS, our primary dependent variable, is illustrated in Figure 1. Detailed description and discussion of the variables and of the hypothesized relationships between variables are the

subject matters of two later parts in this section.
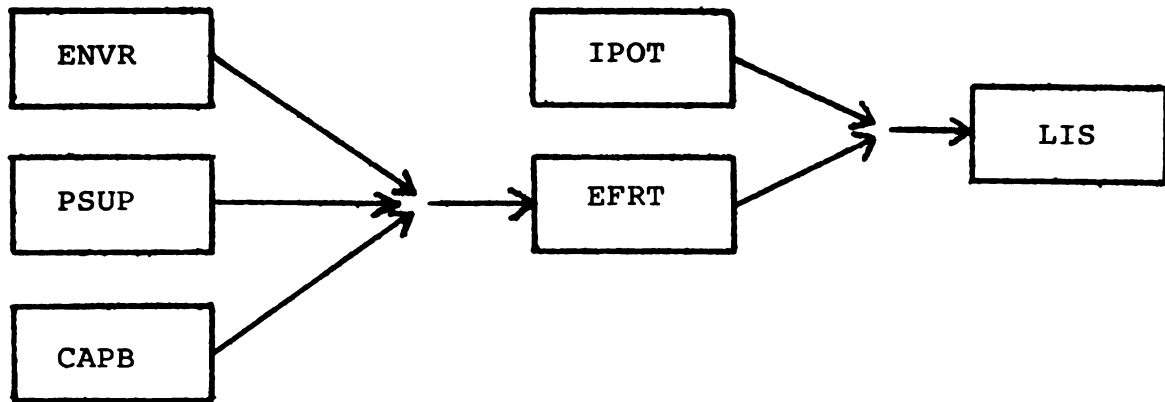
Figure 1. The proposed conceptual model.

We now move to a discussion of the theoretical support for this conceptual model.

Theoretical Support for the Conceptual Model

The conceptual model we just introduced is new for explaining an aspect of organizational performance. However, it is not altogether new because an analytical scheme similar to this one has long been recognized as one of the most powerful theories explaining variations in employee performance in organizational settings. This is the so-called *expectancy theory* in organizational psychology and it deals with motivational determinants of the performance of an individual. Our conceptual model, therefore, can be viewed as an extension of this theory from the level of an individual to the level of an organization responsible for implementing a public policy.

Porter and Lawler and Vroom trace the origins of

expectancy theory to Tolman and Lewin, two of the early advocates of cognitive theories of behavior.[1] Basic to these theories is the argument that (1) behavior is purposeful and goal-directed, (2) individuals have "beliefs" or "expectations" about future events, in particular, about the potential outcomes of the actions they could take, and (3) individuals behave in a manner which they think will lead to the attainment of goals they value.

"Expectancy" is a cognitive concept and it refers to a person's subjective assessment of the chances that a particular act will yield a particular outcome. The same concept has sometimes been called "subjective probability," indicating that expectancies may take on values between zero (for the case where the individual thinks that the specific act will never result in the valued outcome) and one (he is certain that his act will be followed by the outcome).

Expectancies enter the theory indirectly, as one of the determinants of "motivation," the key explanatory variable of the model. In addition to motivation the theory uses two other predictor variables to explain variations in performance: ability and organizational characteristics. Ability, in general, refers to the individual's capacity to

---

[1]See: Lyman W. Porter and Edward E. Lawler, Managerial Attitudes and Performance (New York: Dorsey-Irvin, 1968) p. 9; Victor H. Vroom, Work and Motivation (New York: John Wiley and Sons, Inc., 1964) pp.13-14. The specific works of Tolman and Lewin cited by Porter and Lawler and Vroom are: E.C. Tolman, Purposive Behavior in Animals and Men (New York: Century, 1932) and K. Lewin, The Conceptual Representation and the Measurement of Psychological Forces (Durham, N.C.: Duke University Press, 1938).

perform a given task or a set of tasks. Defined as such, the term covers a host of individual characteristics. Included among these are intellectual factors (verbal, numeric, symbolic, spatial skills and the like), manual factors (such as strength and dexterity), and personality traits. It reflects the capabilities of the individual, i.e., those characteristics which enable him to accomplish a specific task.

Organizational variables enter the model in the form of goal-setting and extrinsic and intrinsic reward mechanisms.

Motivation, the most interesting of the three predictor variables, refers to what we have called "effort," a characteristic which reflects how vigorously the individual will use his abilities in performing a given task. Others have given different names to this concept. Vroom reports that it is similar to the Lewinian concept of "force," Tolman's "performance vector," Atkinson's "aroused motivation," Luce's "subjective expected utility," and Rotter's "behavior potential."[1] Two aspects of motivation, its intensity and its direction, determine how much "force" the individual will exert in performing the given task. Intensity of the individual's motivation reflects his overall state of arousal to engage in any behavior. Direction shows whether his motivation is channeled towards performance of the specific task in question.

Intensity of the individual's motivation is measured

---

[1]See Vroom, Work and Motivation, p. 18 for full reference on these concepts.

in terms of the attractiveness to the individual of various outcomes that could follow the performance of the task. This amount of attractiveness is referred to in the model as the *valence* of a given outcome. There are two types of goals or outcomes which might motivate an individual. Extrinsic outcomes refer to factors such as pay and promotion. Intrinsic outcomes refer to factors such as feelings of accomplishment and satisfaction. The types of extrinsic and intrinsic outcomes that are perceived as valent may differ from individual to individual. Thus, the theory suggests that by focusing on the valence of outcomes one can assess the individual's intensity of motivation to perform the task.

The directionality of motivation, or whether the individual will direct his efforts towards high performance, is viewed as a function of two "contingencies" which can be both expressed in probability terms. Cummings and Schwab[1] summarize these as follows:

> " . . . First, the employee must feel that if he attempts to be a high performer, he has a good chance of doing so. This contingency is referred to as the individual's *expectancy* perception, or *effort-performance probability*. A worker might feel that his effort-performance probability is low because he lacks the ability to perform well, because the task itself is poorly structured for successful performance, or for a variety of other reasons.
>
> "The second contingency is actually a set of contingencies that refers to the individual's expectations that high performance will result in various valent intrinsic and/or extrinsic outcomes. These contingencies are frequently referred to as the

---

[1]L. L. Cummings and Donald P. Schwab, Performance in Organizations (Glenview, Illinois: Scott, Foresman and Company, 1973) .

individual's *instrumentality* perceptions, or *performance-reward probabilities*."[1]

Thus, according to expectancy theory, the individual first examines the task to be performed and assesses the chances that he will be successful if he attempts to perform it. He then calculates the chances that his high performance will lead to a specified set of extrinsic and intrinsic outcomes. Finally, he evaluates the level of attractiveness of these outcomes to himself. If either or both probabilities are low or if the individual does not see the outcomes to be positively attractive to him, the "motivation" portion of the theory predicts that the individual would not be a high performer.

There is a large body of literature relating to one or more aspects of expectancy theory. Unlike most other theories of performance, expectancy theory is general enough to be applicable to a wide variety of situations. According to Cummings and Schwab, the empirical evidence on the predictive ability of expectancy theory is quite promising:

> "The research conducted on expectancy theory to date looks very promising for the basic motivational hypotheses of the theory. Clearly, additional research is needed, and limitations in the theory may eventually surface which are not evident at the present time. Nevertheless, the evidence already available strongly suggests that an expectancy theory approach to understanding and predicting organizational performance behavior is fruitful."[2]

Findings from one of the more recent empirical studies on the predictive power of expectancy theory are reported by

---

[1] Ibid., p. 30.
[2] Ibid., p. 33.

Porter and Lawler. In this investigation Porter and Lawler tested a set of propositions derived from a conceptual model similar to ours, using survey data collected from close to 500 managers in seven organizations. Three of these organizations were divisions of state governments. Specific variables on which data were collected included: manager's attitudes towards their pay, their perception of the type of role behavior required for success on the job and the degree of need satisfaction provided by the job, self-rating of job performance and effort expended, and superior's ranking of job performance and effort expended. Only limited data were collected on abilities.

Porter and Lawler summarize their empirical findings as follows:

> " . . . Managers who saw pay closely tied to performance factors received higher performance and effort ratings than managers who did not see such a close relationship (i.e., who had lower effort-reward probabilities). . . . Furthermore, the strongest relationships between perceptions of pay being based on performance factors and our measures of effort existed for those managers who attached the greatest importance to pay as a reward. This latter finding not only supports the model's contention that both *value of reward* and *perceived effort-reward probability* are involved in determining a manager's *effort,* but it also indicates that the form of the combination of these first two variables in the model is interactive rather than additive."

> " . . . role perception-performance relationships were stronger for managers who were seen as exerting high effort (i.e., were "highly motivated") compared to those seen as exerting relatively low effort. Thus, not only is the notion of a combined effect of *effort* and *role perceptions* on *performance* supported, but also the results indicate the relationship may well be an interactive one. In other words, if extremely high effort were to be combined with extremely inaccurate role perceptions, the prediction would be that the subsequent performance would be evaluated as relatively ineffective."[1]

---

[1]Porter and Lawler, <u>Managerial Attitudes and Performance</u>, pp.160-161.

On the effects of ability and motivation on performance Vroom reports: "The evidence which does exist suggests rather strongly that ability and motivation, as typically measured or manipulated, do not have independent effects on performance but rather interact with one another."[1] The empirical studies cited by Vroom include the following:

- Elizabeth French's study of the joint effects of intelligence and achievement motivation on the problem solving success of airmen;

- Fleishman's experimental study on the effects of ability and motivation on the performance of 400 air force trainees in a complex coordination task;

- Vroom's study of the effects of participation in decision making on supervisor's satisfaction and performance;

- Wyatt's experiments on the effects of economic incentives on the level of performance achieved by female workers in a candy factory.[2]

Clearly, most of these studies are isolated attempts to test one or more aspects of expectancy theory. Although the "universality" of the theory has not been empirically established, the extent of empirical support from different settings indicates that the major propositions of the theory (i.e., those relating (1) ability and motivation to performance and (2) valance, expectancy, and instrumentality to motivation) have held in most test situations.

There are quite a few similarities between the structure of expectancy theory and the conceptual model we

---

[1]Vroom, Work and Motivation, p. 200.

[2]Full references for these studies can be found in Vroom's Work and Motivation, pp. 200–204.

introduced earlier. In particular, there are four parallels
between the components of expectancy theory and the vari-
ables encompassed by our conceptual model.

First, both structures contain a variable which ref-
lects capability. Expectancy theory talks about task-ori-
ented abilities of the individual. In our model this is
paralleled by the concept of implementation potential which,
to a large measure, reflects the agency's capability to
implement the given policy.

Second, the notion of motivation in expectancy theory
is very similar to the concept of effort used in our model.
Expectancy theory predicts that highly motivated individuals
will be more successful than those with low motivation,
other things being equal. We predict that, controling for
their implementation potential, agencies with higher levels
of policy-specific effort will attain a greater level of
implementation success than those with low effort.

Third, the concept of valence in expectancy theory
resembles our notion of policy support. Expectancy theory
refers to the attractiveness to the individual of a given
goal or outcome; we are talking about the level of support
of a given policy by the leaders of the implementing agency.
Expectancy theory argues that if the goal or outcome is
highly attractive to the individual, he will be highly moti-
vated to perform the task. We are postulating that if the
leadership of the agency support the policy strongly they
will take appropriate actions to put forth the effort

necessary to implement it fully.

Fourth, and finally, the <u>effort-performance probability</u> of expectancy theory is parallelled by our notion of <u>self-evaluation of capability</u>. Expectancy theory suggests that individuals who feel that they can, with high probability, successfully accomplish the given task are more likely to perform better than those who view this probability as low. According to our conceptualization, agencies which believe that they have the capabilities to implement the policy are more likely to exert the necessary effort than those who have less confidence in their capacity.

We realize that similarities between the variables and the structure of expectancy theory and those of our conceptual model cannot alone be regarded as constituting full theoretical justification for our explanatory scheme. This is especially true since expectancy theory has been formulated for explaining the performance of an individual, not the success of an agency in implementing a public policy. However, extension of a "micro" theory to a "macro"-level is not uncommon and, if the micro theory happens to have been empirically verified in a number of different situations, such an extension may offer the most promosing starting point for theory building at the macro-level.

This completes our discussion of the theoretical support for the conceptual model. We now turn to a detailed description and discussion of each of the six variables included in the conceptual model.

Variables of the Model

Level of implementation success

This variable refers to how much successful goal achievement has been accomplished as a result of the implementation of the policy in question. As indicated earlier in this section, the yardstick used for assessing the level of goal achievement is the set of implementation goals identified and defined during policy operationalization. In many instances, as in the case of the specific policy studied here, these goals coincide with the policy goals identified during policy formulation.

Implementation success is a measure of organizational performance concerning the execution of a specific public policy. If the agency's mission does not extend beyond the execution of a single public policy, implementation success and organizational performance become identical. On the other hand, if the agency in question is responsible for simultaneous implementation of two or more public policies, the agency's success level in implementing one policy may differ from its success in executing the other policies. In general, however, the overall organizational performance of the agency is strongly correlated with the average success level it attains in implementing all the policies it is mandated to execute.

Assessment of the policy's implementation requires "outcome" rather than "impact" indicators. Briefly, the term "impacts" refers to lasting changes in the lives of people

which have come about as a result of implementation of the policy in question. Outcomes, on the other hand, refer to the immediate (and more visible) effects of the policy. To give an example, the number of hours a participant has been exposed to skill training may be one of the outcomes of a manpower program whereas the program's impact would most likely be measured in terms of the change in the partici- pant's employability that can be attributed to the program in question. Similarly, the immediate outcome of a social services program might be the number of units of counseling and other services provided. The impact of such a program, however, is usually assessed in terms of the extent to which these services have made the recipients economically self-sufficient or less dependent on the welfare system.

The policy goals (or desired impact statements) estab- lished during the policy formulation stage are translated into implementation objectives (or desired outcome statements) during policy operationalization. Thus successfulness of the implementation is purely a matter of level of attainment of the policy's implementation outcomes and not its desired impacts.

Level of effort

By the term effort we are referring to the amount of "energy" expended by the implementing agency in the execution of the policy in question. In simple terms, with this vari- able we are attempting to find a response to the question, "How hard is the agency trying to implement the policy?"

The kind of effort we are talking about is human effort, made up of the individual efforts of persons affiliated with the agency--managers and nonmanagers; professionals, para-professionals, and nonprofessionals; volunteers and nonvolun-teers; full-time staff and part-time staff.

Our definition of effort implies that for a given agency one can talk of an absolute amount -- called total effort -- which is the sum of the individual efforts of all persons affiliated with the agency over a specific period of time. Theoretically, for each individual this amount can be partitioned into different policies and projects the individual has worked on during the period in question. Thus, we can theoretically talk about policy-specific effort for an individual and for the agency as a whole.

In our simple conceptual model we are studying effort as a single variable. In other instances one can define several effort-related variables. For example, instead of total policy-specific effort one can define it in terms of dichotomies such as managerial-nonmanagerial effort, profes-sional-nonprofessional effort, etc. In other instances, total effort could be divided into task-specific effort depending upon the implementation requirements of the policy. Such a breakdown might include as separate variables or as components of a total effort score items such as the follow-ing: effort to organize for the implementation, effort to plan the implementation, effort to coordinate the implementa-tion, effort to mobilize resources, effort to staff, etc.

Effort plays an important role in the model mainly for two reasons. First, it is intuitively obvious that without effort it is virtually impossible to achieve implementation success. It is also quite reasonable to expect, other things being equal, higher levels of effort to be associated with high success. Although there might be a point after which a "diminishing marginal returns" phenomenon takes place, it would be highly unlikely to observe a weak relationship between implementation effort and implementation success. Second and more importantly, within the model effort plays an intervening variable role. From a substantive standpoint, it is more reasonable to expect a causal linkage between ENVR, PSUP, and CAPB and effort than between these three variables and implementation success. It makes more sense to expect policy support to be causally related to effort than to implementation success. High level of support of the policy by the leadership of the agency does not directly and automatically lead to the policy's success. Such support, however, may lead to a greater level of effort which, other things being equal, may lead to greater success. Thus, effort is included in the model not only for its intuitively obvious relationship to implementation success but also for the central role it plays as an intervening variable by connecting the other variables to LIS, the primary dependent variable. Although explanation of implementation success is our primary task, in light of the role of effort in the model, it is equally important that we explain the variation in

effort via other variables.

Implementation potential

This variable refers to the policy-relevant capabilities of the agency to implement the policy and, as such, it relates to relatively stable characteristics of the organization. Included here are policy-specific factors such as staff capabilities, leadership capabilities, monetary and other resource capabilities of the agency. It can be assumed that these capabilities stay relatively constant over the short run, i.e., they do not fluctuate widely over short periods of time. It is possible, of course, to modify these capabilities through change producing activities such as training, adding new staff, moving to new facilities, or acquiring new equipment/materials. For all practical purposes it can be assumed that the potential of an agency to implement a new policy remains stable over the period of one year.

In a sense, implementation potential illustrates the prior probability of implementation success, other things being equal. It shows the variation among implementing agencies in initial conditions that exist prior to start of the implementation. This reflects the comparative advantage/disadvantage of each implementing agency vis-a-vis others at the start of the implementation. It is a highly policy-specific variable; the actual make-up of the variable is largely dependent on the specific policy being examined. To

give an example, the staff of an agency may be highly qual-
ified and capable to implement a given policy, yet it may
be viewed as less capable to implement a different policy,
one that demands a different mix of staff qualifications.

Implementation potential places an upper limit on the
level of success an agency can achieve as a result of the
effort it expends in implementing the policy. An agency
with zero potential, although it would be unlikely to find
such an agency, would be expected to achieve only a low
level of success, regardless of the amount of effort it
spends in implementing the policy. Thus, given two agencies
which have expended equal amounts of effort in implementing
the same policy, the one with greater potential would be
expected to achieve a higher level of implementation success,
other things being equal.

Environmental forces for Implementation

This variable refers to the net effects of two types of
forces: forces for and forces against the implementation of
a given policy. Included in these forces are pressure
directed towards the implementing agency by other agencies,
agencies other than those that administer or implement the
policy, and by interested individuals or groups of individ-
uals.

The specific make-up of this variable is highly depen-
dent upon the individual policy being studied and the envi-
ronmental factors that are relevant to the implementation of

the policy. In the case of most federally formulated and locally implemented policies, reactions to the policy from the competing community agencies and the members of the target group become particularly important. This is especially true in cases where policies similar to the one studied are being implemented by other agencies in the community. It is also important in instances where implementation of the policy in question requires the cooperation of and/or inputs from particular agencies in the community.

The assumption underlying our inclusion of this variable in the model is that, other things being equal, implementation agencies are influenced by the pressures from the environment in making their decisions about the level of effort they will undertake in implementing the policy. If there are no pressures one way or another from the outside, other factors, such as policy support and self-evaluation of capability, will determine the level of implementation effort. If there is strong environmental pressure for the implementation of the policy, this pressure will induce the implementing agency to intensify its effort to execute the policy. On the other hand, if there is strong pressure against the implementation, the implementors are more likely to cut back their effort.

Policy support

This variable refers to the extent to which the leadership group of the agency, i.e., the group of persons within

the hierarchy of the implementing agency who are in positions
of policy-related decision making, support the policy in
question. Agencies in which the policy receives strong sup-
port from the leadership are more likely to allocate the
effort necessary to implement the policy. Conversely, if
the leadership of an implementation agency opposes or does
not support a given policy it is highly likely that that
agency will exert only little effort towards implementing
that policy.

Support for a policy requires, first of all, the agree-
ment of the leaders of the implementing agency with the pol-
icy as a whole. Included among the other determinants of
policy support are the attitudes of the leaders towards the
specific target group to be served according to the policy.

Why the leader of a particular agency does or does not
support a new policy is a complicated matter to assess. He
may agree with the overall intent of the policy but disagree
with the adopted strategy; he may not like the manner in
which it has been operationalized and "passed down" to the
local level; he may think that other agencies are better
equipped to serve that specific target group; or he may have
negative attitudes towards the target group. In the case of
a marginal policy, he may oppose it because he feels that his
agency is already overworked and understaffed and that the
additional effort required for implementing the new policy
may adversely affect the implementation of other policies
concurrently executed by his agency.

In general, policy support reflects the leaders'
estimate of the value of the rewards (positive or negative)
to his agency associated with successful implementation.
Accordingly, he strongly supports a policy if he values
highly the outcomes of successful implementation. These
outcomes may include intrinsic and extrinsic rewards such
as better recognition of his agency by other agencies in
the community, by the public, and by the agency administer-
ing the policy; attainment of a greater status for his
agency; greater cooperation from other agencies; greater
funding; and a sense of accomplishment and self-satisfaction
for himself and his staff. He balances these positive out-
comes against the potential negative rewards associated with
non-implementation or poor implementation. These negative
rewards may include outcomes such as cuts in funding, poor
image, poor relations with other agencies, etc. The value,
to the leader, of the net balance of the potential positive
and negative effects of successful and unsuccessful implemen-
tation determines the extent to which he supports the policy
in question. The actual level of effort spent in implement-
ing the policy according to the proposed model, is a function
of this level of support and two other variables: environ-
mental forces for implementation and self-evaluation of
capability.

Self-evaluation of capability

This variable refers to the capability of the implementing agency to execute the policy successfully, as evaluated by the leadership of the agency. It is a subjective measure and therefore may not coincide with other, more objective measures of capability (such as implementation potential). It illustrates the self-confidence of the leadership in their organization to execute the policy effectively. It can be expressed as a probability measure, as the subjective probability, estimated by the leader of the agency, that the agency can implement the policy successfully.

In the conceptual model self-evaluation of capability is introduced as one of the determinants of level of effort. It is hypothesized that, controling for environmental forces and policy support, agencies with higher self-perceived capabilities to implement the policy are more likely to put a greater effort to its implementation than those with low confidence in their capabilities.

This completes our definition and discussion of the primary variables of the conceptual model. In the next part of the section we will concentrate on the hypothesized relationships between these six variables.

Relationships Between Variables

There are two major functional relationships illustrated in the proposed model. The first relationship connects EFRT and IPOT to the primary dependent variable, LIS. The

second relationship ties ENVR, PSUP, and CAPB to the secondary dependent variable, EFRT. We will now examine these two relationships and formulate alternative models for each relationship. These models constitute our alternative hypotheses regarding each relationship, which are tested empirically in Chapter V.

Relationship of EFRT and IPOT with LIS

In general, our hypotheses regarding the relationship of EFRT and IPOT with LIS are the following: (1) controling for IPOT, LIS increases as EFRT increases, and (2) controling for EFRT,LIS increases as IPOT increases. Stated differently, given two agencies with the same level of implementation potential, the one with the greater level of effort will achieve a greater level of success than the other. Similarly, of two agencies with equal levels of implementation effort, the one with the higher potential is predicted to achieve a greater level of success.

It is intuitively obvious that IPOT and EFRT should be related to LIS in the direction predicted by the two hypotheses stated above. What is not so clear is the specific form of the functional relationship which connects these three variables. In this respect we have formulated and will test the explanatory power of three alternative models. All three are linear models and for easé of identification we have labeled them as the *additive model,* the *interactive model,* and the *hybrid model.*

The additive model.--According to this model, LIS is
a simple linear function of EFRT and IPOT:

$$LIS = a + b(EFRT) + c(IPOT) + e \qquad (1)$$

where LIS, EFRT, and IPOT are as before, b and c are regres-
sion coefficients, a stands for the constant term and e for
the error term.

This model hypothesizes that EFRT and IPOT have inde-
pendent and separate effects on LIS. Agencies with high
levels of EFRT and IPOT are predicted to achieve high levels
of success. Conversely, agencies with low levels of IPOT
and EFRT will achieve very little success. Between these
extreme values of IPOT and EFRT, agencies with high EFRT
but low IPOT and those with low EFRT but high IPOT are
predicted to achieve moderate levels of success. Thus,
according to this model, it is possible for an agency with
very high implementation potential to reach a moderate to
high success level with only minimal effort. The high poten-
tial level compensates for the low effort and, in effect,
increases the productivity of each unit of effort. Alter-
natively, an agency with a very high effort can achieve
moderate to high success even though it may have a low
implementation potential. In this case high effort com-
pensates for the low potential and, even though each unit of
effort has relatively low productivity, the agency still
attains moderate to high success.

The presence of EFRT and IPOT in a relationship explain-
ing LIS makes intuitive sense. The primary independent

variable of the relationship is EFRT since it is policy-specific effort that leads to results. The role of IPOT, as we see it, is secondary. As we have argued above, for a given level effort IPOT serves as a measure of productivity and enhances or dampens the effect of EFRT on LIS.

The predictions based on the additive model are quite reasonable for most combinations of values EFRT and IPOT could take. In two instances the predictions become somewhat cloudy. First, in the case of zero or close to zero effort and very high potential, the model predicts a moderate level of success. This may be questioned since it is unreasonable to expect even moderate success without any effort whatsoever going in the direction of policy implementation. It is like work being produced without force in physics. The second case is the exact opposite of the first one. When EFRT is extremely high but IPOT is zero the model predicts moderate success. One may question if it is plausible for an agency with none of the capabilities required for successful implementation to achieve a moderate success level with increased effort.

The answer to this potential weakness of the additive model lies in one of the other variables, self-evaluation of capability, which is proposed as one of the explanators of level of effort. Since it would be quite unlikely to observe a very wide discrepancy between CAPB, which stands for a subjective measure capability, and IPOT, which stands for an objective measure of capability, when IPOT is extremely

high CAPB cannot be expected to be low. In such a case CAPB would have a positive effect on EFRT and reduce the inconsistency between IPOT and EFRT. Similarly, when IPOT is zero one cannot ordinarily expect CAPB to be high. This would have a dampening effect on EFRT and again reduce the wide gap between the relative levels of IPOT and EFRT. Thus, even though it is theoretically possible to talk about extreme discrepancies between the values of IPOT and EFRT it is unlikely to observe these extremities in the real world. Therefore the additive model makes sense for all combinations of values IPOT and EFRT can be expected to take in an empirical investigation.

The interactive model.--According to this model, LIS is a simple linear function of the multiplicative interaction of EFRT and IPOT:

$$LIS = a + b(EFRT \times IPOT) + e. \qquad (2)$$

The model hypothesizes that LIS is determined by the joint effect of EFRT and IPOT, rather than their independent effects. When both EFRT and IPOT are high (low) LIS is predicted to be high (low). Thus, when the standardized values of EFRT and IPOT are close to one another this model's prediction is quite close to that of the additive model. However, when the values of EFRT and IPOT are far apart, results of the interactive model are different from the predictions based on the additive model. Low effort and high potential, as well as high effort and low potential, produce relatively low success. The two variables play a restraining or dampening role on one

another. Effort places a ceiling on what one can achieve
with increased potential and, alternatively, potential imposes
an upper limit on the effects of increased effort. In a
sense, the interactive model rewards agencies which have a
balance between their potential and effort and punishes those
with a wide discrepancy between their effort and potential
levels.

The logic of this model is somewhat opposite the "com-
pensation principle" of the additive model. According to
this principle agencies can make up for what they lack in
potential by putting in a greater effort. According to the
interactive model, they cannot. Accordingly, this second
model portrays a more rigid and black-and-white view of imple-
mentation than the first model. Agencies can attain high
success only if their potential and effort levels are rela-
tively high.

It is quite likely that the explanatory power of the
interactive model may be greater in cases where the implemen-
tation capabilities of the agencies studied show great varia-
tion or when only a very specialized effort is required for
successful performance. However, implementation of most
human service related policies, such as the handicapped
policy studied here, require an effort that is not extremely
specialized. Agencies implementing such policies have
substantial flexibility to adopt a course of action within
the broad constraints imposed by the operationalized policy.
Furthermore, capabilities of most of these agencies are quite

alike. Thus, we believe that the additive model might be more suitable in exploring implementation success in cases similar to the one studied here. Accordingly, we view the additive model as portaying our null hypothesis concerning the relationship between IPOT, EFRT, and LIS. The interactive model summarizes our first alternative hypothesis and the hybrid model, which we now move to, outlines our second alternative hypothesis.

The hybrid model.--This model, as its name implies, combines the additive and the interactive model and postulates that LIS is simple linear function of EFRT, IPOT, and the multiplicative interaction of EFRT and IPOT.

$$LIS = a + b(EFRT) + c(IPOT) + d(IPOT \times EFRT) + e. \quad (3)$$

The hybrid model is a compromise between the two previous models. It would be applicable in instances where there is a wide discrepancy between the predictions of the additive and the interactive models. Otherwise we do not expect the hybrid model to add much to the explanatory power of either of the two models discussed earlier.

Relationship of ENVR, PSUP and CAPB with EFRT

In general, our hypotheses regarding the relationship of ENVR, PSUP, and CAPB with EFRT are the following: (1) controling for PSUP and CAPB, EFRT increases as ENVR increases; (2) controling for ENVR and CAPB, EFRT increases as PSUP increases; and (3) controling for ENVR and PSUP, EFRT increases as CAPB increases.

There is reason to believe that the effect of ENVR to EFRT is separate from and independent of the effects of PSUP and CAPB. ENVR represents the pressures from forces exogenous to the implementing agency, whereas PSUP and CAPB are internal to the organization. Thus, ENVR can be viewed as increasing or decreasing the level of effort which results solely from the effects of PSUP and CAPB.

The specific form of the relationship which connects PSUP and CAPB to EFRT is what distinguishes the two models we are proposing here. As in the case of the relationship of EFRT and IPOT with LIS, we have formulated an additive and an interactive model. These models are described and discussed below.

The additive model.--According to this model EFRT is a linear function of ENVR, PSUP, and CAPB:

$$EFRT = a + b(ENVR) + c(PSUP) + d(CAPB) + e. \qquad (4)$$

The model hypothesizes that all three independent variables have separate independent effects on EFRT. Because of the additive nature of the model lack of effort due to the low value of any one of the variables can be compensated for by a high value for either or both of the other variables. In particular, PSUP and CAPB are viewed as having independent effects on EFRT. Thus, an agency with high PSUP and low CAPB or low PSUP and high CAPB would be predicted to undertake a moderate level of implementation effort, the final predicted level being subject to the exogenous effect of ENVR.

The interactive model.--According to this model EFRT
is a linear function of ENVR and the multiplicative inter-
action of PSUP and CAPB:

$$EFRT = a + b(ENVR) + c(PSUP \times CAPB) + e. \qquad (5)$$

Interaction of PSUP and CAPB makes substantive sense
in the case of the relationship between these four variables.
Leaders of an agency usually do not jump into the implementa-
tion of a new policy if they do not support the policy or if
they believe that the agency does not possess the capabilities
needed for doing a good job. A high value for either of
these two variables is a necessary but not sufficient con-
dition for high effort to be put forth in the implementation
of the policy.

This completes our discussion of the relationships
between the variables of the model. To recapitulate, we
proposed three alternative models for the realtionship of
EFRT and IPOT with LIS and two alternative models for the
relationship of ENVR, PSUP, and CAPB with EFRT. In the case
of both relationships the additive model was put forth as
summarizing our null hypotheses regarding that relationship.
The interactive model discussed for each relationship por-
trayed our primary alternative hypotheses. The explanatory
power of each of these models are examined in Chapter V
utilizing the data from a sample of Head Start programs. The
models with the greatest explanatory power for explaining the
implementation of the handicapped policy in Head Start are
identified and discussed in that chapter of the dissertation.

Methodological Notes on the
Conceptual Model

Our overall purpose in this dissertation is the expla-
nation of implementation success. To achieve this, we intro-
duced a few new concepts and variables and formulated a con-
ceptual model and a set of associated alternative hypotheses
which are empirically tested in a later chapter. An impor-
tant methodological question that has not been answered so
far is whether the overall approach used here can be
regarded as "proper" scientific explanation. Here we first
address this question and then discuss the adequacy of the
theoretical support for the conceptual model.

Formulation and empirical testing of a conceptual model
and a set of hypotheses, alone, do not automatically lead to
explanation of a specific phenomenon. The model and the
hypotheses need to be justified, i.e., there is a need to
draw up a rationale which backs up each major hypothesis and
a need to clarify the specific role a particular hypothesis
plays in the overall explanation scheme.

Methodologically, there are two general approaches
to the question of hypothesis justification. These are the
so-called deductive theory and the empirical generalization
approaches to the study of political phenomena.

The deductive theory approach, as its name implies,
requires a formal, axiomatized, and deductively related
theoretical system from which the hypotheses can be deduced.
The rationale for the hypothesis lies in the deductive

framework. The hypothesis simply expresses one empirical consequence of the deductive theory. If this consequence is empirically verified, the theoretical framework upon which the hypothesis is based becomes more credible. According to this approach, therefore, confirmation of a single hypothesis is directly related to confirmation of a deductive theory; no hypotheses "stand alone", without a tie to a formal deductive framework.

The empirical generalization approach does not require a deductive framework for the justification of the hypothesis. An empirical generalization is defined as "an isolated proposition summarizing observed uniformities of relationships between two or more variables."[1] Kaplan calls these "descriptive generalizations".[2] Glaser and Strauss view them as "establishing the generality of a fact".[3] Accordingly, the empirical generalization approach does not strictly demand formal justification of the hypotheses in the same manner as deductive theory does.

This does not mean that those utilizing the empirical generalization approach do not tie their hypotheses to a theory. Their goals, like the goals of deductive theorists, are explanation and prediction of political phenomena. They,

---

[1]Robert K. Merton, On Theoretical Sociology (New York: The Free Press, 1967), p. 149.

[2]Abraham Kaplan, The Conduct of Inquiry (San Francisco: Chandler Publishing Co., 1964), pp. 113-115.

[3]Barney G. Glaser and Anselm L. Strauss, The Discovery of Grounded Theory (Chicago: Aldine Publishing Co., 1967), p. 24.

like deductive theorists, attempt to accomplish these goals
by formulating what Hempel, Nagel and others have described
and Goodman has called "lawlike generalizations".[1] If we
examine the requirements for lawlike generalizations we see
that (1) the generalization must be unrestricted as to space
and time, i.e., it must be universal; (2) the generalization
must not be true only because nothing satisfies the conditions
stated; (3) the evidence for the generalization must not
coincide with the range of its application; and (4) the
generalization must be derivable from other laws.[2] Needless
to say, these four conditions are difficult to satisfy within
the overall framework of the empirical generalization approach.
In particular, the fourth condition limits severely the
chances that a student of politics can arrive at lawlike gen-
eralizations via the empirical generalization approach. In
most cases empirical generalizations are not strictly deriv-
able from existing scientific laws (i.e., lawlike generaliza-
tions that have been found to be empirically true.) If they
were, they would be called deductive theories instead of
empirical generalizations. Usually, the connections between
empirical generalizations and the set of existing scientific
laws are rather loose. Inference rather than deduction is
the norm and in most cases the theories which form the bases
of an empirical generalization are themselves nondeductive.

---

[1]See: Abraham Kaplan The Conduct of Inquiry, pp. 91-92.

[2]Ibid., pp. 91-92.

Clearly, our overall approach in this dissertation is closer to the empirical generalization than the deductive theory approach. We do not have an axiomatic or a deductive framework which forms the theoretical base of our hypotheses. Instead, the hypotheses of the dissertation are based upon a conceptual model which we have formulated for explaining implementation success. Thus, rather than looking for a theoretical justification for each specific hypothesis, we need to concern ourselves with the adequacy of the theoretical support for the whole conceptual model.

The conceptual model is similar in many ways to the analytic-conceptual schemes reviewed earlier in Chapter II. However, the specific structure of our scheme is quite different from any of the works reviewed. The similarities between our model and the works of others relate mostly to the inclusion of a specific variable within the explanatory framework used. For example, one of our variables, policy support, is quite similar to Bunker's notion of "policy salience."[1] Similarly, our notion of implementation potential corresponds closely to the capability-related variables proposed by Smith, Bunker, and Bailey and Mosher.[2] Finally, what we have called "environmental forces for implementation" have been used under different names by most of the studies

---

[1]See: Bunker, "Perspectives on Implementation Processes".

[2]See: Smith, "The Policy Implementation Process", Bunker, "Perspectives on Implementation Processes", and Bailey and Mosher, ESEA.

reviewed in Chapter II. The remaining two variables in our model, self-evaluation of capability and level of effort, were not explicitly used in any of the works examined previously.

The model is also similar to and borrows heavily from expectancy theory. Four of our five explanatory variables are almost identical to the basic predictor variables of expectancy theory. Furthermore, both structures attempt to explain the variation in two similar dependent variables: implementation success (performance) and effort (motivation). In addition, additive as well as interactive models have been postulated for the key relationships in both structures.

Expectancy theory provides strong support for our model but the theory itself has been formulated in response to a different problem: performance of individuals within organizational settings, On the other hand, other students of policy implementation who are studying the same problem have used variables similar to ours, but these similarities in variables constitute, at best, weak theoretical support for our model. Thus, we are faced with a situation where there is only weak support for the conceptual scheme from within the field but strong support for it outside. Can this be regarded as "adequate" support? We believe it can.

In studies such as ours where there are no previously tested and directly relevant theories, the investigator is in constant search of schemes which might prove to be useful to his explanation. In our instance, we formulated

several alternative models before arriving at the final model used here. These earlier models represented alternative explanatory structures that seemed to "fit" the problem or the data at hand (or both) but they had one important deficiency: none of them had the kind of theoretical support the present model has.

This is a "plus" for the final model but it does not prove that there is "adequate" support for it. Adequacy is something that needs to be judged in the context of the purposes of the inquiry. Our view is that in exploratory studies such as ours what is needed is a theoretically promising starting point. An empirically-verified theory from another field which also has substantive relevance to the problem at hand provides perhaps the best stepping stone for further theory building in the discipline.

CHAPTER IV

DATA

In this chapter we introduce the specific policy that
is examined here, summarize the methodology and the proce-
dures used in collecting the data, and describe how the pri-
mary variables of the conceptual model have been measured.

The chapter is divided into two broad sections. In
the first section, entitled "The Setting," we first review
the background, the concept, and the organization of Head
Start. We follow this by a discussion of the legislation
which defines the handicapped policy and a review of its
operationalization by the Office of Child Development. In
the final part of the section we describe the purposes and
the major tasks of the project which enabled collection of
the data used here. Also included in the final part of the
section is a presentation and discussion of the major find-
ings reached in the tasks of the project that are relevant
to our work.

We start the second section of the chapter, which is
entitled "The Data," with a summary of the research design
and the methods and procedures used. (The full description
of the procedures can be found in Appendix A.) This is fol-
lowed by a description and discussion of the index

90

construction methodology adopted for measuring four of the six variables in the conceptual model. In the final part of the section we mainly present the empirical results of the index construction process. (Measurement of the secondary variables used in constructing the indices is described in Appendix B.)

<p style="text-align:center">The Setting</p>

<p style="text-align:center">The Program</p>

Background

Head Start was one of the most popular programs devised during the Johnson Administration's war on poverty. As a concept it gained instantaneous acceptance. Its transition from concept to implementation was handled with remarkable swiftness. Implementation of the program started in the summer of 1965, less than a year after the establishment of the first planning committee and six months after the public announcement of the program.[1] The summer 1965 program was allocated a federal budget of close to 100 million dollars and it served 651,000 disadvantaged preschool children.[2] During the following fiscal year (1966) a total of 733,000 children were enrolled and the federal Head Start budget

[1]See: Marshall S. Smith and Joan S. Bissell, "Report Analysis: The Impact of Head Start," Harvard Education Review, Vol. 40, No. 1, Winter 1970, p. 54.

[2]The primary source for these and other statistics presented in this section is: U.S. Department of Health, Education and Welfare, Project Head Start Statistical Fact Sheet-Fiscal Year 1972 (Washington, D.C.,: Office of Child Development, U.S. Department of HEW, 1973).

rose to approximately $200 million. During the eight years

between Summer 1965 and Summer 1973 Head Start served almost

five million children (4,941,500 to be exact) with a federal

budget of $2.75 billion.[1] Taking a recent year as an exam-

ple, during the Fiscal Year 1972 the total Head Start bud-

get was $376.3 million. Of this amount $338.9 million or

90.1% went to the operation of Full Year and Summer Pro-

grams.[2] The rest of the federal funds were used to finance

other Head Start related activities such as the Parent and

Child Centers, Career Development and Technical Assistance

programs, experimental programs, program direction, and eval-

uation. In 1972 there were 269,500 children enrolled in

1,605 full-year programs across the nation. These programs

were staffed with 23,761 professional, 40,151 paraprofes-

sional, and 85,000 volunteer personnel. Sixty per cent of

the full year programs were in urban, the rest in rural

---

[1]Ibid, pp. 2-3. The fiscal Year 1973 figures included in these totals are estimates.

[2]The Head Start Manual defines Summer and Full Year programs as follows: "Summer Head Start programs operate for the most part during school vacation. They are only for children who are eligible for kindergarten or first grade, and will be attending school for the first time in the fall. . . . The minimum length of a summer program is 120 hours. The minimum length of the weekly program is 15 hours. . . . Full Year Head Start programs may operate for periods of up to twelve (12) months for either a part of a day or a full day. The minimum length of a full year program is eight months, of at least 15 hours a week. Children should move directly from full year Head Start into school (kindergarten or first grade). . . . These programs are primarily for children of age three up to the age the child enters the school system, but may include some younger children." See: U.S. Department of Health, Education and Welfare, Head Start Child Development Programs: A Manual of Policies and Instructions (Washington, D.C.: Office of Child Development, U.S. Department of HEW, 1967), p. 4.

areas. The average cost of serving each child in the Full Year 1972 program was $1,118.

Figures for the Summer 1972 program were as follows: There were 434 programs in operation, enrolling 86,400 children. The summer program staff consisted of 5,700 professionals, 9,000 paraprofessionals and 17,738 volunteers. Average cost per child in the Summer 1972 program was $230.

In summary, Head Start was a program that started "big" and remained so throughout its ten-year history. Its immediate acceptance and continued funding are indicators of its popularity. Its scope is also illustrated by the fact that approximately two per cent of the present population of the U.S. has attended Head Start sometime during the last ten years.

The origin of Head Start dates back to the Title II-A of the Economic Opportunity Act of 1964, as amended. Like other programs administered by the Office of Economic Opportunity (OEO) it was designed in such a way as to give local communities broad flexibility in program planning and community involvement. The program called for "maximum feasible participation" of the parents of children. The laissez-faire attitude of OEO in designing and implementing the program led to wide variations among Head Start centers.[1] OEO sponsorship of the program continued until mid-1969. Then, the responsibility to administer Project Head Start was delegated to The Office of Child Development (OCD), a

---

[1]See: Smith and Bissell, Report Analysis, pp. 56-60.

newly-created agency under HEW's Office of the Secretary.

Initially, OCD's management of the program was not

different at all from that of OEO. The same "Manual" was

used to guide local operations, no major policy changes

were made. In short, during the first years of OCD sponsor-

ship local program operations were not any different from

under OEO.

During 1970 and in the following years OCD initiated

several major efforts to strengthen Head Start. The first

one of these was a comprehensive monitoring program launched

in 1970. This was followed by a significant undertaking,

called the "Head Start Improvement and Innovation Program"

(I and I), which aimed at generating improvements in local

program performance. Attempts were made to clarify Head

Start program goals; a set of "performance standards" were

drafted to induce local programs to evaluate and improve

their programs; training and technical assistance activities

were expanded; local programs were offered several "program

options,"[1] several experimental projects were funded in the

tradition of earlier experimental efforts such as Follow

---

[1] There were five basic options offered: (1) the "Standard" Head Start model (the five-day-per-week, center-based classroom format), (2) the "variations in center attendance" model (serving some or all children on a less than five-day-per-week basis), (3) the "double sessions" model (scheduling two classes of children per day), (4) the "home-based" model (serving children in their home, with the parent as the primary factor in the child's development), and (5) the "locally-designed options" model (local programs designing and proposing their own "model"). See: U.S. Department of HEW, OCD Notice N-30-334-1 on Program Options for Project Head Start, (Washington, D.C.: U.S. Department of HEW, Office of Child Development, 1972).

Through, Planned Variations, Health Start, Parent and Child Centers, and Home Start. The initial concept of Head Start, however, remained the same.

Concept

Head Start, as its name implies, is based upon the simple concept that preschool children from disadvantaged families, like their parents, suffer from various ills of poverty and that giving them a "head start" may help break the vicious circle these children would most likely find themselves in while growing up and after becoming adults. This head start was to be given the preschoolers through a comprehensive child development program which aims at bringing about a greater degree of social competence in disadvantaged children. "Social competence," as used here, refers to "the child's everyday effectiveness in dealing with his environment and later responsibilities in school and life."[1] The term, according to Head Start literature, "takes into account the interrelatedness of cognitive and intellectual development, physical and mental health, nutritional need, and other factors that enable a child to function optimally."[2]

Thus, the rationale underlying the Head Start concept is based on a line of reasoning similar to the following:

● All children have certain basic needs (such as physical

---

[1] U.S. Department of HEW, OCD Notice N-30-364-1 on Head Start Program Performance Standards, (Washington, D.C.: U.S. Department of HEW, Office of Child Development, 1973, p. 6).

[2] Ibid., p. 6.

in

pro

As

di

of

co

and mental health, balanced nutrition, cognitive and intellectual growth, and social and emotional health).

● Preschool children whose basic needs are not met are less likely to be socially competent.

● Preschool children from disadvantaged famililties are less likely to have their basic needs met than children from more advantaged families.

● Therefore, unless an intervention is made to meet their basic needs, preschool children from disadvantaged families are less likely to be socially competent.

● Therefore, if we want all our preschoolers to be socially competent, there is a need for an intervention program such as Head Start.

Our simplified version of its concept is also reflected in the program goals for Project Head Start. These goals provide for:

"A. The improvement of the child's health and physical abilities.

B. The encouragement of self-confidence, spontaneity, curiosity, and self discipline which will assist in the development of the child's social and emotional health.

C. The enhancement of the child's mental processes and skills with particular attention to conceptual and verbal skills.

D. The establishment of patterns and expectations of success for the child, which will create a climate of confidence for his present and future learning efforts and overall development.

E. An increase in the ability of the child and his family to relate to each other and to others in a loving and supporting manner.

F. The enhancement of the sense of dignity and self-worth within the child and his family."[1]

As implied by these goals, Head Start's approach is interdisciplinary and comprehensive, allowing for the provision of a broad range of services. In addition, parental and community involvement are viewed as essential elements of

_____

[1]Ibid., p. 7; also see: Head Start Manual, pp. 2-3.

the Head Start approach. Parental involvement is given high
priority because of the belief that the family is the prin-
cipal influence on the child's development. Community
involvement is required for the reason, among others, that
Head Start, alone, does not have the resources required to
provide the broad range of services a disadvantaged child
might need.

Organization

Head Start's organization and the principal guidelines
governing its operation are fully described in the Head
Start Manual and the OCD Transmittal Notices which revise
some of the material in the Manual.[1] Here we provide a
brief summary of some of the pertinent aspects of operating
a local program as background to studying the problem at
hand.

1. Sponsorship.—Local Head Start programs are usually
   sponsored by Community Action Agencies established under
   Section 204 and 205 of the Economic Opportunity Act, school
   systems, or other non-profit agencies. The applicant agency
   is referred to as the "grantee agency" and need not neces-
   sarily operate a program. A grantee agency may subcontract
   all or part of the operation of the program to one or more
   "delegate agencies."

2. Funding.—Applications for Head Start grants are made to
   and reviewed by one of the ten Regional Offices of the
   Office of Child Development. Applicants are usually
   required to finance twenty per cent of the total cost of
   the program from non-federal sources. Part or all of the
   twenty per cent can be met through non-federal in-kind
   contributions, such as use of space and facilities in non-
   federally owned buildings.

---

[1]Another useful background document is the following: U.S. Depart-
ment of Health, Education and Welfare, Project Head Start 1969-70:
A Descriptive Report of Programs and Participants (Washington, D.C.:
Office of Child Development, U.S. Department of HEW, 1972).

3. Types of programs.--Head Start is primarily for preschool children of age 3 and up. Summer Head Start programs operate mostly during the summer months and are only for children eligible for kindergarten or first grade. The minimum length of a summer program is 120 hours and it must operate at least 15 hours per week. Full Year Head Start programs may operate for a period of 8 to 12 months for either part of a day or a full day. The minimum length of the weekly program is also at least 15 hours.

4. Eligibility.--Head Start is primarily for children from poor families. Children from families that are on welfare or that have an annual family income that is less than the federally established poverty line are eligible for Head Start. Up to ten per cent of the children enrolled in each class can be from non-poor families.

5. Parent Participation.--Each Head Start program is required to have an "effective" parent participation program. Parents may be involved in Head Start in one or more of the following ways:

● participation in program-level decision making (in such capacities as a member of the Policy Advisory Group or an elected member of the Parent Committee)

● participation in classroom activities as paid employees, volunteers or observers;

● working with their child in cooperation with Head Start staff;

● participation in educational or other activities for parents which they have helped to develop.

6. Program components.--Head Start has four basic program components: education, social services, parent involvement, and health services. Health services are further broken into four subcomponents: medical, dental, mental health, and nutrition. Each of these basic services is headed up by a program director. Other activities such as recruitment, volunteer activities, and career development are each administered through a coordinator. Other program staff include teachers (usually one per 15 children), teacher aides (usually two per 15 children, one paid and one voluntary), nurses, health aides, social workers, community aides, and cooks.

The Policy

The Legislation

On September 19, 1972, President Nixon signed into law
the Economic Opportunity Amendments of 1972. With this sign-
ing, the new law, P.L. 92-424, opened a new chapter in the
efforts of the federal government to serve the handicapped.
The act contained the following provision:

> "The Secretary of Health, Education and Welfare shall
> establish policies and procedures designed to assure that
> not less than 10 per centum of the total number of enroll-
> ment opportunities in the Nation in the Headstart program
> shall be available for handicapped children (as defined
> in paragraph (1) of section 602 of the Elementary and
> Secondary Education Act of 1965, as amended) and that
> services shall be provided to meet their special needs.
> The Secretary shall implement his responsibilities under
> this paragraph in such a manner as not to exclude from any
> project any child who was participating in the program
> during the fiscal year ending June 30, 1972. Within six
> months after the date of enactment of this Act, and at
> least annually thereafter, the Secretary shall report to
> the Congress on the status of handicapped children in
> Headstart programs, including the number of children
> being served, their handicapping conditions, and the
> services being provided such children."[1]

A definition for the term "handicapped children" in the
legislation is provided in the Elementary and Secondary
Education Act of 1965, as amended:

> "The term "handicapped children" means mentally retarded,
> hard of hearing, deaf, speech impaired, visually handi-
> capped, seriously emotionally disturbed, crippled, or
> other health impaired children who by reason thereof
> require special education and related services."[2]

With this legislation the Congress made eligible close
to 40,000 disadvantaged and handicapped preschool children

---

[1]Economic Opportunity Amendments of 1972, sec. 3(b), 86 Stat. 688 (1972).

[2]Education of the Handicapped Act, sec. 602, 84 Stat. 175 (1970).

of ages three through five for participation in Head Start
programs throughout the nation. Prior to this legislation
Head Start guidelines did not specifically <u>exclude</u> handicapped
children from participation in the program. And, indeed, some
local Head Start programs did enroll several disadvantaged
handicapped preschoolers and served them along with the non-
handicapped. However, with this Congressional mandate the
Office of Child Development was being told that they <u>must</u>
make at least ten per cent of the Head Start enrollment
opportunities available to handicapped children.

LaVor traces the legislative history of P.L. 92-424
to 1969 and the House Education and Labor Committee. In
1969 and in 1970 the Committee considered H. R. 13520 and
later, H. R. 19362, a comprehensive child development act.
The latter bill contained a provision directing the states
to provide programs for the handicapped and to assure that
at least seven per cent of the annual allotment of each state
are used for such programs.[1] This bill was reported out of
the Select Education Subcommittee before the adjournment of
the 91st Congress, but no final action was taken on the bill
during that session.

In 1971, while the House Subcommittee was still working
on a comprehensive child development legislation, the Senate
was considering another child development proposal sponsored
by Senator Mondale. Originally this bill, S. 1512,

---

[1]Martin L. LaVor, "Economic Opportunity Amendments of 1972, Public
Law 92-424," <u>Exceptional Children</u>, Nov. 1972, pp. 249-253.

contained no language on behalf of the handicapped. An amendment to S. 1512 offered by Senator Prouty contained the following language:

> " . . . not less than 10 per cent of appropriate funds under Section 513 (c) be reserved by the Secretary in order to guarantee that handicapped children are included in child care programs."[1]

S. 1512 was incorporated into the Economic Opportunity Act legislation (S. 2007), but in December 1971 the President vetoed the entire bill. It was not until June 29, 1972 that the Senate passed S. 3010, which included an amendment containing the language present in P.L. 92-424.

According to LaVor the rationale for the Senate action was summed up in the Committee report of S. 2007:

> "The history of Headstart clearly shows that severely handicapped children have been systematically excluded from programs and, in fact, children with only moderate handicaps have generally been refused access to such services. These refusals have normally been based on the feeling that the national program is not primarily oriented toward treating handicapping conditions, and expertise is not available at the local level for developing effective programs. Thus, those children with the greatest need for help, and those parents with the heaviest burden to bear, are denied the help to which they should be entitled. . . . We find it discouraging that this program, created to help the children in need, has excluded or failed to help a significant portion of that population."[2]

There are two major implications of this new legislation. First, the Congressional mandate makes about 38,000 handicapped children eligible to receive federal assistance through the Head Start program. Prior to this action,

---

[1] Ibid., p. 250.

[2] Ibid., p. 250.

according to an estimate by LaVor, less than 25,000 of the total population of about one million handicapped children under the age of six were receiving any type of service at all as a result of Federal funds.[1] Thus the new law more than doubles the number of handicapped preschool children eligible to receive federal assistance.

Secondly, with the passage of the new law a new concept was introduced in the Federal approach to the handicapped. For the first time a federal program for the handicapped had no funds earmarked for special services to the handicapped. The ten per cent enrollment mandate was to be met within the budgetary limitations of the regular Head Start grant. Head Start programs were expected to make the necessary arrangements with other community agencies to meet the special needs of handicapped children enrolled in their centers.

The legislation clearly spells out what outcome was to be accomplished by HEW. It does not, however, specify how HEW was going to go about enrolling and serving the handicapped without any additional funds. The mechanics of "how to do it" were left up to the Secretary of HEW. Furthermore,

---

[1]Ibid., p. 253. LaVor's estimates are based on total dollar expenditures for preschool handicapped children in all Federal programs. According to his account, only four federal programs had any expenditures at all for handicapped preschool children: Title I of the Elementary and Secondary Education Act ($465,000 per year for preschool handicapped), Title VI-B of the Elementary and Secondary Education Act ($2,800,000 per year for preschool handicapped), P.L. 89-313, Children in State Supported Schools ($3,264,000 per year for preschool handicapped), and P.L. 90-538, Handicapped Children's Early Education Assistance Act ($7,500,000 per year for preschool handicapped). See: Ibid., p. 251.

the Secretary was mandated to report back to the Congress by March 19, 1973, within six months after the date of enactment of the Act, and at least annually thereafter on the status of handicapped children in Head Start.

Operationalization by OCD

Operationalization of the new policy by OCD can be best understood by reviewing the Transmittal Notice it sent to local programs in Head Start services to handicapped children.[1] This Transmittal Notice outlines the revisions to be made in the Head Start Manual in lieu of the new law and, as such, it summarizes the manner in which the policy has been operationalized by HEW.

The Notice starts with the claim that Head Start has never discriminated against handicapped children:

"Head Start has always had a national policy of enrollment for all eligible children, including handicapped children. . . . As noted in the Head Start Manual of September 1967, 'Head Start encourages the inclusion of mentally or physically handicapped preschool children in an integrated setting with other Head Start children.' . . . "

"More recently, the Office of Child Development has mounted initiatives to give priority to services to handicapped children. These relate, in part, to the overall thrust of the Head Start Improvement and Innovation effort which emphasizes the importance of individualizing services to meet the unique needs and potential of every child participating in the program."[2]

This is followed by a statement that explains OCD's

---

[1] U.S. Department of Health, Education and Welfare, OCD Notice N-30-331-1 on Head Start Services to Handicapped Children (Washington, D.C.: U.S. Dept. of HEW, Office of Child Development, 1973).

[2] Ibid., p. i.

understanding of the reasons for the scarcity of handicapped children in Head Start:

> " . . . there is reason to believe that there are substantial numbers of eligible children who are not registered in Head Start because families are not aware that they are accepted or because local programs do not encourage their participation. Some grantees, contrary to Head Start policy, have even discouraged the participation of handicapped children."[1]

The Notice, according to OCD, "sets forth the policy to implement this legislative mandate and requirements affecting each grantee."[2] Sanctions against not conforming with the contents of the Notice are spelled out as follows: "This issuance constitutes Head Start policy and compliance is required as a condition of further funding."[3]

Salient aspects of the policy set by OCD for implementing the mandate, as indicated in this Notice, can be summed up in the following manner:

- Regional Offices.--The ten per cent mandate is passed down to the ten Regional Offices of OCD. Thus, each Regional Office is required to ensure an average of at least ten per cent enrollment in their geographic area. Furthermore, local programs are encouraged to "negotiate" with their respective Regional Office on their handicapped enrollment targets based on the make up of their target handicapped population, resources and capabilities of the program to serve the handicapped, and desires of the parents to enroll their children in Head Start.

- Comprehensiveness of service delivery.--Local programs are asked to ensure that the full range of services normally available to non-handicapped children are made available to handicapped children as well.

---

[1]Ibid., p. ii.

[2]Ibid., p. ii.

[3]Ibid., p. 1.

- Integration.—Local programs are asked to serve the handicapped in an integrated fashion with the non-handicapped. Exceptions to this policy could be made if it is judged that an integrated setting would not benefit the child.

- Outreach and recruitment.—Head Start grantees are asked to develop, in cooperation with other community groups and agencies, plans and procedures for identifying, attracting and enrolling handicapped children from their respective service areas.

- Screening and diagnosis.—Local programs are required to develop procedures for diagnosing a child's handicap and for assessing his special service needs. This is also to be accomplished in cooperation with other community agencies. These procedures are to cover all handicapping conditions.

- Scheduling.—OCD set a three-phased schedule for implementing the new mandate. In Phase I local programs were asked to give immediate priority to handicapped children in filling their normal enrollment vacancies. In Phase II, Summer 1973 Head Start Programs were asked to make comprehensive plans for serving handicapped children. In Phase III, Full Year Head Start Programs were required to implement the mandate to the fullest extent.

In addition to these, local programs were reminded not to drop non-handicapped children from the program in order to make room for handicapped children. Other provisions of the OCD policy included a warning against mislabeling of children, a clause about the conditions under which a handicapped child could be terminated, and a reminder that some of the program options might be more suitable in serving the handicapped than the standard Head Start model.

Conspicuously missing from the OCD policy are specific definitions of the handicapping conditions included in the mandate and "how to do it" guides for planning and implementing the congressional requirement. The National Office of OCD delegated the implementation responsibility to the regions and the regions, in turn, delegated it to each

grantee. Given the non-specific nature of the policy guide-lines, this meant that the local programs were given full responsibility and flexibility in interpreting and implement-ing the mandate.

In March 1973, as per the reporting requirements of the legislation, OCD presented its first Annual Report on the status of handicapped children in Head Start to the Congress.[1] The Report, which was principally based on responses of 712 out of 1,000 Head Start grantees to a mail-out survey con-ducted in August-September 1972, concluded that "roughly four to five per cent of total Head Start enrollment con-sists of children with special needs. . . . The principal handicapping conditions reported in the survey were speech impaired (31%), seriously emotionally disturbed (14%), and mentally retarded (8%)."[2] The Report did not include any other significant findings than these two. Detailed analy-ses of the survey data were not submitted to the Congress, nor were they made available to the general public.

OCD's First Annual Report to the Congress, having been prepared on the basis of whatever data were available at the time of the signing of the legislation (September 1972) or could be gathered within six months from this date, did not

---

[1] U.S. Department of Health, Education and Welfare, Head Start Services to Handicapped Children - First Annual Report of the U.S. Depart-ment of Health, Education and Welfare to the Congress of the United States on Services Provided to Handicapped Children in Project Head Start (Washington, D.C.: U.S. Dept. of HEW, Office of Child Development, 1973).

[2] Ibid., pp. 7-8.

adequately assess the true status of the handicapped in Head Start. Validity of the reported figures was not assessed. Furthermore, in light of OCD's three-phased schedule for implementing the mandate, full compliance with the Congressional requirement was not expected to occur before the start of the 1973-74 program year. Thus, OCD started looking at the Second Annual Report to the Congress, which was scheduled for submission in or before March 1974, as the first comprehensive report on Head Start's efforts on behalf of the handicapped. Thus, in order to have the data required for preparing the Second Annual Report collected OCD decided to procure the services of an independent contractor and issued a Request for Proposals (RFP) for a major study on Head Start.

The Project

RFP 53-73-HEW-OS, entitled "Evaluation of the Handicapped Effort In the Head Start Program", was actually made up of six separate RFPs combined in one package. Studies on six major aspects of the handicapped effort, which were originally planned as separate projects with separate RFPs, were combined at the last minute into one RFP. Each of the six topics was made a "major task" of this new project and a budget of close to one-half million dollars was allocated for its completion. OCD estimated that completion of the project would require about nine person-years of effort over a sixteen month period starting

in July 1973.

Following an extensive review process, OCD selected the Systems Research Incorporated-Syracuse University proposal from among the responses to the RFP.[1] HEW signed a contract with Systems Research Incorporated (SRI), the prime contractor. SRI, in turn, subcontracted portions of the work to Syracuse University.[2]

---

[1]See: "Technical Proposal - Response to RFP 53-73-HEW-OS - Evaluation of the Handicapped Effort in the Head Start Program" (proposal submitted to the Department of Health, Education and Welfare by Systems Research Incorporated, Lansing, Michigan, May 24, 1973).

[2]Systems Research Incorporated--a Lansing, Michigan based consulting firm with previous experience in management of large scale governmental projects, evaluation of public programs, survey research, and computer applications--was to carry the bulk of the study. The Project Director, Mr. Alan F. Bogatay, was a vice-president of SRI. In addition to Mr. Bogatay, the SRI study team included Mr. Selcuk Ozgediz, a Co-Manager of the study team, Mr. William B. Baucom, chief of the Special Surveys Group, Mr. Allan D. Dale, chief of the Information Systems Development Group, and several other professional staff members.

The Syracuse University study team consisted almost exclusively of the staff of the Division of Special Education and Rehabilitation. Professor Gail Ensher served as a Co-Manager of the study team. Professors Robert Bogdan, Daniel Sage, and Burton Blatt headed up the Assessment and Evaluation Group, the Cost Analysis Group, and the Policy Analysis Group, respectively. Professor Blatt, Director of the Division of Special Education and Rehabilitation and of the Center for Human Policy at Syracuse University, also served as the chairman of the Senior Consultant Group assembled for the project. Serving with Professor Blatt in this group were some of the leading authorities on child development and special education. Among the persons in this advisory group were Dr. Edward Newman (former Commissioner of the Rehabilitation Services Administration at HEW), Professor Seymour Sarason (Yale University), Professor Merle Karnes (University of Illinois), Professor Howard Spicker (Indiana University), Professor Frank Garfunkel (Boston University), Professor Wolf Wolfensburger (Syracuse University), Dr. Julius Richmond, and Dr. John Johnson. In addition, graduate students from Syracuse University handled the bulk of the fieldwork.

Major tasks of the project

The project consisted of six major tasks, as indicated earlier. Two of these tasks, Tasks II and III, involved extensive data gathering and analysis. The other four tasks relied heavily on the data generated in Tasks II and III. The data used in this dissertation were gathered, almost exclusively, within Task III.

The following is a description of the pusposes and the scope of the three major tasks of the project which are relevant to our work.

Task II

The primary purpose of this task was to make sure that the information required by Congress regarding the status of handicapped children in Head Start is collected and ana- lyzed in time to develop the Annual Report due in March 1974. In addition, data collected in Task II provided a basis of selection of Head Start programs visited for in-depth assess- ment during Task III. Task II also helped establish several key parameters of program costs analyzed in Task IV.

Task II was completed in essentially seven stages. In the first stage, immediately after the contract award, a mail- out survey was conducted to assess the status of the handi- capped effort in the Summer, 1973 programs. During the second stage, definitions for the nine handicapping conditions were developed and a "master" list of Head Start grantees and delegate agencies was developed. In the third stage,

the survey instrument and accompanying instructions were developed, pretested, and cleared with the Office of Management and Budget. During the fourth stage the questionnaires were mailed (October 1974). The fifth stage involved coding, keypunching, verifying, programming, processing, and analysis of the returns. In the sixth stage two special telephone surveys were conducted. First, a random sample of about 100 of the responding grantees and delegate agencies were telephone interviewed to assess the test-retest reliability of the survey instrument. Second, a telephone interview was conducted with each non-responding grantee to find out if the non-respondents differed from the responding programs in terms of their handicapped effort. Finally, in the seventh stage, a draft and a final report was prepared and submitted to OCD. The draft report was used by OCD in preparing their Annual Report to Congress.[1]

Task III

The primary purpose of Task III was to collect in-depth, on-site information on the implementation of the Congressional mandate so that an assessment could be made of the ways of improving the implementation. Task III involved site visits to three types of programs: a set of regular Head Start programs, a set of special "experimental

---

[1]See: "The Status of Handicapped Children in Head Start - Final Report Based on Surveys of Summer 1973 and Full Year 1973-74 Head Start Programs" (report submitted to the Office of Child Development by Systems Research Incorporated, October 1974).

projects," and a set of "exemplary projects." The experi-
mental projects, they were fourteen in all, were regular
Head Start programs which had received special funding from
either OCD or the Bureau of the Education for the Handicapped
for experimenting with and demonstrating new approaches to
providing comprehensive services to pre-school handicapped
children in a program setting with non-handicapped children.
The exemplary projects consisted, primarily, of ten non-
Head Start preschool enrichment programs which could offer
new models for serving the handicapped within Head Start.

Task III site visits were conducted in two rounds.
In the first round, which took place during November-Decem-
ber 1974, visits were made to sixteen regular Head Start
programs and eleven experimental projects. During the
second round, which was conducted in April-May 1974, thirty-
six regular Head Start programs, three experimental projects,
and ten exemplary projects were visited.

Approaches to the first and the second round of visits
to regular Head Start programs differed substantially. In
the first round the visits were exploratory and open-ended.
Participant observation was the principal technique used and
data collection centered around eleven major "areas of inqui-
ry." An open-ended interview guide was used to guide the
field observations.

In contrast with the first round visits, the second
round visits were more structured and they concentrated on
better defined issues. A total of thirty-six regular Head

Start programs were visited in this round. Since the data used in this dissertation comes primarily from these thirty-six programs, the procedures used in selecting them are discussed in detail in Appendix A.

Task V

This task involved an extensive review of the need for and the supply of services for preschool handicapped children in the country. The overall aim of the task was to provide OCD with data that can be used for future planning and policymaking. A literature review was conducted on serving the preschool handicapped and estimates were made of the numbers of preschool handicapped children by handicapping condition. Data collected from the exemplary programs were studied to identify some of the best current approaches to serving the preschool handicapped. Also, an extensive telephone and mail survey of agencies serving the preschool handicapped was conducted in each of the fifty states. Finally, the data generated in the other tasks of the study were examined by the Senior Consultant Group and their impressions and recommendations were forwarded to OCD.[1]

This completes our review of the major tasks of the project that are relevant to our work. We now turn our attention to some of the major findings from the study.

_____

[1]See: "Statement on Policy Recommendations to the Office of Child Development" (statement prepared and signed by the members of the Senior Consultant Group, September 1974).

Major findings

We will not go over the detailed findings from all the
tasks of the project. Instead we will summarize the most
important, and the most contraversial, findings from the work
done by SRI and Syracuse University. Some of these findings
have led to debates within as well as outside the project
team. Let us first review the national survey conducted by
SRI.

The survey of Full Year 1973-74 programs yielded a
response rate of about 79%, with at least one response from
83% of all grantees.[1] Since a separate telephone survey was
conducted with the non-respondents, if one combines the
responses to key items obtained through this survey with the
responses to the mail survey, the overall response rate for
key items of data reaches approximately 90%. A third survey,
conducted in the form of telephone interviews with 100 of
the responding programs for ascertaining the reliability of
the responses to the mail survey, has yielded a correlation
coefficient of .97 between the total number of diagnosed
handicapped children reported by the sampled programs to the
Full Year survey and the telephone survey.

The following are the principal findings from the
national survey:

---

[1]The unit of analysis for the SRI survey was a delegate agency.
The 79% figure represents the response rate from all delegate agencies.
The 83% figure refers to the grantees. This means that not all of the
delegate agencies under a grantee may have responded to the survey.

- There were about 29,000 handicapped children enrolled in Full Year programs as of November-December 1973. "These 29,000 were children whose handicapping conditions were reported to have been diagnosed by qualified professionals. In addition, the handicapping conditions of 5,000 children had only been partially diagnosed as of November-December 1973; and 4,000 other children were believed by Head Start staff to be handicapped, even though their handicaps had not yet been confirmed."[1] Thus, percentagewise the handicapped enrollment in Head Start was at least 10.1%; this figure could be as high as 13.2% if all the children partially diagnosed or suspected to be handicapped are eventually confirmed as "handicapped."

- About one of every five handicapped children were reported to have multiple handicaps. The primary handicapping conditions of children reported as handicapped were as follows:[2]

Speech impaired	35.0%
Health or developmentally impaired	20.0%
Seriously emotionally disturbed	12.2%
Physically handicapped	9.5%
Hearing impaired	7.9%
Mentally retarded	7.4%
Visually impaired	6.6%
Deaf	1.0%
Blind	0.5%

- 46.6% of the handicapped children were reported to require "a fair amount" or "practically constant" special assistance. Since the amount of special assistance required by a child was used as a measure of the severity of the child's disability, roughly one-half of the children reported as handicapped were only "mildly" handicapped. The other half were judged to be "moderately" or "severely" handicapped.[3]

- Handicapped children, in general, were served in physically integrated settings. "In 57% of all Head Start programs at least one handicapped child was present in 80 to 100% of the classrooms of these programs."[4]

---

[1] See: "The Status of Handicapped Children in Head Start - Final Report Based on Surveys of Summer 1973 and Full Year 1973-74 Head Start Programs", p. 2.

[2] Ibid., p. 5.

[3] Ibid., p. 6.

[4] Ibid., p. 7.

- Roughly 60% of the programs serving the handicapped provided or arranged special pre-service training, the same percentage of the programs provided in-service training to improve the abilities of their staff to work with handicapped children. Special equipment and materials were provided for handicapped children by one out of six programs serving the handicapped. About 35% of the handicapped children in Head Start were referred to Head Start as handicapped (by their parents or by other agencies); and special diagnostic services were provided by Head Start for an estimated 75% of the handicapped children enrolled.[1]

- Finally, "the average handicapped child in Head Start was receiving about two of the seven special services listed below."[2]

  - provision of special counsel to parents of handicapped children related to their child's handicap

  - provision of planned special experiences to handicapped children to increase "adjustment skills" related to their handicap

  - special, individualized counseling

  - speech or physical therapy

  - medication and/or drug therapy

  - prosthetic devices

  - other special services

Let us now turn to a review of the principal findings from Task III of the study.

Findings from the first round of visits to sixteen regular Head Start programs were based primarily on the impressions of the numbers of the research teams who have conducted the visits. "Participant observation" being the main technique used in these visits, results from each visit was reported through a narrative "site visit reports," which

---

[1]Ibid., p. 8.

[2]Ibid., p. 9. Special services provided are listed in order of decreasing frequency, i.e., the first listed service is reported to have been provided to more number of children than any of the others.

sometimes reached seventy-five typewritten pages in length. Although these insightful reports were extremely useful in getting a feel for the manner in which the mandate was being implemented, they were too heterogeneous and non-uniform to permit adoption of any systematic analysis scheme. Thus, the main conclusions from the first round visits were arrived at through a modified delphi exercise. Participating in these lengthy analysis sessions were close to all of the field work staff. The prime objective of these sessions was to reach a consensus on observations about the implementation of the mandate. The following were some of the key conclusions reached in the analysis sessions.

- Head Start staff were confused by and unclear of the term "handicapped." As a result, the mandate had the deleterious effect of mislabeling some children as "handicapped."

- All of the programs visited reported that they had always served the handicapped and that they had not excluded any handicapped child from participation in the program, with the exception of a few severely disabled children.

- The handicapped effort at the local level proceeded more as an "evolving" than a pre-planned process.

- Overall, the number of handicapped children enrolled in the Full Year 1973-74 program was only slightly higher than the number of such children enrolled the previous program year.

- The majority of the children identified and reported as handicapped were only mildly handicapped; severely handicapped children constituted only a small percentage of all children identified as handicapped.

- Programs, in general, were making more special efforts to serve the handicapped this year than heretofore. Significantly greater efforts were being made this year to diagnose handicapping conditions professionally than in earlier years. In addition, the mandate resulted in more detailed individualized assessments of the developmental needs of handicapped as well as non-handicapped children.

● Parents of both handicapped and typical children expressed
very positive attitudes towards the efforts of Head Start
on behalf of the handicapped. The handicapped effort also
resulted in increased relationships with community agencies.

● Head Start staff expressed a strong need and desire for
additional training which would help them better serve the
handicapped. Receptivity of the staff to severely handi-
capped children seemed to increase as their contacts with
such children increased.

● Staff of most of the programs visited indicated that they
disagreed not so much with the intent of the mandate but
with the manner in which the National and the Regional
Offices of OCD "passed it down" to the local level, with
very little guidance, support, and direction.[1]

We need to clarify that all of the programs visited

in the first round had previous experiences in serving the

handicapped. If at all, as a group they are more represen-

tative of programs with relatively larger numbers of handi-

capped children. Quoting from the Task III Interim Report,

" . . . if the findings from the first round of visits are

to be generalized, they can be generalized not to total

Head Start, but to top 30 per cent of the programs as far

as the handicapped effort is concerned."[2]

Findings from the second round visits, which were

based on a larger sample of programs and a more rigorous

methodology, did not differ significantly from the find-

ings from the first round visits. The most prevalent issue

was mislabeling of children. Quoting from the Final Report,

" . . . we found that almost without exception the programs

---

[1]Sources for the conclusions listed here are the Interim Report
and the Final Report on Task III.

[2]See: "Interim Report . . . ", p. 22.

we visited were applying the term 'handicapped' to some children with very minor difficulties, who required no special classroom assistance or services, who have always been enrolled in Head Start without such classifications. . . . In our view, this has been the most serious and deleterious effect of the new legislation. . . . "[1] To the question "How successful has Head Start been in providing new services for handicapped children this year?" the Task III Final Report provides the following answer: "Head Start services for children with special needs have basically remained the same this year and in order to really fulfill the intent of serving more serious disabled children, the legislation needs to be further clarified and new approaches with greater resources developed."[2]

Conclusions of the Syracuse University team from both the first and the second round visits are summarized as follows:

> "Severely retarded children comprised a very small percentage of the total enrollment of Head Start centers visited, said population significantly less than the 10 per cent Congressional demand; there was great variation among programs vis-a-vis attitudes towards the handicapped—especially the severely handicapped—and program opportunities for the handicapped. By and large, most mildly and moderately handicapped children were physically and psychologically integrated in Head Start programs, with such integration usually assured upon admission; exclusion or exemption was the more serious problem than the integration of those admitted. The mandate appeared to have positive effects in increasing a coordinated involvement with families and other community agencies. Lastly, Head Start staff continued to feel very

---

[1]See: "Final Report . . . ", p. 202

[2]Ibid., p. 230 (emphasis deleted).

strong general needs for both in-service training and improved and increased technical assistance and consultation." [1]

The conflict between the findings reported in the Task II and the Task III final reports, prepared by the SRI and the Syracuse University teams, respectively, is quite apparent. While the Task II report claims that, based on the figures reported by the local programs in November-December 1973, the ten per cent Congressional mandate has been surpassed, the Task III report makes an opposite claim, based on data collected during site visits to a total of fifty-two programs, that the ten per cent mandate has not been met in terms of the severely handicapped children enrolled and that imposition of the mandate has led to unnecessary and improper labeling of some children with very minor difficulties as "handicapped."

The issue, clearly, is a definitional matter. In fact, there would even be no conflict between the two reports if one assumes that the intent of the mandate covers only the moderately and the severely handicapped. If one were to measure "moderately" and "severely" in terms of numbers of children requiring "a fair amount of" or "almost constant" special assistance, the Task II survey yields a handicapped enrollment figure between 4.7 and 6.1 per cent. And this

---

[1] Ibid., pp. 252-252.

is clearly well below the mandated 10 per cent figure.[1]

There is considerable evidence that the intent of the mandate lies in the direction of the severely handicapped. Congressman Albert Quie (R-Minnesota) remarked, in a floor speech on the day the House voted on the S.3010 Conference Report, as follows: " . . . It is my hope that as a result of this amendment that Headstart will now provide services for children who are totally blind, totally deaf and severely physically or mentally handicapped. . . . It has been said for years that a child who is poor is handicapped. I would add to that that a child who is poor and also severely handicapped is doubly handicapped . . . "[2]

The OCD policy on the handicapped effort also excludes mildly disabled children from the scope of the mandate:

> "While children with milder handicapping conditions
> (e.g. children with visual problems correctable with eye-
> glasses) will continue to be identified and receive appro-
> priate Head Start services, they fall outside the scope
> of this issuance. The intent is rather to insure that
> Head Start serves more fully children who have severe
> vision and hearing impairment, who are severely physically
> and mentally handicapped and who otherwise meet the legis-
> lative definition of handicapped children in terms of

---

[1]The 4.7% figure takes into account only children diagnosed as handicapped as of November-December 1973 and reported as requiring "a fair amount of" or "almost constant" special assistance. The 6.1% figure includes these children plus children whose diagnosis had not been completed at the time of the survey and children who were judged to be handicapped by Head Start staff prior to any diagnosis. In computing the 6.1% figure we assumed that the same percentage (i.e., 46.6%) of these additional children would require "a fair amount of" or "almost constant" special assistance once their diagnoses were completed.

[2]See: LaVor, "Economic Opportunity Amendments of 1972, Public Law 92-424", p. 253 (emphasis added).

their need for special services."[1]

Finally, the individual definitions of the nine handi-
capping conditions used in the surveys of Summer and Full
Year programs, which were developed jointly by the SRI and
the Syracuse University teams, clearly refer only to chil-
dren with moderate or severe handicaps.[2]

In light of these interpretations of the policy, there-
fore, one cannot readily accept the claim that the ten per
cent mandate has been met. By the same token, it cannot be
conclusively asserted that Head Start programs, nationally
have failed in implementing the mandate. It can be stated
with no or few reservations, however, that, in general, local
programs with severely handicapped children were more success-
ful in implementing the Congressional mandate than programs
with the same number of mildly or moderately disabled
children. This measure of policy success, in our opinion,
is more in line with the intent of the policy than one which
does not distinguish between levels of severity. And the
primary success measure we have used in the dissertation
assigns a greater weight to each severely handicapped child
enrolled.

---

[1]See: OCD Notice N. 30-331-1 on Head Start Services to Handicapped Children, p. 3.

[2]See: "Appendix A: Survey Questionnaire and Definitions - Full Year 1973-74 Programs" in "The Status of Handicapped Children in Head Start - Final Report Based on Surveys of Summer 1973 and Full Year 1973-74 Head Start Programs".

The Data

In this section we concentrate on methodological and data-related issues. Our main concern is the development of empirical indicators for the six primary variables of the conceptual model. This measurement process involves the construction of indices for each primary variable. The elements or secondary variables entering each index are generated from the data collected during the second round visits to the thirty-six Head Start programs. Our choice of index construction as the principal method, as we will further discuss below, was dictated by the nature of the data that were collected in the project.

We have divided this section into three parts. We first present an overview of the methodology and the procedures used in the second round visits. This overview consists mainly of a summary of the detailed description of the research design and the methodology presented in Appendix A. In the second part of the section we describe and discuss the specific procedure used in constructing the indices. The final part of the section is devoted mainly to presentation of the empirical results from the index construction process.

An Overview of Methodology and Procedures

Research design

The research design used for generating the data is primarily a cross-sectional survey design. The data were

collected at approximately one point in time (April-May 1974)
from a sample of thirty-six regular Head Start programs.
Longitudinal data were generated by asking the respondents
to report information about earlier points in time and by
having the respondents describe certain processes over time.
Due to the lack of a control group and pre and post measure-
ments, the design suffers from the phenomena that typically
limit the internal validity of all non-experimental designs.

Questionnaire design

The questionnaire was designed in order to find answers
to the following three questions: (1) How well are Head
Start programs presently serving the handicapped, in partic-
ular the severely handicapped? (2) Why are some Head Start
programs presently not serving the severely impaired?
(3) What are the factors which account for the differences
between programs in enrollment of severely handicapped chil-
dren? The instrument was developed over a two-month period
after a careful analysis of the exploratory data collected
during the first round visits to sixteen programs. Data
from only the "Program-Level Information" part of the ques-
tionnaire are used in the dissertation; child-specific and
classroom observation data did not serve our purposes. The
finalized version of the instrument and its accompanying
documents were fielded only after obtaining formal approval
from HEW and the Office of Management and Budget.

Selection of programs

A total of thirty-seven regular Head Start programs were originally selected for the second round visits. One of these programs withdrew from participation in the study and we dropped one program due to problems caused by missing data. Empirical analyses found in the dissertation are, therefore, based on data from a sample of thirty-five programs.

The sampling frame consisted of the list of 1,353 Full-Year Head Start programs which responded to the mail survey conducted during November-December 1973. The original sample size of thirty-seven was determined on the basis of resource constraints and contractual obligations. The final sample was selected through a complicated process which involved stratification of programs based on their size and severely handicapped enrollment, reduction of the size of the sample frame through random sampling, further reduction of the reduced sample frame through application of a set of screening criteria, random selection of programs with no reported handicapped children, and purposive selection of the remaining programs. Although randomization has entered the selection process at several points, the final sample cannot be considered truly random. However, there is no reason to believe that the sample is not representative of the overall population of Head Start programs.

Pre-fieldwork and fieldwork
activities

Almost all of the interviewers/observers who conducted the fieldwork during the second round had previous fieldwork experiences, mainly as a result of their participation in the first round visits. In addition, they were subjected to a two-day formal training session and they participated in a three-day classroom observation program conducted in carefully selected preschool and special education settings. All of the ten persons who conducted the fieldwork were faculty members and graduate students from Syracuse University.

Visits to the selected programs lasted one to three days and, except for five large programs, were conducted on a one interviewer per program basis. The principal respondent to the "Program-Level Data" section of the questionnaire was the Head Start director at the sampled program. No major problems were encountered during the fieldwork.

Post-fieldwork activities

These activities included preparation of the codebook, which included codes for 1,035 variables, editing of the collected data, reinterviews with the respondents, coding of the quantitative and the anectodal data, data processing, and analysis. Analysis of the data by the Syracuse University team was mainly descriptive and did not involve any model building or multivariate analysis.

Our conceptualization of policy implementation and analysis of the second round data were done independently of

the work of the project team from Syracuse University. The
only thing common to our work and that of the Syracuse Uni-
versity team is the data base generated during the second
round visits.

Index Construction Methodology

The second round visits to the sampled programs were
conducted in order to assess the status of the implementation
of the policy at the field level and to investigate the fac-
tors which may account for the differences in the success
levels achieved by programs in implementing the congressional
mandate. The questionnaire was not designed and the data
were not collected with a specific model of implementation
in mind. Instead, the data collection effort centered
around key issues and concerns which surfaced as hypotheses
at the conclusion of the exploratory research conducted dur-
ing the initial stages of the project. For this reason,
the correspondence between the data at hand and the vari-
ables of our conceptual model is not as close as we would
have liked to have.

It is possible to measure two of our variables, level
of implementation success and self-evaluation of capability,
directly, i.e., without extensive data manipulation. This
is because the available data allow us to develop simple
indicators for measuring these two variables. For the other
four variables, however, it is possible to construct a vari-
ety of indicators which relate to one or more aspects of

of each variable. Unfortunately, none of the indicators that relate to a given variable is "superior" to the others in terms of validity or reliability. Thus, in developing measures for these four variables we are faced with a difficult choice between (1) selecting only one of the relevant indicators and using that as our measure for that variable, and (2) developing a new measure by combining all the indicators relevant to that variable. We have chosen the second alternative because it allows us to keep the "measuring power" of all indicators relevant to a variable. This alternative involves construction of an index for each of the following four variables: implementation potential, level of effort, environmental forces for implementation, and policy support. All of the secondary variables which were considered in developing the indices for these variables are defined, described, and discussed in Appendix B.

Index construction is a data reduction technique. It is similar to and at least as widely used as scaling, although there is much less methodological literature on index construction than scaling. Both techniques yield ordinal measures and there are slight differences between them. Babbie summarizes these differences as follows:

> "An index is constructed through the simple cumulation of scores assigned to specific responses to the individual items comprising the scale. A scale is constructed through the assignment of scores to *response patterns* among the several items comprising the scale. A scale differs from an index by taking advantage of any *intensity structure* that may exist among the individual items."[1]

---

[1] Earl R. Babbie, <u>Survey Research Methods</u> (Belmont, California: Wadsworth Publishing Company, Inc., 1973), p. 254.

Construction of an index typically involves completion of the following steps:

- selection of the items considered for inclusion in the index,
- examination of the relationships among the items selected,
- determination of the scoring procedure,
- construction of the initial composite index,
- internal validation of the composite index through item analysis, and
- external validation of the index.

These steps correspond roughly to the procedure we used in constructing the indices for EFRT, IPOT, ENVR, and PSUP. However, because of the unavailability of a criterion measure we could use for external validation of each index, we decided to construct each index twice, using different indicators for each item considered in the index. Our rationale for doing this is that if the same items end up in the two alternative indices and if the two final indices correlate strongly with one another, our confidence in both indices would be greatly increased. If, however, the indices developed from two different sets of indicators for the same initial items seem to differ from one another significantly, we would have less confidence in the index and would have to explain the differences between the two indices before using either index in the statistical analyses.

There is a second reason for our choice to develop two alternative measures for each item considered in the index. As can be observed in Appendix B, most of the indicators we

have for the items (secondary variables) initially included
in the indices for EFRT, IPOT, ENVR, and PSUP are measured
at the ordinal level.[1] Theoretically, we cannot add two
or more ordinally measured items and obtain a composite
score. However, if we can transform these ordinal measures
into an interval measure, then we can conduct the arithmetic
operations necessary for the construction of the composite
index. And if the index based on interval-level measures is
found to be very strongly correlated with the one based on
ordinal measures, we would have reason to argue that the two
indices can be used interchangeably.

How can we transform an ordinal scale into an interval
measure? The only alternative that seemed reasonable is to
assign to each ordinal code of an item the cumulative per-
centage which corresponds to that code in the univariate fre-
quency distribution of that item. Cumulative percentile
scores can be interpreted as cumulative probabilities (which
are measured at the ratio level). With this transformation,
each cumulative percentage shows the probability that a pro-
gram randomly selected from the sample at hand would achieve
a given ordinal score or lower. Since the original ordinal
codes are all previously arranged such that the lowest code
corresponds to the least desirable outcome (from the view-
point of the policy's success), the cumulative percentage

---

[1]We consider the "yes-no" type dichotomies we have as ordinal
measures because the two categories of all of our dichotomies are
directional and ordered.

distribution preserves the original order among the programs.
The range of the cumulative percentage distribution is
between zero and one. Although the minimum cumulative per-
centage would never equal zero, it could be very close to it.
However, the distribution always attains the maximum value
of one.

The transformation described above yields, for each
item considered for inclusion in the index, a new distribu-
tion which has a measurement level higher than ordinal.
These new sets of values can now be used in constructing what
can be called an "alternative index" for each of the four
primary variables. This index is alternative to the "orig-
inal index" which is based on a simple linear transformation
of the original codes. The linear transformation is neces-
sitated because the range of the ordinal values assigned to
each item differ from one another. In the case of dichoto-
mies, for example, the originally assigned codes are zero
and one. For a trichotomy these codes are zero, one, and
two; and in one instance the values of an item assumes
integer values between zero and eight. Simple addition of
these codes in constructing the index amounts to assigning
a weight to an item on the basis of the number of values it
assumes. In order to correct for this imbalance we linearly
transformed the original codes of each item into the [0,1]
interval. This transformation assigned a value of zero to
the minimum and a value of one to the maximum original code.
Codes inbetween the minimum and the maximum were transformed

into the [0,1] interval on a linear basis. This transformation maintains the original order among the programs and assures equality of the intervals between successive values of the item.

In summary, we constructed two indices, an original and an alternative index, for each of our four primary variables which are measured through indices. The original index was based on a simple linear transformation of the ordinal code for each item into the [0,1] interval. The alternative index was based on a transformation which assigned each original ordinal code the cumulative percentage corresponding to that code in the univariate frequency distribution of the item. The two indices were constructed independently of one another and the same procedures were used in arriving at the sets of items which are included in the two final indices.

The following is an outline of the specific procedures used in developing the two indices:

- The items to be included in the *item pool* for each variable were selected on the basis of their face validity after a careful review of the data. A conscious attempt was made to include as many valid items as possible in the item pool. The full list of items (secondary variables) initially included in the item pools appears in Appendix B.

- A *correlation matrix* was constructed in order to examine the bivariate Pearson correlations among the variables and their associated significance levels.

- The *signs of the bivariate correlation coefficients* were

studied. Since the codes for the items were arranged such that in all cases the lowest value of the item corresponded to the least desirable outcome, items which had consistently negative correlations with the others were dropped from the the item pool.

- An *initial index* was constructed by adding the scores of items remaining in the item pool. This represents assignment of equal weight to each item. We had no a priori reasons for assigning differential weights to the items.

- An *item analysis* was conducted by studying the bivariate correlations between each item and the initial index minus that item. If this correlation coefficient was not significant at the 0.10 level the bivariate coefficients between that item and the others which entered the initial index were examined. If the item did not correlate with any of the others (again at the 0.10 level of significance) the item was dropped from the item pool.

- A *final index* was obtained by revising the initial index based on the results of successive item analyses. In all instances, items were dropped one at a time to observe the effects of dropping a variable on the correlations between a remaining item and the new index minus that item.

We compared the alternative index and the original index, which were independently constructed using the above procedure, in terms of the final set of items remaining and the bivariate correlation coefficient between the two final indices. In all four cases the two final indices were almost identical. They yielded the same final set of items and the

Pearson correlations between the two indices were all signif-
icant at the 0.001 level. We review these results and the
measurement of the other two variables (i.e., LIS and CAPB)
in the next part of this section.

Measurement of Variables

We describe how we have measured the six primary vari-
ables of the model in the order these variables were intro-
duced and defined in Chapter III.

Level of implementation success

The overall implementation objective of the policy is
to enroll and serve in Head Start sufficient numbers of hand-
icapped children to account for at least ten per cent of the
total enrollment. This is the goal specified in the legisla-
tion and the objective passed down to local programs by OCD.
Thus, the simplest success measure we could use is a dichot-
omy based on breakdown of the programs into two groups on
the basis of their percentage handicapped enrollment. How-
ever, this measure is unsatisfactory because the use of the
phrase "at least ten per cent" in the legislation implies
that a program with, say, twenty per cent handicapped would
be considered more successful than one with ten per cent
handicapped. Adding to this the fact that a great number of
programs did not in fact reach the ten per cent figure, it
is reasonable to measure success in terms of the actual per-
centage of the handicapped enrollment.

Enrollment of handicapped children reflects one aspect

of policy success. Another success related issue is the question of how well each program is serving the population of handicapped it has enrolled. This is a very difficult question to address empirically. Eventually the question boils down to how well the program is serving each handicapped child enrolled. This could be perhaps best measured by estimating the extent to which Head Start is meeting the special needs of each handicapped child. Although we have some child-specific data on this issue, the sample size per program is too small to lead to any meaningful empirical analyses.[1] Thus, the success measure used in our analyses can only be based on the enrollment of handicapped children.

How can one determine if a child is handicapped? This is a highly complex and technical issue. The legislation specifies nine handicapping conditions[2] and in the course of the project definitions were developed for each of these impairments.[3] In addition, a distinction was made between levels. Children who required little or some special assistance within Head Start, i.e., assistance that would not be normally provided to non-handicapped children, were referred

---

[1]Seventy-four child-specific case studies were developed in the second round. These children were enrolled in twenty-five of the thirty-five programs in our sample. This yields only a sample of three children per program.

[2]These are: blindness, visual impairment, deafness, hearing impairment, physical impairment, speech impairment, health or developmental impairment, serious emotional disturbance, and mental retardation.

[3]These definitions were sent to local programs as a part of the Task II questionnaire. See: "Appendix A: Survey Questionnaire and Definitions-Full Year 1973-74 Programs" in "The Status of Handicapped Children in Head Start-Final Report Based on Surveys of Summer 1973 and Full Year 1973-74 Head Start Programs."

to as "minimally handicapped;" those that required a fair
amount of assistance were called "moderately handicapped;"
and children who required practically constant attention
and/or assistance were classified as "severely handicapped."
These measures of severity, however, were only suggestive,
the specific handicapping condition and its severity level
was to be determined by a qualified professional. In sum,
although some attempt was made to provide guidance and
assistance to local programs regarding the definitional
question, the final decision was left to the judgment of
qualified medical personnel affiliated with the program.

Measurement of implementation success in terms of per-
centage of handicapped, where the term "handicapped" covers
all handicapping conditions and severity levels, has a major
disadvantage. It places mildly or moderately handicapped
children on a par with the severely impaired. As we pointed
out in the first section of this chapter, the intent of the
policy clearly lies in the direction of the severely handi-
capped. Given two identical programs each with only one
handicapped child with the same impairment, the program
serving a severely handicapped child would be considered as
more successful than the one serving a mildly or a moderately
handicapped. This is not only because of the intent of the
mandate but also because of the fact that, given the insuf-
ficiency of educational and other services for the preschool
handicapped nationwide, the severely handicapped would in
all likelihood benefit more from participation in Head Start

than the mildly or the moderately handicapped. Adding to these the consideration that in many of the programs children classified as mildly handicapped were in fact either non-handicapped or were clearly outside the scope of the policy, we reach the conclusion that the success measure should be based more heavily on the percentage severely handicapped.

One alternative is to use as our success measure only the severely handicapped percentage. This has two disadvantages. First, it does injustice to programs serving "legitimate" moderately handicapped children, which, for one reason or another, have enrolled only few or no severely handicapped children. Secondly, as illustrated in Table 1, the distribution of the severely handicapped percentage is extremely skewed. If we were to use this measure, our dependent variable would have very little variance. In particular, we would not be able to meaningfully study the twelve programs with no severely handicapped children.

The specific measure we are using here is based on differential weighing of the percentage severely handicapped and the percentage mildly/moderately handicapped. We assigned a weight of two to the percentage severely handicapped and a weight of one to the percentage mildly/moderately handicapped and added the weighed percentages to obtain the success scores illustrated in Table 2.

Differential weighing of the two percentages reflects both the assignment of a higher level of success to a program with severely handicapped children and the discounting of

TABLE 1

PERCENTAGE OF SEVERELY AND MILDLY/MODERATELY
HANDICAPPED CHILDREN

Percentage of Severely or Mildly/Moderately Handicapped	Programs with the Indicated Percentage of Severely Handicapped		Programs with the Indicated Percentage of Mildly/Moderately Handicapped	
	Number	Percentage	Number	Percentage
0	12	34.3	4	11.4
0.1 - 5.0	11	31.4	7	20.0
5.1 - 10.0	6	17.1	8	22.9
10.1 - 15.0	5	14.3	8	22.9
15.1 - 20.0	1	2.9	4	11.4
20.0	0	0	4	11.4
Total	35	100.0	35	100.0

the reported numbers of mildly/moderately handicapped. The weight of two for severely handicapped is not totally arbitrary. It is based, in part, on the assumption that roughly half of the children reported as mildly/moderately handicapped can be considered as "normal" children with only minor disabilities. This is supported by the findings from the Task II survey as well as the observations of the interviewers who conducted the fieldwork.

According to this procedure, then, a program with three per cent severely handicapped and seven per cent

TABLE 2

IMPLEMENTATION SUCCESS SCORES

Success Score	Number of Programs	Percentage of Programs
0	4	11.4
0.1 - 5.0	2	5.7
5.1 - 10.0	3	8.6
10.1 - 15.0	5	14.3
15.1 - 20.0	9	25.7
20.1 - 25.0	2	5.7
25.1 - 30.0	1	2.9
30.1 - 35.0	2	5.7
35.1 - 40.0	5	14.3
40.1 or over	2	5.7
Total	35	100.0

mildly/moderately handicapped would be assigned a success score of thirteen. The potential range of the success score is between zero and 200. The actual range is between zero and 51.6. The mean score is 19.1 and the median is 15.4. The distribution is somewhat skewed to the left, however, the skewness coefficient for this distribution is considerably less than those for the distributions of percentage severely handicapped and the percentage mildly/moderately handicapped (0.59 vs. 1.26 and 1.18, respectively). So is the coefficient of variation. It is 72.5% for the distribution of success

scores, as compared to 131.4% and 79.3% for the distributions of percentage severely handicapped and percentage mildly/moderately handicapped, respectively.

Level of effort

We do not have any direct measures of the level of "human effort" expended by each sampled program during the implementation of the mandate. Instead we have eight indicators which roughly measure the amount of task-specific effort that has gone into eight different aspects of the implementation. These eight items, which are defined and discussed in Appendix B, relate to the following aspects of the implementation:

ORGS : effort to organize for the implementation of the policy;

PLAN : planning effort as indicated by the comprehensiveness of the plan prepared for implementing the policy;

COOR : effort to coordinate the implementation of the policy with other agencies within OCD;

MOBR : effort to mobilize the resources of the community;

FUND : effort to secure additional funds for the implementation;

SIZE : effort to increase the size of the implementation staff;

TATA : policy-related training and technical assistance effort;

MFAC : effort to modify the physical facilities in order to make them more appropriate for serving handicapped children.

These eight areas do not fully exhaust all possible

policy-specific activities which were undertaken by the
implementing agencies. Activities for which we do not have
indicators of level of effort include, in particular, recruit-
ment and enrollment of handicapped children; screening, test-
ing, and diagnosis of the handicapping conditions; and activ-
ities directed towards the parents of handicapped children.
Although we do not have indicators for these variables, it
is quite likely that the items eventually included in the
index would account for most of the variation in each of
these.

Let us now review the results of the index construction
process.

- The two indices yielded identical results. Only one vari-
  able, COOR, was removed from the initial item pool in both
  instances. The bivariate correlation coefficient (Pearson)
  between the two final indices is 0.96.

- The two correlation matrices were almost identical. For
  this reason we illustrate in Table 3 only the matrix based
  on the original codes. The largest correlation coefficient
  among the 28 shown in Table 3 is .52 and the smallest is
  .04 (absolute value). The variable with the poorest cor-
  relation with the others is COOR, which was eventually
  dropped from the index. COOR is also the only variable with
  two negative correlation coefficients.

- Before removing COOR from the item pool we examined the cor-
  relation coefficients between each variable and the initial
  index (made up of eight items) minus that variable. The two
  item analyses yielded correlation coefficients which were
  significant at the .10 level for all variables with the
  exception of COOR.

- The two item analyses were repeated after removing COOR from
  the item pool. The new set of correlation coefficients were
  all significant at the .10 level. These results are illus-
  trated in Table 4.

The final index based on the original codes of the
variables is the one we chose to use in further analyses.

TABLE 3

PEARSON CORRELATION COEFFICIENTS BETWEEN
VARIABLES CONSIDERED FOR THE EFRT INDEX[a]

	PLAN	COOR	MOBR	FUND	SIZE	TATA	MFAC
ORGS	.34*	-.04	.44*	.22	.46*	.27*	.26*
PLAN		.29*	.23*	.31*	.13	.24*	.19
COOR			.19	.31*	.08	-.04	.10
MOBR				.41*	.26*	.52*	.27*
FUND					.24*	.12	.21
SIZE						.16	.25*
TATA							.21

[a](*) indicates that the coefficient is significant at the 0.10 level.

TABLE 4

FINAL ITEM ANALYSES FOR THE EFFORT INDEX[a]

| Variable | Pearson Correlation Coefficients between the Variable on the Left and the Index Minus the Variable on the Left | |
	Index Based on Original Codes	Index Based on Cumulative Frequencies
ORGS	.53*	.54*
PLAN	.37*	.35*
MOBR	.57*	.61*
FUND	.39*	.40*
SIZE	.40*	.40*
TATA	.37*	.35*
MFAC	.36*	.35*

[a](*) indicates that the coefficient is significant at the 0.10 level.

Our use of the cumulative percentages in building an alternative index did not result in any surprises. The similarity of the findings at every step increased the credibility of the index based on original codes and provided partial justification for carrying arithmetic operations with the original codes.

The frequency distribution of the final index is illustrated in Table 5. For simplicity, we converted the values of the final index into the [0,100] interval through a linear transformation. Thus, the maximum possible effort score a

TABLE 5

LEVEL OF EFFORT SCORES

Level of Effort Score	Number of Programs	Percentage of Programs
0 - 10.0	2	5.7
10.1 - 20.0	1	2.9
20.1 - 30.0	6	17.1
30.1 - 40.0	4	11.4
40.1 - 50.0	5	14.3
50.1 - 60.0	4	11.4
60.1 - 70.0	4	11.4
70.1 - 80.0	4	11.4
80.1 - 90.0	5	14.3
Total	35	99.9

program can have is 100 and the minimum is zero.

The distribution of effort scores has a mean of 49.8 and a median of 50.0 The distribution is quite symmetric and close to the normal distribution (the skewness coefficient is -.14, indicating slight skewness to the right). The coefficient of variation is 48.2%, indicating no abnormalities in the dispersion of the scores.

Implementation potential

This variable is highly policy-specific and, therefore, needs to be operationalized on the basis of the specific

capability requirements of the policy. We were able to
identify five indicators which related to different capabil-
ities of the local Head Start programs to implement the
handicapped policy. These five variables are the following:

EXPR : previous experience of the agency in serv-
ing severely handicapped children;

LEAD : lead time the agency has had prior to the
start of the implementation;

KNOW : the agency's knowledge of the size and the
characteristics of the handicapped children
who are eligible for participation in Head
Start;

QUAL : policy-related qualifications of the imple-
mentation staff;

AFAC : appropriateness (prior to the start of
the implementation) of the physical facil-
ities of the agency for serving handicapped
children.

These five variables do not exhaust all possible
capacity-related factors. Two of the most important factors
missing from the list are the leadership and the management
abilities of the principals of the implementing agency.[1]
Nevertheless the five variables on which we have data cover
widely different aspects of the agency's potential to imple-
ment the mandate.

The results of the index construction process in the
case of this variable were similar to those reviewed for
effort. The following is a summary of our findings.

---

[1]The data collection instrument used in the second round visits
included a set of questions on the leadership abilities of the Head
Start director. However, the indicators used were highly judgmental
and several interviewers felt uneasy about completing this portion of
the questionnaire. This resulted in a decision to cease the coding and
processing of all leadership questions.

- The two indices yielded identical results. In both instances two variables, AFAC and LEAD, were removed from the initial item pool. The Pearson correlation coefficient between the two final indices is 0.95.

- The two correlation matrices were again very similar. In Table 6 we illustrate the one based on the original codes. These coefficients are considerably smaller than the ones for EFRT. The variable with the poorest correlation with the others is AFAC, which was dropped first from the item pool because of both the weakness and the unexpected sign of the correlations it has with the other four variables.

- After removing AFAC from the item pool we examined the correlation coefficients between each of the remaining four variables and the initial index minus that variable. This item analysis resulted in the removal of LEAD from the item pool. The set of correlation coefficients which led to the removal of LEAD are shown in Table 7.

- A final item analysis was conducted after dropping LEAD from the index. This time the correlation coefficient between KNOW and the total score minus KNOW became insignificant at the 0.10 level. However, we decided to keep KNOW in the item pool mainly because of its correlation with EXPR. The final index for IPOT is thus made up of three variables: EXPR, KNOW, and QUAL.

The distribution of the final IPOT index (based on original codes) has only four values due to the fact that all three of the variables making up the index are dichotomies. Although it was not absolutely necessary, we transformed these four values to the [0,100] interval in order to maintain a measurement scale similar to that of EFRT. The resulting distribution of the four implementation potential scores are illustrated in Table 8. Briefly the mean score of the distribution is 52.4; the median score is 66.7; the distribution is highly symmetric (skewness coefficient = -.10); and the coefficient of variation is not high (C.V. = 60.3%).

TABLE 6

PEARSON CORRELATION COEFFICIENTS BETWEEN
VARIABLES CONSIDERED FOR THE IPOT INDEX[a]

	LEAD	KNOW	QUAL	AFAC
EXPR	.14	.23*	.40*	-.25*
LEAD		.21	.03	-.06
KNOW			.10	-.01
QUAL				-.03

[a](*) indicates that the coefficient is significant at the 0.10 level.

TABLE 7

INITIAL ITEM ANALYSIS FOR THE IPOT INDEX[a]

Variable	Pearson Correlation Coefficient between the Variable on the Left and the Index Minus the Variable on the Left	
	Index Based on Original Codes	Index Based on Cumulative Frequencies
EXPR	.41*	.43*
LEAD	.18	.21
KNOW	.26*	.24*
QUAL	.30*	.33*

[a](*) indicates that the coefficient is significant at the 0.10 level.

TABLE 8

IMPLEMENTATION POTENTIAL SCORES

Implementation Potential Score	Number of Programs	Percentage of Programs
0	5	14.3
33.3	11	31.4
66.7	13	37.1
100.0	6	17.1
Total	35	100.0

Environmental forces for
implementation

We have data on only two variables that are related to pressures directed towards the implementing agency by other agencies and/or groups of individuals. The first of these variables, COMM, refers to the reaction of the agencies in the community, who serve handicapped children, to the new policy mandated to the local Head Start program. The second variable, TARG, relates to the reaction of the parents of children in Head Start. As described in Appendix B, both of these variables are coded in terms of a five point scale ranging from strongly disagree to strongly agree.

We believe that COMM and TARG are quite adequate for measuring ENVR. It would have been desirable to have separate data on the reaction of the parents of handicapped as well as non-handicapped children. Similarly, data on the

reactions of local agencies which do not serve handicapped children but which have strong linkages with Head Start may have improved the index.

Our findings from the index construction process for ENVR can be summarized as follows:

- As before, the results obtained from the two alternative index construction schemes were similar. Both of the initial variables remained in the item pool in both instances. The bivariate correlation coefficient between the two final indices is 0.98.

- The two correlation matrices each have only one entry since there are only two variables in the item pool. COMM and TARG were closely related to one another as indicated by a coefficient of 0.57 in the case of original codes and 0.58 in the case of cumulative frequencies.

- Because we are starting with only two variables, we did not conduct an item analysis. The correlation between each variable and the total score minus that variable is identical to the correlation between the two variables in the item pool.

The distribution of the final ENVR index (based on original codes) is illustrated in Table 9. As before the values of the final index were linearly transformed in order to obtain ENVR scores which range from zero to 100. The mean of the distribution is 62.9 and the median is 62.5. These two figures indicate that, on the average, the policy was slightly supported by parents and competing agencies.[1] This is also illustrated by the fact that the distribution is slightly skewed to the right (skewness coefficient = -.47). Finally, the dispersion of the ENVR scores is in the normal range, as indicated by a coefficient of variation figure

---

[1]This conclusion is based on the observation that the point of "neutrality" to the policy by both the parents and the competing agencies is 50.0.

TABLE 9

FINAL SCORES FOR ENVIRONMENTAL
FORCES FOR IMPLEMENTATION

Final Score	Number of Programs	Percentage of Programs
0	1	2.9
25.0	2	5.7
37.5	1	2.9
50.0	9	25.7
62.5	9	25.7
75.0	6	17.1
87.5	4	11.4
100.0	3	8.6
Total	35	100.0

of 34.9%.

Policy support

The policy support index is based on two indicators.
The first, AGRM, illustrates the Head Start director's level
of agreement with the policy in general. The second, ATTD,
relates to the attitudes of the director towards serving
severely handicapped children in Head Start. AGRM is meas-
sured directly on the basis of a five-point scale ranging
from strongly disagree to strongly agree. ATTD is measured
in terms of a Likert-type attitude scale constructed from
the responses of the director to the five attitude items

quoted in Appendix B. Construction of this scale involved steps identical to the procedure used in building the indices for EFRT, IPOT, ENVR, and PSUP. The ten correlations among the five attitude items were all positive and significant at the 0.10 level. Similarly, the five correlations between each attitude item and the total attitude score minus that item were also significant at the 0.10 level. Thus, none of the initial attitude items were dropped and the final attitude score was constructed by adding the coded responses to the five items.

Here are the results of the index construction process for PSUP:

- The two alternative indices yielded identical results illustrated by a correlation coefficient of 0.98 between the two final indices. Both AGRM and ATTD remained in the index after they were each subjected to the screening criteria.

- The coefficients of correlation between AGRM and ATTD are 0.49 and 0.50 for the two alternative procedures. Thus, both variables were included in the index for PSUP.

The frequency distribution of the final PSUP index, transformed into the [0,100] interval is illustrated in Table 10. The arithmetic mean of the distribution is 63.7 and the median PSUP score is 65.0. The skewness coefficient is -.50, indicating skewness to the right; and the coefficient of variation is 38.1%, indicating that the dispersion of final PSUP scores is not abnormal.

Self-evaluation of capability

The indicator we have for this variable is based on the program director's estimate of the specific types of severe

TABLE 10

POLICY SUPPORT SCORES

Policy Support Score	Number of Programs	Percentage of Programs
10.1 - 20.0	2	5.7
20.1 - 30.0	2	5.7
30.1 - 40.0	2	5.7
40.1 - 50.0	6	17.1
50.1 - 60.0	3	8.6
60.1 - 70.0	4	11.4
70.1 - 80.0	5	14.3
80.1 - 90.0	7	20.0
90.1 -100.0	4	11.4
Total	35	99.9

handicapping conditions the program can accomodate. Thus, the range of the original indicator is from zero (indicating that the program director believes his program does not have the capabilities which he thinks would be required to serve any type of severe impairment) to ten (which corresponds to the director's opinion that the program, overall, has the capabilities to serve severely handicapped children with any of the ten handicapping conditions). We transformed the integer values of this variable into the [0,100] interval in order to be consistent with all the other previous

indicators. This means that after this transformation the scores for CAPB illustrate the percentage of the ten severe handicapping conditions the program has the capabilities to serve, as estimated by the program director. The frequency distribution of these CAPB scores are displayed in Table 11.

TABLE 11

SELF-EVALUATION OF CAPABILITY SCORES

CAPB Score	Number of Programs	Percentage of Programs
0	5	14.3
0.1 - 10.0	3	8.6
10.1 - 20.0	3	8.6
20.1 - 30.0	3	8.6
30.1 - 40.0	0	0
40.1 - 50.0	3	8.6
50.1 - 60.0	3	8.6
60.1 - 70.0	6	17.1
70.1 - 80.0	2	5.7
80.1 - 90.0	2	5.7
90.1 -100.0	5	14.3
Total	35	100.0

The average CAPB score is 50.6 and the median score is 56.7, indicating that, on the average, the program directors believe their programs have the capabilities required

to serve severely handicapped children with about half of the ten handicapping conditions. The distribution is quite symmetric (skewness coefficient = -.10) and the dispersion of the scores is not too high (coefficient of variation = 69.0%).

This completes the presentation of the indicators that are being used for measuring the six variables of the model. Selected characteristics of the distributions of these six indicators are illustrated in Table 12. Since we have already commented on these statistics earlier, we will close this chapter simply by observing that none of the six distributions seem to have any "abnormalities" which would make their utilization in multiple regression analyses questionable.

TABLE 12

SELECTED CHARACTERISTICS OF THE DISTRIBUTIONS
OF INDICATORS USED FOR MEASURING THE
VARIABLES OF THE MODEL

Variable	Potential Range	Arithmetic Mean	Median	Skewness Coefficient	Coefficient of Variation
LIS	0 - 200.0	19.1	15.4	0.59	72.5%
EFRT	0 - 100.0	49.8	50.0	-0.14	48.2
IPOT	0 - 100.0	52.4	66.7	-0.10	60.3
ENVR	0 - 100.0	62.9	62.5	-0.47	34.9
PSUP	0 - 100.0	63.7	65.0	-0.50	38.1
CAPB	0 - 100.0	50.6	56.7	-0.10	69.0

CHAPTER V

FINDINGS

In this chapter we present the empirical findings from
the study. This is accomplished in two stages. First we
concentrate on the more descriptive and simpler generaliza-
tions from the data. We then move towards the analytical and
more complicated findings. Accordingly, the simpler findings
are summarized in the first section of the chapter in the
form of description and discussion of the secondary variables
which were used in developing the indices. The more compli-
cated findings are reviewed in the second section through
comparisons of the alternative models presented earlier in
Chapter III in light of the Head Start data.

A Descriptive Profile of the Implementation

Our purposes in this section are threefold. First, we
will draw a profile of the implementation in terms of the 18
indicators used for defining the four independent and one
intervening variable of the conceptual model. This is con-
ducted in terms of a review of the univariate frequency dis-
tributions of these 18 indicators. Second, we will examine
the statistical association between each of the 18 variables
and LIS, our major dependent variable. In order to

154

accomplish this we reviewed the frequency distribution of LIS and divided the 35 programs into three groups representing low, moderate, and high success. This resulted in classification of 11 programs with the lowest LIS scores as "low success," 10 programs with the highest LIS scores as "high success," and the 14 remaining programs as "moderate success" programs.[1] Therefore, statistical associations between the 18 secondary variables and LIS are examined through a set of bivariate tables where LIS takes on only three values. Our third and final purpose is the examination of statistical relationships among the 18 variables.

In order to be consistent with the preceding chapters, we organized presentation of the simpler findings along the lines of the explanatory variables used in the conceptual model. Thus, we first examine the variables used in defining IPOT. This is followed by descriptive profiles of the 35 programs in terms of variables used for measuring ENVR, PSUP, CAPB, and EFRT. The most significant tables referred to in the discussion are included in the text, others are contained in Appendix C.

---

[1]We attempted to make the sizes of the three groups as equal as possible. The final categorization is based on examination of the natural breakpoints in the frequency distribution of LIS. LIS scores of the programs in the low success group range from 0 to 11.1; the range for those in the moderate success group is 12.9 to 21.0; and the high success group includes programs with a minimum LIS score of 27.5 and a maximum of 51.6.

Pre-Implementation Differences
Among Programs

Head Start program directors, in general, found out about the requirement to serve handicapped children through official notification from their respective regional offices. Some had been notified directly by the national office of OCD and a few first heard about the mandate through informal contacts. When the grantee agencies were notified about the new policy, they, in turn, informed their delegate agencies about the new development in Head Start. The number of agencies involved in the communication chain (Congress $\longrightarrow$ HEW $\longrightarrow$ OCD $\longrightarrow$ regional office $\longrightarrow$ grantee agency $\longrightarrow$ delegate agency) and the manner in which each regional office handled the initial notification resulted in wide differences among the sampled programs in terms of lead time they had prior to the start of the implementation. The directors of 12 programs indicated that they heard about the policy only within the month prior to the start of their program year (See Table 13). On the other hand, 12 program directors said that they knew about the handicapped mandate at least 5 months prior to the start of the 1973-74 program year. The mean of the frequency distribution for LEAD is 3.6 months and the median is 3 months.

One would ordinarily expect that programs with longer lead times would have a greater potential to get ready for the implementation. In particular, since most of the recruitment for the program year which starts in fall takes

TABLE 13

LEAD TIME BY IMPLEMENTATION SUCCESS

Lead Time	Level of Implementation Success						Total	
	Low		Moderate		High			
	#	%	#	%	#	%	#	%
1 month or less	6	54.5	3	21.4	3	30.0	12	34.3
2-4 months	3	27.3	6	42.9	2	20.0	11	31.4
5-9 months	2	18.2	5	35.7	5	50.0	12	34.3
Total	11	100.0	14	100.0	10	100.0	35	100.0

place during the spring months, programs which had at least three months lead time would be in a position to recruit handicapped children along with the non-handicapped they identify in the spring. Furthermore, a longer lead time means a greater potential to plan the implementation, recruit staff, develop programs for the handicapped, and establish linkages with community agencies for arranging provision of special services which cannot be met with Head Start's resources. Thus, we expected lead time to be strongly correlated with implementation success. However, the data do not confirm this hypothesis. Kendall's tau b, which measures the strength of the association between the two ordinal variables, is 0.25, indicating a slight positive correlation. The chi square for Table 13 is 4.66, which, for 4 degrees of freedom, is not significant even at the 0.30 level.

One possible explanation for the lack of association between the length of lead time and implementation success is that the programs did not start "working on" the implementation immediately after being notified about the policy. The official notification did not provide answers to the many questions the program directors had about the mandate. There were no clear guidelines as to who would be considered "handicapped." Besides, the handicapped policy was not the only policy being implemented by Head Start and most program directors were not ready to take their own or their staff's time away from the other Head Start activities. Thus, they chose to wait until they received clearer guidelines, which were promised but were never sent. For some of the programs, however, the policy was not something new since they had been serving handicapped children prior to the notification.

The sampled programs also differed in terms of their previous experiences in serving the severely handicapped. 16 of the 35 programs, according to the directors interviewed, had no severely handicapped children enrolled during the 1972-73 program year. The remaining 19 programs each had at least one severely handicapped child during the year preceding the congressional requirement. For many of the programs in the second group handicapped policy was formalization of what they had been doing all along. However, since they were not required to keep records on their involvement with the handicapped, data were not consistently available on the numbers, handicapping conditions, and severity levels of the

children enrolled during the previous years. The following comments by the directors exemplify some of the most frequently mentioned reasons for lack of this information.

> (They) did not keep separate records.
> (They) did not label.

> (The children) were not labeled as handicapped—only identified as children with special needs.

> (Staff) were not documenting at the time.

> We had them and knew of hearing and speech problems . . . but we didn't pull the information out as separate information.

We expected previous experience in serving the severely handicapped to be strongly correlated with implementation success and the data confirm this hypothesis. As shown in Table 14, all 10 of the programs in the high success group had previous experiences with the severely handicapped as compared to only 18.2% of those in the low success group. The chi square value (14.3, 2 d.f.) is significant at the 0.01 level and a tau c value of 0.68 indicates a very strong positive correlation between past experience and implementation success.[1]

In addition to lead time and past policy experience, there are three other capacity related factors we will examine briefly. These relate to knowledge, by the program director or his staff, of the size of the handicapped population eligible for Head Start, qualifications of the implementation staff, and appropriateness of the physical facilities

---

[1] We are using Kendall's tau b in instances where the table being studied is a square table and tau c in cases where the table is rectangular.

TABLE 14

PREVIOUS EXPERIENCE IN SERVING
THE SEVERELY HANDICAPPED BY
IMPLEMENTATION SUCCESS

Previous Experience	Level of Implementation Success						Total	
	Low		Moderate		High			
	#	%	#	%	#	%	#	%
No previous Experience	9	81.8	7	50.0	0	0.0	16	45.7
Some previous Experience	2	18.2	7	50.0	10	100.0	19	54.3
Total	11	100.0	14	100.0	10	100.0	35	100.0

of the program for serving the severely handicapped. The programs in the sample did not exhibit as much variation in terms of any of these three factors as they did in the cases of past policy experience and lead time. For this reason we will review our findings on these factors in a summary fashion.

● Close to 80% of the program directors indicated that they had no knowledge of, nor could they estimate, the number of handicapped children in their respective service areas who would be eligible for Head Start. Of the 8 directors who said that they knew the Head Start-eligible handicapped population in their areas, 6 (or 75%) were heading up programs classified as high success. Despite the unevenness of the frequency distribution of KNOW, the chi square and tau c values (11.0--which is significant at the 0.01 level--and 0.41, respectively) indicate positive statistical correlation between KNOW and LIS. However, we do not know whether past and present involvement with handicapped children led to increased knowledge or knowledge of the size of the handicapped population resulted in a better recruitment effort and eventual success. In any event, it appears that a great majority of the programs do not know the size of their

potential "market" and those that do know also happen to be in the more successful groups. (See Table C. 1)

● Four-fifths of the program directors interviewed indicated that their staffs included at least one person who had completed at least one college level course in special education. All 10 of the programs on the high success group had at least one such person in the staff, as compared with 54.5% of the programs in the low success group (See Table C. 2). QUAL and LIS were statistically associated, as indicated by a chi square value significant at the 0.05 level and a tau c value of 0.39.

● About three-fourths of the program directors indicated that their physical facilities were not appropriate at the beginning of the program year to serve all types of severely handicapped children. The low, moderate, and high success programs did not statistically differ from one another in terms of appropriateness of physical facilities. In fact, a greater proportion of the low success programs (36.4%) had appropriate physical facilities than in the moderate and high success groups (21.4% and 20.0%, respectively). The tau c value is quite insignificant (-0.14); so is the chi square (0.96, which becomes significant only at the 0.65 level). These results are quite surprising, especially in light of the fact that one of the most frequent excuses given for not enrolling severely handicapped children is the inappropriateness of the program's physical facilities. Figures in Table C.3 show that appropriate physical facilities is not a "must" for serving most severely handicapped children.

In summary, the "typical" Head Start program in the sample was notified about the mandate three months prior to the start of the implementation; it had some previous experience in serving the handicapped; the program staff had no knowledge of their potential "market;" the program included at least one person with background in special education; and the physical facilities of the program were not appropriate for serving all types of severely handicapped children. Two of the variables studied, lead time and appropriateness of facilities, were not significantly related to implementation success, the remaining three were significantly associated

with it. The strongest correlation found was between previous experience in serving the severely handicapped and implementation success.

Pressures from the Outside

An important issue in policy implementation is the extent to which the implementing agency carries out the mandate partly as a result of pressures from the outside. By the term "outside" we are referring to any and all interested parties other than the agencies administering or implementing the policy. In the case of the handicapped policy, there are two such local parties which may have influenced the policy-related decisions in Head Start. One of these is the other community agencies which provide services to handicapped children and the second one is the group or groups representing the parents of children in Head Start.

Reactions of community agencies to Head Start's new policy are important for two reasons. First, Head Start's entry into the "preschool handicapped market" implies that in some instances Head Start would be competing against these agencies for sponsorship of Head Start eligible children. Second and more importantly, because the mandate is not accompanied with additional funding, Head Start has to rely on these agencies for providing specialized services to the handicapped. Prominent among these services are periodic screening, testing, and evaluation; diagnosis, treatment, and therapy; and specialized counseling and consultation.

Local agencies which most frequently provide these services
include the following:

- hospitals, medical centers, clinics

- public health departments

- community mental health programs

- special education departments of school districts or colleges/universities

- specialized agencies such as crippled children's services and commissions for the blind

- civic or community clubs such as the Lion's Club.

The data on the reactions of these agencies to the
handicapped policy are based on the perceptions of the Head
Start director. Accordingly, as shown in Table C.4, about
half of the Head Start directors felt that these agencies
were in general supportive of the policy. 15 directors felt
that they were by-and-large "neutral." Only 3 directors
indicated that they perceived negative sentiment towards the
mandate. The low, moderate, and high success groups dif-
fered little in terms of this variable (Kendall's tau c =
0.38; chi square is not significant at the 0.10 level).

Reactions of the parents of Head Start children to the
policy are also quite important. In particular, since some
of the parents serve on the Policy Council or the Parent
Advisory Board, they could potentially influence Head Start's
actions during the implementation. It is also probable that
members of these boards, who are primarily parents of non-
handicapped children, may not want to have their children
associate with the severely disabled.

Findings on the reactions of the parents are almost identical to the findings about the reaction from community agencies (See Table C.5). About half of the program directors felt that the parents were neutral to the policy. 42.9% indicated the parents supported the mandate. Only 3 program directors felt negative sentiments on the part of the parents. The three success groups did not differ markedly in terms of this variable (tau c = 0.37; chi square value not significant at the 0.05 level).

We also crosstabulated the parents' reaction against reaction from the community groups (See Table C.6). The two variables were significantly associated with each other (tau b = 0.42; chi square significant at the 0.01 level). Part of this association could be attributed to the fact that both variables were measured in terms of the perceptions of the same person, i.e., the Head Start director.

In summary, about one-half of the directors felt that agencies in the community which also serve handicapped children were supportive of the new policy. Opposition to the policy by community agencies was rare. Findings on the reaction from the parents of Head Start children were almost identical. Finally, the three success groups were not appreciably different from one another in terms of either of these two variables.

Head Start Directors' Reactions
to the Policy

Head Start directors, in general, had more negative
reactions towards the ways they were notified about the
policy than to the policy itself. 60% of the directors were
in agreement with the intent of the requirement, 17% were
neutral, and 23% disagreed with it (See Table 15). The impo-
sition of the mandate, without prior warning by the regional
offices, and the establishment of a quota were viewed very
negatively.

TABLE 15

AGREEMENT WITH THE POLICY BY
IMPLEMENTATION SUCCESS

Director's Agreement With the Policy	Level of Implementation Success						Total	
	Low		Moderate		High			
	#	%	#	%	#	%	#	%
Strongly disagree	1	9.1	0	0.0	0	0.0	1	2.9
Disagree	4	36.4	3	21.4	0	0.0	7	20.0
Neutral	3	27.3	2	14.3	1	10.0	6	17.1
Agree	3	27.3	4	28.6	2	20.0	9	25.7
Strongly Agree	0	0.0	5	35.7	7	70.0	12	34.3
Total	11	100.1	14	100.0	10	100.0	35	100.0

Lack of definitions and comprehensive guidelines gave most program directors the feeling that the Congress, HEW, and OCD were "passing the buck" down to the local level without any concern for the difficulties of the implementation.

The following quotes from the questionnaires exemplify the initial reactions of the program directors:

" . . . Initial reaction was one of apprehension and worry over children (e.g., blind) walking around with no assistance."

" . . . Initial reaction was neutral, but more concern arose when it was realized that severely handicapped children were to be served."

"If the children are mildly handicapped, I (the program director) have no problem; but with the severely handicapped, we would need more money, staff, and transportation."

"(The staff was) very concerned about how they were going to handle seriously handicapped children. (They) disagreed because there were not enough facilities or staff, but agreed with the philosophy of it all."

"We had no real feelings (about the mandate). We had already been doing it."

"We always had a large number of handicapped. We had over 10 per cent last year without the mandate."

"(We) felt that Head Start always served the handicapped anyway."

Agreement with the policy differed significantly across the three groups. 72.8% of the program directors in the low success group were either neutral to the policy or they were against it. The corresponding figures for the moderate and high success groups are 35.7% and 10.0%, respectively. Strong statistical association between agreement and success is also indicated by the value of tau c (0.56).

Agreement with the mandate shows only part of the

picture regarding support of the policy by the program direc-
tors. Another aspect of support is reflected by the atti-
tudes of the directors towards serving the severely impaired.
To assess this, we constructed a Likert-type attitude scale
the frequency distribution of which is displayed in the last
column of Table 16.[1]

TABLE 16

ATTITUDES TOWARDS SERVING THE SEVERELY
HANDICAPPED IN HEAD START BY
IMPLEMENTATION SUCCESS

Total Attitude Score	Level of Implementation Success						Total	
	Low		Moderate		High			
	#	%	#	%	#	%	#	%
10 or less	4	36.4	1	7.1	0	0.0	5	14.3
11-15	4	36.4	2	14.3	2	20.0	8	22.9
16-20	2	18.2	4	28.6	3	30.0	9	25.7
21-25	1	9.1	7	50.0	5	50.0	13	37.1
Total	11	100.1	14	100.0	10	100.0	35	100.0

The findings on attitudes are similar to the earlier find-
ings on agreement with the policy. 62.8% of the program
directors had favorable attitudes towards serving the
severely handicapped. Only 14.3% expressed negative

---

[1]The items which entered the scale are listed in Appendix B.
Results of the item analyses conducted prior to scale construction
are summarized in Chapter III.

attitudes. Attitudes and implementation success were sta-
tistically correlated (tau c = 0.41; chi square significant
at the 0.10 level) but not as strongly as in agreement vs.
success.

We also investigated the statistical association
between agreement and attitudes but did not find as strong
a relationship as we expected (tau c = 0.36; chi square
significant only at the 0.10 level). An examination of
Table 17 reveals that of the 8 directors who disagreed with.
the policy, 2 had favorable attitudes towards serving the
severely handicapped (those with attitude scores between 16
and 25). Alternatively, of the 21 directors who agreed with
the policy, 2 had negative and another 2 had "neutral" atti-
tudes towards serving the severely handicapped. These
results indicate that in a few instances program directors
may disagree with the policy for reasons other than unfavor-
able attitudes towards the target group. Reasons for the
disagreement may include factors such as opposition to the
manner in which they were notified, negative reaction from
the parents or other community agencies, or complete dis-
agreement with an aspect of the policy such as the belief
that the 10 per cent quota would lead to mislabeling of some
children. At the other extreme, the data indicate that in
a few instances program directors may agree with the policy
even though they have negative attitudes towards the severely
handicapped. This inconsistency can be partially explained
by the argument that a director who has negative attitudes

TABLE 17

ATTITUDES TOWARDS SERVING THE SEVERELY
HANDICAPPED IN HEAD START BY AGREEMENT
WITH THE POLICY

Total Attitude Score	Director's Agreement with the Policy							
	Strongly Disagree or Disagree		Neutral		Agree or Strongly Agree		Total	
	#	%	#	%	#	%	#	%
10 or less	3	37.5	0	0.0	2	9.5	5	14.3
11-15	3	37.5	3	50.0	2	9.5	8	22.9
16-20	1	12.5	1	16.7	7	33.3	9	25.7
21-25	1	12.5	2	33.3	10	47.6	13	37.1
Total	8	100.0	6	100.0	21	99.9	35	100.0

towards the target group may still agree with the policy because he feels his agency is obligated to implement a decision made by the Congress. Alternatively, and perhaps more likely, the director may interpret the intent of the policy as including mainly the less disabled children and agree with it despite the fact that he has negative attitudes towards serving the severely impaired.

We tested three other statistical hypotheses related to relationships of either agreement or attitudes with other variables. The first two of these concern statistical association of agreement with reaction from community agencies and parents' reaction to the policy. The third one relates

to the connection between previous experience of the program

in serving the severely handicapped and the director's atti-

tudes towards the severely handicapped.

- The crosstabulation of the director's perception of the reaction of the community agencies and his own agreement with the policy is displayed in Table C.7. There seems to be a moderate level of association between the two variables (tau b = 0.41; chi square significant at the 0.05 level). Two thirds of the 21 program directors who agreed with the policy believe that the community agencies also support the Head Start mandate. The statistical association between these two variables suggests that the director's own reactions to the policy may have been shaped, in part, by his perceptions regarding the amount of support he can get from other agencies in the community. Without such support Head Start could not meet the special needs of particularly the severely handicapped since the programs were not given additional funds for purchasing the services needed by these children. The director's agreement with the policy, therefore, could have been influenced by his perceptions about the degree of cooperation he would receive from the agencies which provide specialized services to the handicapped.

- In contrast with the finding reported above, Table C.8 shows that the director's perception of the reaction of parents and his own agreement with the policy are not statistically related (tau b = 0.31; chi square significant only at the 0.35 level). This could be due to a lack of influence on the part of the parents or due to the fact that close to half of the directors believed that the parents were "neutral" to the policy (as compared to only 17.1% of the directors who were neutral). It is also likely that, in instances where the reaction of the parents differed from the reaction of the director, the director may have thought he could convince the parents that Head Start should serve/ not serve handicapped children.

- An important question about implementation concerns the relationships between attitudes and past behavior. The most common hypothesis is that the directors of programs which have previously served severely handicapped children would have more favorable attitudes towards serving the severely handicapped than the directors of programs with no such previous experience. This is based on the argument that as a result of their past experience, the directors of programs in the former group will have overcome the "worries" associated with serving severely handicapped children and observed the benefits—for the children, their parents, and the program—resulting from their participation in Head Start.

The data at hand do not confirm this hypothesis. As shown in Table C.9, attitudes and experience are not statistically related (tau c = 0.28; chi square significant only at the 0.50 level). There are three possible explanations for this finding. First, some of the directors were new and did not share the experiences of the other staff. Their attitudes therefore, could not have been influenced by the program's past experience. Secondly, in some of the programs with previous experience, these experiences may have led to unpleasant impressions about serving the severely impaired. Finally, it is possible that the current attitudes of the directors were formed independently of the past experience of the program.

In sum, more than half of the directors were in agreement with the policy and a similar proportion had favorable attitudes towards serving the severely handicapped in Head Start. We found strong statistical association between agreement and implementation success. Attitudes were also positively related to success, but not as strongly as in agreement vs. success. The relationship between agreement and attitudes was also not strong. The director's agreement with the policy was positively correlated with his perceptions of the support of the policy by the community agencies, but it was not significantly correlated with his perception of the parents' reaction to the mandate. Finally, we found that attitudes of the directors towards serving the severely handicapped were not at all related to the program's previous experience in serving the severely handicapped.

Directors' Perceptions of Head Start's Capabilities

Only 7 out of 35 directors believed that their programs had the capabilities to serve all types of severely handicapped children (See Table 18). The average number of severe

TABLE 18

SELF-EVALUATION OF CAPABILITY BY
IMPLEMENTATION SUCCESS

Number of Handicapping Conditions[a]	Level of Implementation Success						Total	
	Low		Moderate		High			
	#	%	#	%	#	%	#	%
0	4	36.4	1	7.1	0	0.0	5	14.3
1 - 2	4	36.4	2	14.3	0	0.0	6	17.1
3 - 4	1	9.1	1	7.1	1	10.0	3	8.6
5 - 6	1	9.1	2	14.3	3	30.0	6	17.1
7 - 8	1	9.1	4	28.6	3	30.0	8	22.9
9 -10	0	0.0	4	28.6	3	30.0	7	20.0
Total	11	100.1	14	100.0	10	100.0	35	100.0

[a]This column shows the program director's estimate of the number of
different severe handicapping conditions of the program has the capa-
bilities to serve.

handicapping conditions which, in the opinions of the direc-

tors surveyed, Head Start could serve is 5 and the median is

6. The severe disabilities the directors believed to be most

difficult to cope with in Head Start were total blindness,

total deafness, mental retardation and severe physical impair-

ments. Alternatively, the directors thought that their pro-

grams could more easily serve severely speech impaired,

health or developmentally impaired, visually impaired, or

hearing impaired children. Five program directors indicated

that they believed their programs could not serve children

with any type of a severe impairment.

Capability assessments by program directors differs sharply across the three success groups. 72.8% of the directors of programs in the low success group believe that their programs can accomodate no more than 2 severe impairments. The corresponding percentages for the programs in the moderate and high success groups are 21.4% and 0%, respectively. The tau c value is 0.52, indicating strong statistical association between self-evaluation of capability and implementation success.

We also investigated the statistical association between the directors' assessment of capability and their level of agreement with the policy (See Table C.10). As we suspected, the relationship between these two variables was statistically significant (tau c = 0.46). Of the 11 directors who believed their programs could accomodate no more than 2 severe impairments 6, or 55%, disagreed with the policy. At the other end of the spectrum, 87% of the 15 directors who thought their programs could handle at least 7 handicapping conditions were in agreement with the policy. These findings suggest that a program director's level of agreement with the policy could have been influenced, in part, by his perceptions about the implementation capabilities of his agency.

Disabling Conditions of Children
Enrolled in Head Start

The first and perhaps the most confusing issue the programs faced at the start of the implementation was the matter of definition for the handicapping conditions. OCD did not provide clear guidelines in this regard and the programs either developed their own definitions--with assistance from other agencies--or used the definitions which were included in SRI's mail survey of all programs. Confusion over the definitional issue led to initial mislabeling of some normal children as "handicapped." The issue had not completely been resolved by the time the second round visits were conducted, i.e., approximately 8 months after the start of the implementation. The reliability of the reported number of handicapped children, therefore, is not known, especially for the mildly/moderately handicapped group.

There were a total of 8,860 children enrolled in the 35 programs included in the sample. This represents about 3.1% of the total number of children enrolled in Head Start during the 1973-74 program year. 1,218--or 13.7%--of the 8,860 children enrolled in the sampled programs were reported to be handicapped. Of this number, only 21.3% were reported as severely impaired, the remaining group of children were considered as mildly/moderately handicapped. The reported disabling conditions of the children included in the severely handicapped and the mildly/moderately handicapped groups are shown in Table 19.

TABLE 19

PRIMARY DISABLING CONDITIONS OF HANDICAPPED
CHILDREN ENROLLED IN HEAD START BY SEVERITY
OF THE HANDICAP

Disabling Condition	Percentage of All Severely Handicapped Children	Percentage of All Mildly/Moderately Handicapped Children
Blind	1.2	0.6
Visually impaired	6.5	6.4
Deaf	1.9	0.3
Hearing impaired	6.2	10.3
Health or developmentally impaired	12.3	13.7
Physically impaired	12.7	8.8
Speech impaired	34.6	35.4
Seriously emotionally disturbed	11.5	17.1
Mentally retarded	9.6	7.4
Undifferentiated diagnosis	3.5	0.0
Total	100.0	100.0
(N)	(260)	(958)

Speech impairment is the most frequent disability reported by the sampled programs. In the case of the severely impaired, this is followed by physical and health or developmental impairments. For the mildly/moderately handicapped, speech-related disabilities are followed by serious emotional disturbance and health or developmental

impairments. In both instances blindness and deafness are the least frequent disabilities.

The prevalence figures for Head Start reported above are not significantly different from the distributions of such disorders reported by others. Burton Blatt, in a special study he undertook within the SRI-Syracuse University project, summarizes the findings from recent studies on prevalence estimates of handicapping conditions in school-age children.[1] We took three of the four such estimates reported by Blatt (the three which corresponded most closely to our handicapping condition categories) and for each estimate calculated the relative share, within the handicapped population, of each disabling condition. We also combined our severely and mildly/moderately categories to obtain a single distribution. This distribution is shown in Table 20 along with the average percentages derived from the three estimates for school-age children.

An examination of the percentages shown in Table 20 reveals that the two distributions are strikingly similar, especially for the speech impaired, emotionally handicapped, and the health or developmentally disabled. The large discrepancy between the percentages for the mentally retarded can be explained by the fact that " . . . mental retardation is unidentified in the preschool years, except when it is

_____

[1]Burton Blatt, "Handicapped Children in Head Start: A Discussion on Epidemiology, Policy Development, and Program Planning" in "Interim Report on Assessment of the Handicapped Effort in Experimental and Selected Other Head Start Programs Serving the Handicapped."

TABLE 20

DISABLING CONDITIONS OF HANDICAPPED
SCHOOL-AGE AND HEAD START CHILDREN[a]

Disabling Condition	Percentage of All Handicapped Children in Head Start	Average Percentages for School-Age Children[b]
Visually handicapped	7.3	0.6
Hearing impaired	10.1	4.4
Health or developmentally impaired	13.4	11.2
Physically impaired	9.6	4.3
Speech impaired	35.2	38.1
Emotionally handicapped	15.9	21.5
Mentally retarded	7.9	19.1
Other	0.7	0.5
Total	100.1	99.7

[a]"Visually handicapped" and "hearing impaired" categories include the blind and the deaf which were previously reported as separate conditions.

[b]This column is based on a simple arithmetic average of the percentage distributions derived from the data reported by Blatt (See the text).

accompanied with central nervous system pathology, clinical stigmata, and, therefore, with usually moderate or severe disability."[1] Thus, in spite of the difficulties faced by most Head Start programs in defining the handicapping conditions, the figures reported by the sampled programs are

---

[1]Ibid., p. 20.

generally in line with the prevalence estimates for the school-age population.[1]

<div align="center">

Selected Aspects of the
Implementation
</div>

Most of the children reported as handicapped came to Head Start through one or more of the following: (1) regular recruitment efforts of the Head Start staff, (2) referrals by community agencies, and (3) referrals by parents of handicapped children. The disabling conditions of most of these children were reportedly diagnosed by "qualified professionals". In many instances the child's handicapping condition had been diagnosed prior to his enrollment in Head Start. In all the programs studied here, handicapped children were totally integrated with the nonhandicapped. There were no special classrooms for the handicapped and--except for regular therapy, treatment, or counseling sessions conducted outside the center--the handicapped and the nonhandicapped were served within the same physical setting.

We have specific data on eight different aspects of the policy's implementation by the programs in the sample. We previously referred to these data in connection with the index constructed for measuring the level of effort expended by each agency in implementing the policy. Here we will review our basic findings concerning each of these different aspects of the implementation.

---

[1]We have compared the Head Start data with data for the school-age population mainly because of the unavailability of prevalence estimates for preschool-age children.

Organizing for the
Implementation

In most of the programs one person was in charge of coordinating the implementation. In some programs this person was the Head Start director himself, in others the health services coordinator or another staff member was assigned the overall responsibility to implement the mandate. Some of the larger programs hired a person specifically for this purpose. In some cases more than one person was in charge.

We examined the percentage of the coordinator's time devoted to the implementation (See Table C.11). In 8 cases there was no coordinator (or someone was named as the coordinator but he did not devote any of his time to the implementation). In 14 programs the coordinator devoted up to 25% of his time to the implementation of the policy and in 13 programs the person in charge of the mandate spent between 25% and 100% of his time on the implementation. This variable is very strongly correlated with implementation success (tau b = 0.62; chi square significant at the 0.01 level) indicating that in the more successful programs greater amounts of effort were being expended to organize for and coordinate the implementation.

Planning

Eight of the thirty-five program directors indicated that they did not prepare any kind of a written implementation plan. In 7 programs a plan was prepared but it covered

up to 4 of the 8 planning areas the directors were probed
about. The directors of the remaining 20 programs indicated
that they had prepared a plan and that it covered at least
5 of the 8 planning areas (See Table C.12). Comprehensive-
ness of the plan (as measured by the number of planning
areas covered) is not strongly associated with implementa-
tion success (tau b = 0.39; chi square significant only at
the 0.15 level). This is because some of the more success-
ful programs implemented the policy without a written plan.
Conversely, some of the programs with comprehensive plans
did not succeed in enrolling handicapped children. This
was partly because they could not execute their plans
successfully. It is also possible, of course, that they
may have started with a "bad" plan in the first place.

Coordination of the Implementation with other agencies

All but 2 of the program directors indicated that they
were coordinating their efforts with either the regional
office or other Head Start programs or both (See Table C.13).
This coordination was in the form of attending statewide or
regional meetings on the handicapped policy, communication
with the regional office and/or other Head Start programs
about specific aspects of the implementation, and organizing
and/or participating in workshops. Level of coordination
effort, as we have measured it here, and implementation
success were not related at all (tau b = 0.16; chi square

si

be

re

me

nc

M

o

p

o

c

s

r

i

significant only at the 0.35 level). This is primarily because all of the programs in the low success group sent representatives to the regional or statewide coordination meetings/workshops even though they were serving very few or no handicapped children.

Mobilization of resources

Implementation of the policy depended, in large part, on arrangement of the special services required by handicapped children through other agencies in the community. Little over one-half of the Head Start directors interviewed indicated that 2 different community agencies were providing specialized services to the handicapped children in their programs. In 3 instances the programs were reportedly receiving services from 4 or more different agencies; and in 7 cases (all in the low success group) the programs were not utilizing any other agencies in the provision of services to handicapped children (See Table C.14). The statistical association between this variable and level of implementation success is very strong (tau c = 0.66). This finding provides empirical support to the view that successful implementation of this policy depends to a large extent upon how effectively the program staff can mobilize the resources available in the community.

Additional funding

The congressional mandate, as we indicated in Chapter IV, was not accompanied with additional funding. This left

the programs with two options: (1) to cover the extra costs
of serving the handicapped through in-kind assistance from
other agencies in the community, and (2) to obtain funds
(from sources other than OCD) which can be used in purchas-
ing the services that cannot be arranged any other way.
This second option involved soliciting donations from sources
such as the grantee agency, charitable organizations, state
or local governmental units, school districts, and other
federally funded human service programs (e.g., Model Cities).

15 of the 35 directors interviewed indicated that they
were able to obtain additional funds for the purpose of pro-
viding services to handicapped children (See Table C.15).
There are statistical differences between the three success
groups (tau c = 0.44; chi square significant at the 0.05
level), although these differences are not as significant as
in the case of mobilization of resources.

Additional staff

Only 14 of the 35 directors (40%) indicated that they
added at least one person to their staff for the primary
purpose of better serving handicapped children. The remain-
ing programs implemented the mandate with their existing
staff resources (See Table C.16). Funding difficulties were
the primary reason cited for not hiring new staff. Almost
all of the program directors in the sample expressed a need
for more staff, especially for persons with one or more of
the following skills: speech therapists, psychologists,

expe

psyc

istr

addi

at

fac

low

and

tiv

Tra
ass

ca

in

of

fo

T&

ti

sh

pr

edu

in

ica

fac

sen

experienced outreach specialists, special education teachers,
psychiatric workers, physical therapists, handicapped admin-
istrators, and medical consultants.

The three success groups differ slightly in terms of
addition of new staff (tau c = 0.43; chi square significant
at the 0.05 level). The differences are mainly due to the
fact that all but one (i.e., 90.0%) of the programs in the
low success group had no new staff (as compared to 50.0%
and 40.0% for the moderate and high success groups respec-
tively).

Training and technical assistance

All but 2 of the program directors interviewed indi-
cated that they or their staff received some form of train-
ing and technical assistance (T&TA) for the primary purpose
of serving handicapped children. In fact, close to three-
fourths of the directors identified 3 or more different
T&TA activities (See Table C.17). Most of the training iden-
tified were in the form of in-service or pre-service work-
shops for the Head Start staff. In a few cases some of the
program staff were taking college-level courses in special
education or related fields.

The three success groups do not differ statistically
in terms of this variable (tau c = 0.24; chi square signif-
icant only at the 0.20 level). This is in part due to the
fact that some of the programs in the low success group were
sending their staff to training sessions although they had

very few or no handicapped children enrolled during the 1973-74 program year.

Modification of physical facilities

Twenty-seven of the thirty-five program directors indicated that they made no changes in their physical facilities for serving handicapped children (See Table C.18). Nine of these programs had no need to modify their facilities because they were quite appropriate for serving severely handicapped children in the first place. Of the 8 programs which modified their physical facilities, 6 were in the high success group and the remaining 2 were in the moderate success group. These figures yield significant statistical differences between the three success groups in terms of modification of facilities (tau c = 0.50; chi square significant at the 0.01 level). The types of changes mentioned by the program directors include building ramps, enlarging passage areas, and adding fixtures to toilets.

Let us now recapitulate what we have done so far in this chapter. We studied the univariate frequency distributions of all of the secondary variables used in constructing the indices for ENVR, PSUP, CAPB, IPOT, and EFRT. In addition, we examined a few selected statistical associations among these variables and investigated the relationships between each secondary variable and implementation success. As a result of these we are now in a position to do two things. First, we can draw the profile of a "typical" Head

Sta.

Sec

str

mer

und

St

sa

Start program in terms of the 18 implementation variables.
Second, we can identify the secondary variables which had the
strongest and the weakest statistical associations with imple-
mentation success. These are the final two tasks we will
undertake in this section.

The following is a description of how a fictitious Head
Start program, which is typical of the 35 programs in the
sample, has gone about implementing the handicapped policy.[1]

> The program was notified about the policy about 3 months
> prior to the start of the 1973-74 program year. The
> notification came from the regional office and was not
> accompanied with detailed guidelines.
>
> The program had some previous experience in serving the
> handicapped. There was at least one severely handi-
> capped child enrolled during the previous year and there
> was at least one person on the program's staff who had
> some college-level training on special education. The
> program's physical facilities were not appropriate for
> serving some of the severely impaired such as the blind
> or the physically handicapped and no one involved with
> the program had any idea about the size of the Head
> Start eligible handicapped population within the
> geographical area served by the program.
>
> The director was in general agreement with the intent of
> the policy but disliked the manner in which it was
> forwarded to the program by the regional office. He also
> had slightly favorable attitudes towards serving the
> severely handicapped. Agencies in the community which
> provide special services to the handicapped were neutral
> to the policy. So were the parents of children already
> enrolled in Head Start. The director believed that his
> program had the capabilities to accommodate perhaps 5
> or 6 types of severe handicapping conditions but not
> the blind, the deaf, the severely mentally retarded, or
> the severely physically impaired.

---

[1]This description is based principally on a review of the
univariate frequency distributions presented above and selection of
the most appropriate statistic that describes the central tendency of
the distribution. In most instances this statistic is the sample mean,
however in some cases we used the median or the mode.

The program staff did not recruit handicapped children
in any different manner than they did nonhandicapped
children. Many of the handicapped children who were
eventually enrolled were referred to Head Start by
community agencies or parents who had heard about the
mandate. After conducting a series of tests and diag-
noses, medical professionals associated with the pro-
gram identified 35 of the 253 children enrolled in
the 1973-74 full-year program as "handicapped". Of
these 7 were considered to be severely handicapped and
the remaining 28 were judged as mildly/moderately hand-
icapped. Slightly over one-third of the 35 children
had speech impairments. There were no totally blind
or totally deaf among the 35.

A person from the staff of the program was assigned the
responsibility to coordinate the implementation. This
person spent about 20% of his time on matters related to
the policy. In addition, an implementation plan was
prepared, which covered 5 of the 8 important aspects of
the implementation. The program's efforts on behalf of
the handicapped were coordinated with both the regional
office and other Head Start programs. No new staff were
added just for the purpose of serving the handicapped,
but the existing staff went through 3 separate pre-
service or in-service training programs. The program
staff were not able to obtain additional funding for the
implementation. However, they were able to obtain in-
kind assistance from 2 community agencies. Finally, the
physical facilities of the program were not modified in
order to make them more suitable for serving the severely
handicapped.

The above profile of a typical program is purely

descriptive and it does not tell us anything about what

could have made the program successful or unsuccessful. In

order to study this we need to examine the statistical asso-

ciation between each of the 18 implementation variables and

level of implementation success.

In the preceding pages we reported the Kendall's tau *b*

(or tau *c*) value as the principal measure of association

between each of the 18 variables and implementation success.

A review of these values indicates that 6 of the 18

variables had very strong statistical relationships, 8 had

moderately strong correlations, and 4 had statistically

insignificant associations with the dependent variable.

These variables and their respective tau b (or tau c)

values are listed below.

Variables Strongly Correlated with Success	Variables Moderately Correlated with Success	Variables not Correlated with Success
EXPR (0.68)	FUND (0.44)	LEAD (0.25)
MOBR (0.66)	SIZE (0.43)	TATA (0.24)
ORGS (0.62)	TARG (0.42)	COOR (0.16)
AGRM (0.56)	KNOW (0.41)	AFAC (-0.14)
CAPB (0.52)	ATTD (0.41)	
MFAC (0.50)	QUAL (0.39)	
	PLAN (0.39)	
	COMM (0.38)	

We do not know exactly why some of these variables are

strongly correlated with success and other are not. However,

we can speculate about the reasons for the presence or com-

plete absence of a strong empirical relationship. The points

we make below summarize our interpretation and analysis of

the magnitudes of the coefficients listed above.

- LEAD is not related to success because the indicator used in the study shows the length of time between notification and start of the program year, not the time period between the point at which work on implementation commenced and the start of the program year.

- TATA and COOR are not related to success because programs engage in these activities more for reasons of visibility (in the eyes of the regional office and other Head Start programs) than for improving the implementation. Also, unsuccessful programs with heavy TATA and COOR activities can always argue that they are getting prepared for the next year.

- AFAC is not related to success because physical barriers are not the real reason for exclusion of the handicapped.

Also, programs usually decide first if they are going to implement/not implement the policy and then, if they decide to implement, investigate the means of removing the barriers to successful implementation.

● EXPR is strongly correlated with success because past behavior is one of the best predictors of present behavior.

● AGRM is strongly related to success because when clear-cut rewards (positive and negative) are not spelled out, the decision to implement or not to implement is largely left to the discretion of the program.

● CAPB is strongly related to success because, under assumptions of rationality, individuals hardly perform the tasks they think they are going to fail in.

● MOBR, ORGS, and MFAC are strongly related to success because these three variables denote real commitment to the implementation, as opposed to artificial commitment (e.g., TATA and COOR) or commitment on paper (e.g., PLAN).

Needless to say, these speculations and interpretations are after-the-fact and they do not collectively point to a distinct and all-inclusive explanatory framework. The conceptual scheme we presented in Chapter III ties together some of these otherwise isolated generalizations and provides a means for a more coherent explanation. We now move to an empirical analysis of this model in light of the Head Start data.

Empirical Comparison of the Alternative Models

In this section we study the empirical support for the conceptual scheme presented in Chapter III. In particular, we compare the empirical evidence for each of the alternative models proposed for explaining LIS and EFRT. Our overall purpose is to identify the specific causal structure which has the greatest empirical support and discuss its implications for studying policy implementation.

The organization of the section is as follows. First we present the set of criteria used in comparing the results of the alternative regression models. This is followed by a comparison of the additive, the interactive, and the hybrid models showing the relationship of IPOT and EFRT to LIS. Next we compare the additive and the interactive models used for the relationship of ENVR, PSUP, and CAPB to EFRT. Finally, we investigate the predictive power of the alternative models by comparing the values of LIS estimated on the basis of different combinations of models used for the two relationships. The last part of the section is devoted to a summary and interpretation of the results.

Criteria for Comparing the
Regression Results

We conducted the comparison of the regression models in two stages. In the first stage we investigated the statistical significance of each of the five models described in Chapter III. In this investigation we paid particular attention to two issues. First we tested each regression equation for overall goodness of fit using the F test. Second, we tested each regression coefficient associated with the independent variables included in the model. In all five cases the F values were significant at the 0.01 level, but in two of the models the regression coefficients associated with some of the independent variables turned out to be insignificant at the 0.05 level. In these instances we dropped the insignificant variables from the regression

equation one at a time until we were left with coefficients which were all significant at the 0.05 level. Thus, the first stage in the comparison of the models involved the identification of the specific regression equations which were all statistically significant and which included only those variables with regression coefficients significantly different from zero.

In the second stage we compared the alternative models in terms of a set of criteria. These criteria include the following:

- overall explanatory power of the model

- appropriateness of the form of the functional relationships used in the model

- randomness (or independence) of the residuals

- homoscedasticity of the residuals

- simplicity of the model

The most commonly used measure of the *explanatory power* of a regression model is R^2, the square of the multiple correlation coefficient. R^2, which shows the proportion of the variation in the dependent variable that is explained by all the variables included in the regression equation, is an excellent indicator of the overall accuracy of the model and we have chosen it as our first criterion in comparing the alternative models.

The alternative models we are studying here are all linear models. It is possible, of course, that the form of one or more of these models may be inappropriate for

explaining LIS or EFRT with the data we have at hand. In
such cases it may be necessary to add new explanatory
variables or work with nonlinear model in order to improve
the *appropriateness* of the form of the initial functional
relationship. Although there are no clear-cut rules for
determining the adequacy of the form of a given model, the
scatter of the residuals usually indicates if there is a
need for adding polynomial or other terms or transforming
the original variables. Therefore, our second criterion
relates to the appropriateness of the form of a given model
in light of the pattern of the residuals when plotted against
the estimated values of the dependent variable.

Our third criterion concerns the *randomness of the resid-
uals*. Here we are interested in finding out whether one of
the basic assumptions of the general linear model, i.e.,
independence of the disturbance terms, is satistified. In
cross-sectional studies non-randomness of the residuals is
usually due to omission of a key variable, incorrect speci-
fication of the form of the relationship, or errors in mea-
suring the dependent variable. In order to test for random-
ness of the residuals we ordered the cases in the sample in
terms of the estimated value of the dependent variable and
performed an analysis of runs of signs of the residuals.
According to this test the actual number of runs of signs
of the residuals is compared against the number of runs that
would be expected under the assumption of randomness of the

sign of each residual.[1] If the expected and the actual number
of runs differ from each other significantly, there is reason
to believe that the residuals are not random. Otherwise,
one has to conclude that the model at hand satisfies the
assumption about the independence of the disturbance term.

Homoscedasticity of the residuals, our fourth criterion,
concerns another important assumption of the linear regres-
sion model. This time the variance of the disturbance term
is at issue. The general linear model is based on the assump-
tion that the disturbance term has constant variance. When
this is violated it may be necessary to transform the depen-
dent variable or use generalized least squares instead of
simple least squares.[2] In the case of the models examined
here, our test for homoscedasticity is only through visual
inspection of the scatter of the residuals.

Our last criterion, *simplicity,* refers to the complex-
ity of the regression equation. In this respect, we examined
two aspects of each equation: (1) the number of independent
variables, and (2) the nature of the independent variables.
More specifically, when comparing two models, we regarded
the one with fewer explanatory variables to be "simpler"
than the other. With respect to the nature of the indepen-
dent variables, we viewed the model based on an additive

---

[1]See: Sidney Siegel, <u>Nonparametric Statistics</u> (New York: McGraw-
Hill, 1956), pp. 52-58.

[2]See: J. Johnson, <u>Econometric Methods</u>, 2nd Edition (New York:
McGraw-Hill, 1972), pp. 208-221.

relationship to be simpler than one which is based on an interaction term. Thus, of the following two relationships $Y = a + bX + cZ$ and $Y = a + bXZ$ we regard the former to be simpler because, although both equations are based on two independent variables, it is based on an additive model.

These five criteria do not exhaust all possible yard-sticks that could be used in comparing the alternative models. Furthermore, a decision based only on these five criteria alone cannot still be regarded as conclusive because of the smallness of our sample and because of the unknown reliability of the indices. Moreover, the empirical results need to be reviewed in light of the substantive support for each model. Therefore, the final decisions resulting from comparisons of the models should be based on statistical as well as theoretical grounds.

<div style="text-align:center">

Relationship of IPOT and EFRT
with LIS

</div>

In Chapter III we formulated and discussed three alter-native linear models for studying the relationship of IPOT and EFRT with LIS. Here we examine the empirical support for these three models.

In the first stage of our comparison of the three models we used the stepwise regression routine with a very low F value for inclusion and deletion of the variables. This permitted us to examine the order in which the variables entered the equation and the change in R^2 and overall F at every step of the process. The resulting equations,

along with the significance of each regression coefficient (numbers in parantheses), are shown below. The correlation matrix which led to these equations is displayed in Table 21.

TABLE 21

CORRELATION MATRIX FOR THE
LIS EQUATIONS

	IPOT	EFRT	IPOTxEFRT
LIS	0.65	0.78	0.85
IPOT		0.45	0.85
EFRT			0.79

The additive model:

$$LIS = -7.14 + 0.36 \ (EFRT) + 0.17 \ (IPOT) \tag{6}$$
$$(0.01) \qquad (0.01)$$

The interactive model:

$$LIS = 4.70 + 0.49 \ (EFRTxIPOT) \tag{7}$$
$$(0.01)$$

The hybrid model:

$$LIS = 2.84 + 0.11 \ (EFRT) - 0.06 \ (IPOT) + 0.47 \ (EFRTxIPOT) \tag{8}$$
$$(0.35) \qquad (0.59) \qquad (0.03)$$

The overall F values for each of the three equations is significant at the 0.001 level. However, in the hybrid model the regression coefficients for both IPOT and EFRT were statistically insignificant at the 0.05 level. This implies that the addition of either EFRT or IPOT to the equation which contains the interaction term IPOT x EFRT does not significantly contribute to the explanation of

LIS. This is also confirmed by the stepwise results for the hybrid model. After IPOT x EFRT entered the equation the partial correlation coefficients for IPOT and EFRT were both insignificant at the 0.05 level. This result could have been predicted after examining the correlation matrix. The correlation coefficients showing the statistical association between IPOT x EFRT and IPOT and between IPOT x EFRT and EFRT are very high. This implies that once IPOT x EFRT is in the equation it is doubtful that either IPOT or EFRT will add anything significant to the explanation of LIS.

As a result of these findings we dropped the hybrid model from our pool of alternative models. Without IPOT or EFRT in the equation, this model becomes identical with the interactive model. Therefore, we deal only with the additive and the interactive models in the second stage of the comparison.

The following is a comparison of the two models in terms of the five criteria introduced earlier. A summary of this comparison can be found in Table 22.

- **Explanatory power**.--The additive and the interactive model yielded identical results in terms R^2. Standard errors of of the two estimates, expressed as a percentage of the mean of LIS, were 39.5% and 38.2% for the additive and the interactive model, respectively. The high R^2 figures (0.72 and 0.73) confirm our earlier expectation that IPOT and EFRT are the most important determinants of LIS.

- **Appropriateness of the form of the model**.--The two models again yielded similar results in terms of this criterion. Plots of the residuals against the values of LIS estimated by the additive and the interactive model (Figures 2 and 3) show similar spreads. Both models significantly underpredicted the LIS value of the most successful program (the residuals for this program were beyond the 2 standard
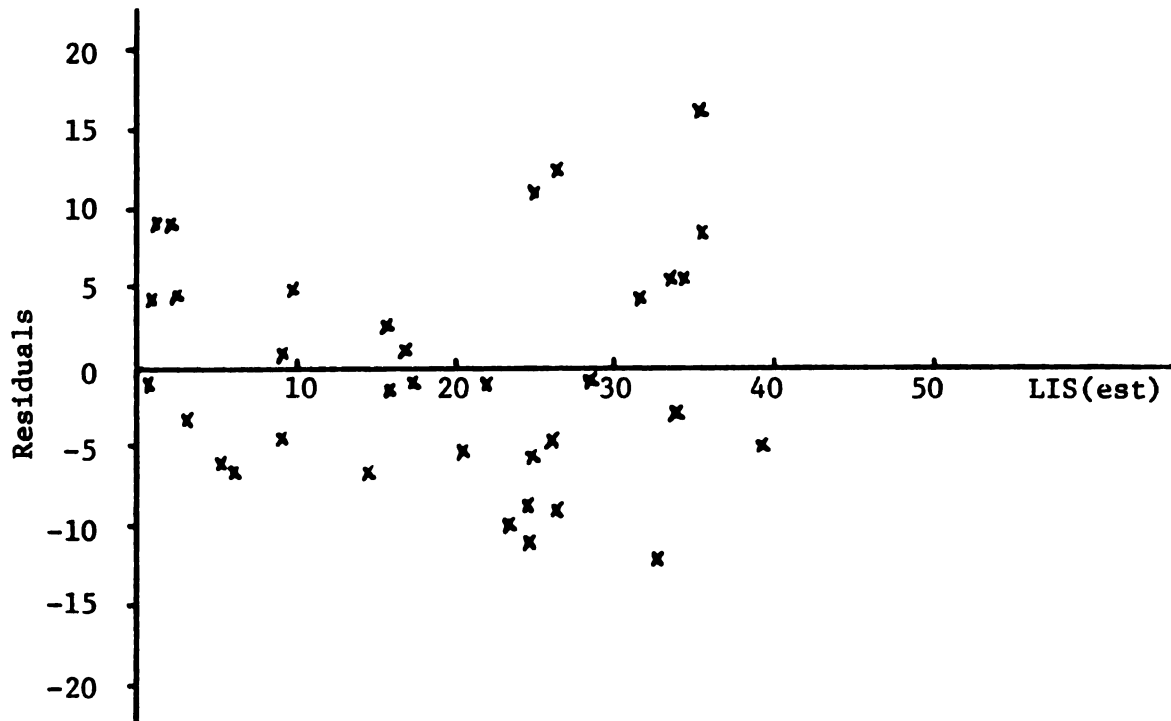
Figure 2. Scatter of the residuals from the estimated values
of LIS--the additive model
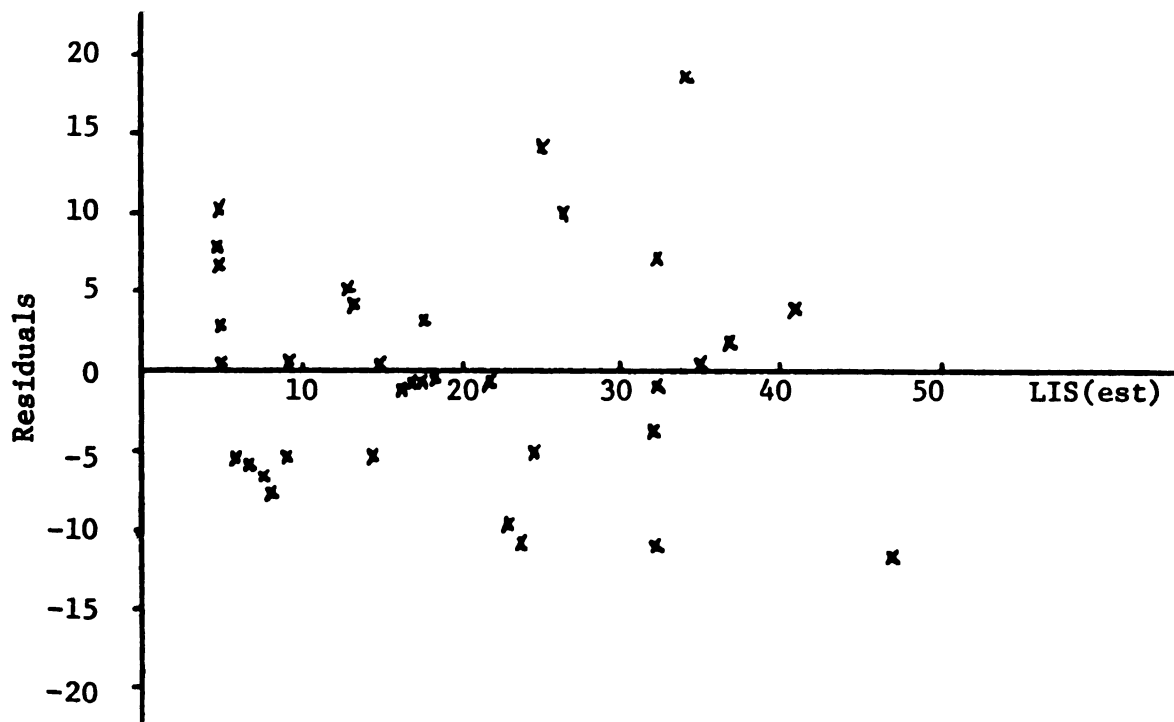
Figure 3. Scatter of the residuals from the estimated values
of LIS--the interactive model

TABLE 22

COMPARISON OF THE ADDITIVE AND THE INTERACTIVE
MODELS--EXPLANATION OF LIS

Criteria	Additive Model	Interactive Model
Explanatory power	$R^2 = 0.72$	$R^2 = 0.73$
Appropriateness of the form of the model	Appropriate	Appropriate
Randomness of the residuals	Random	Nonrandom
Homoscedasticity of the residuals	Slightly heteroscedastic	Slightly heteroscedastic
Simplicity	2 variables; simple model	2 variables; complex model

deviation limit in both cases). Also, both models exhibited
inconsistencies in predicting the higher values of LIS.
Given the high R^2 values, we believe that this is due to
heteroscedasticity rather than the need for adding new terms
to either model.

● Randomness of the residuals.--The expected number of signs
of runs for 35 observations (under the null hypothesis of
randomness) is 18. The additive model yielded 19 runs, the
interactive model yielded 12. The difference between the
expected and the observed number of runs is statistically
significant for the interactive model but not for the
additive model (at the 0.05 level). This implies that there
is statistical evidence against the randomness of the resid-
uals obtained from the interactive model.

● Homoscedasticity of the residuals.--In both cases there is
evidence of some heteroscedasticity. Visual inspection
of the scatter of the residual indicates that the variance
of LIS is larger for higher values of LIS (there are more
residuals beyond the ± 1 standard deviation limit for the
high values of LIS than for the low values). However, the
two plots did not differ from one another significantly
to warrant the conclusion that one model is more
heteroscedastic than the other.

- Simplicity.--Both models are essentially based on the same two explanatory variables. However, since the additive model is relatively simpler than the interactive model we have to conclude that the former is simpler overall.

Clearly, both of these models are quite good in predicting LIS. The five criteria we are using here slightly favor the additive model because of its simplicity and the apparent randomness of its residuals. Why did the two models yield similar results? One explanation is that both IPOT and EFRT are strong predictors of LIS (the Pearson correlation coefficient between IPOT and LIS is 0.65; between EFRT and LIS it is 0.78). In addition, these two explanatory variables are correlated with each other--although not too strongly--which implies that programs with low IPOT values in general also have low EFRT values and regardless of whether we multiply or add IPOT and EFRT we still predict a low LIS. The same argument applies to the other extreme of the IPOT and EFRT indices. We have kept IPOT in the equation for the additive model despite its correlation with EFRT (r = 0.45). The reason for this is that addition of IPOT to the equation containing only EFRT increases the R^2 by 0.11 (from 0.61 to 0.72) and this increment in explanation of LIS is statistically significant at the 0.01 level.

Relationship of ENVR, PSUP, and CAPB with EFRT

In Chapter III we formulated and discussed two alternative linear models for the relationship of ENVR, PSUP, and CAPB with EFRT. Again, one of these was an additive and the

other was an interactive model. The correlation matrix for this relationship is shown in Table 23. As a result of stepwise regression we obtained the following equations (the numbers in parentheses again show the significance level of each regression coefficient):

The additive model:

$$EFRT = 10.61 + 0.25 \ (ENVR) + 0.38 \ (PSUP) + 0.13 \ (CAPB) \ (9)$$
$$(0.15) \qquad\qquad (0.05) \qquad\qquad (0.28)$$

The interactive model:

$$EFRT = 15.54 + 0.34 \ (ENVR) + 0.34 \ (PSUP \ x \ CAPB) \qquad (10)$$
$$(0.045) \qquad\qquad (0.01)$$

TABLE 23

CORRELATION MATRIX FOR THE
EFRT EQUATIONS

	ENVR	PSUP	CAPB	PSUP x CAPB
EFRT	0.49	0.63	0.52	0.56
ENVR		0.50	0.33	0.41
PSUP			0.66	0.78
CAPB				0.95

The overall F values for each of these two equations is significant at the 0.01 level. The two regression coefficients in the interactive model are also significant at the 0.05 level. The coefficients for ENVR and CAPB in the additive model, however, are not significant. In light of this finding, we ran the regression routine for the additive model

by forcing the variables into the equation one at a time and
by changing the order of insertion of the variables (i.e.,
a total of six separate regression runs). The results indi-
cated that each time PSUP, the variable with the highest
correlation with EFRT, was in the equation the coefficients
for either of the other two variables became insignificant
(at the 0.05 level). On the basis of this finding we
dropped both ENVR and CAPB from the additive model and
obtained the following as our revised additive model:

$$EFRT = 10.20 + 0.62 \text{ (PSUP)} \qquad (11)$$
$$(0.01)$$

Next, we compared the revised additive model with the
interactive model in terms of the five criteria. Table 24
shows a summary of this comparison. Discussion of the two
models in terms of the five criteria follows.

- Explanatory power.--The two models yielded identical results.
Percentage of variation in EFRT explained by each of the
two models is not as high as in the relationship of EFRT and
IPOT with LIS (40% and 39%, respectively, for the additive
and the interactive model). The standard errors, expressed
as percentage of the mean of EFRT, were 38.0% and 42.7% for
the additive and the interactive model, respectively.

- Appropriateness of the form of the model.--The scatters of
the residuals plotted against the values of EFRT estimated
by the two models are shown in Figures 4 and 5. The two
spreads are similar but not identical. The additive model
significantly overestimated the EFRT value of a program
which had a very low actual EFRT score. Alternatively, the
interactive model underpredicted the EFRT value of a pro-
gram with a relatively high EFRT score. Neither of the
two plots indicates a definite need for revising either
model (i.e., by adding a polynomial or an interaction
term).

- Randomness of the residuals.--The actual number of runs of
signs of residuals for the additive model is 21; for the
interactive model it is 15. Given that the expected number
of runs (under the assumption of randomness) is 18, both
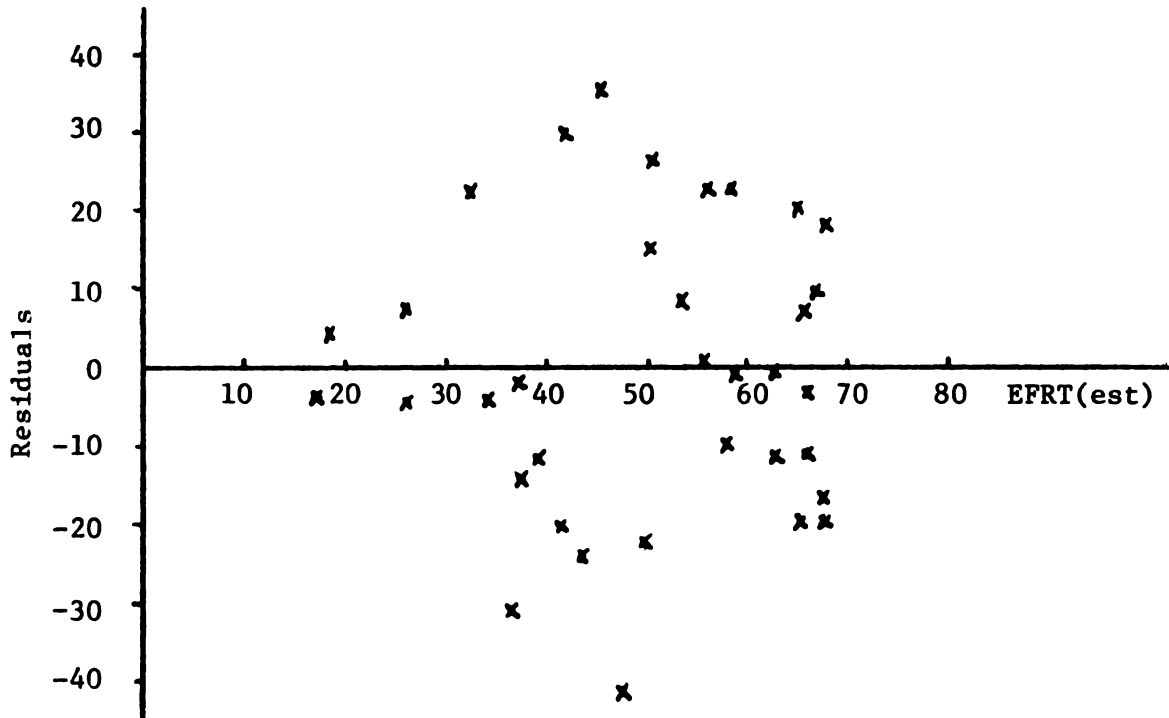
Figure 4.   Scatter of the residuals from the estimated values
of EFRT--the revised additive model
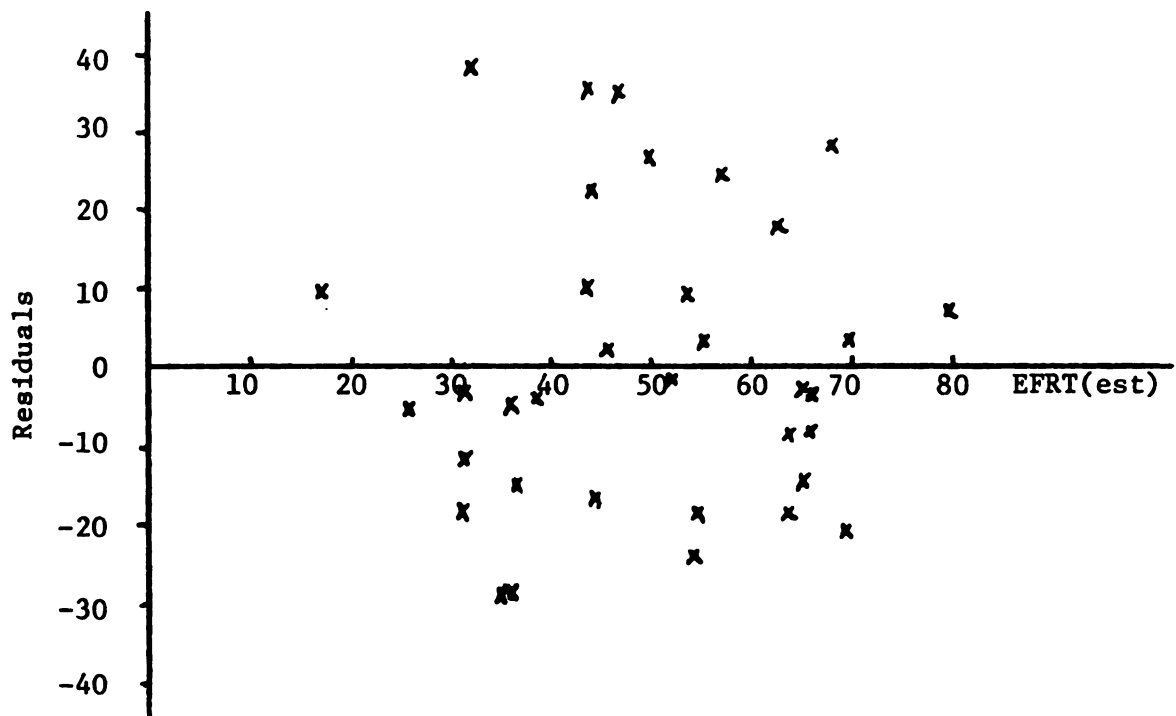
Figure 5.   Scatter of the residuals from the estimated values
of EFRT--the interactive model

TABLE 24

COMPARISON OF THE ADDITIVE AND THE INTERACTIVE
MODELS--EXPLANATION OF EFRT

Criteria	Additive Model (Revised)	Interactive Model
Explanatory power	$R^2 = 0.40$	$R^2 = 0.39$
Appropriateness of the form of the model	Appropriate	Appropriate
Randomness of the residuals	Random	Random
Homoscedasticity of the residuals	Homoscedastic	Homoscedastic
Simplicity	1 variable; simple model	3 variables; complex model

of these figures are within the 95% confidence interval.
Thus, we have no statistical reason to believe that the
residuals for either model are nonrandom.

● Homoscedasticity of the residuals.--Visual inspection of
the two plots of residuals indicates that the disturbance
terms are quite homoscedastic. Both models predict high
EFRT values with consistency. Absolute values of the
residuals are slightly larger for moderate and low effort
scores. We do not think that these observations warrant the
conclusion that the disturbance term of either model is
heteroscedastic.

● Simplicity.--The revised additive model is clearly superior
in terms of this criterion.

The five criteria we are using here slightly favor the

revised additive model because of its simplicity. With this

model we are able to explain 40.0% of the variation in EFRT

with a single explanatory variable. The interactive model

reaches the same level of explanation with three variables, two of which are combined in the form of an interaction term.

We do not have a powerful explanation of the finding that the two models yield practically the same R^2 figure. An examination of the correlation matrix indicates that PSUP, alone, has greater explanatory power than PSUP x CAPB (Pearson correlations between these two terms and EFRT are 0.63 and 0.56, respectively). Therefore, the interaction term does not add anything to the explanation one would not be able to obtain with PSUP alone. The distribution of PSUP x CAPB is almost identical with that of CAPB (r = 0.95) and since the correlation of CAPB with EFRT is smaller than the correlation of PSUP with EFRT (0.52 vs. 0.63), it is reasonable to expect a PSUP x CAPB to have weaker explanatory power than PSUP alone.

Predictive Power of Combinations of Models

The data from the 35 Head Start programs in our sample did not permit a clear choice between the additive and the interactive models in terms of the R^2 criterion. In view of this, we decided to conduct one final test, this time examining the predictive power of combinations of models. Since we have two alternative models for explaining LIS, in both of which EFRT is an explanatory variable, and two additional models for explaining EFRT, we have a total of four alternative model combinations.

	EFRT equation	LIS equation
Alternative AA	Additive	Additive
Alternative AI	Additive	Interactive
Alternative IA	Interactive	Additive
Alternative II	Interactive	Interactive

The computational aspects of the alternative model combinations are quite straightforward. In the "additive-additive" combination for example, we first obtained an estimate of EFRT using the revised additive model. We then used these estimated EFRT values, along with the original IPOT scores, to estimate LIS through the additive model for LIS. Calculation of LIS estimates based on the other three alternatives were made in a similar manner.

Each of the four alternative combinations has an interesting feature. The variables used in obtaining the alternative estimates of LIS are either PSUP and IPOT or PSUP, CAPB, ENVR, and IPOT. Since all of these variables can be measured at a point in time immediately prior to the start of the implementation, they make it feasible for us to predict LIS before any implementation activity takes place. Therefore, comparison of the four model combinations not only permits examination of the explanatory power of each combination, it also allows us to study the real predictive ability of these models.

Operationally, we first obtained a set of predicted LIS scores for each model combination. Next, we computed the

deviations of the predicted scores from the actual values

of LIS. We then used the four sets of deviations to cal-

culate three statistics: the average deviation (per pro-

gram), the average absolute deviation (disregarding the sign

of the deviations), and the average squared deviation (sum

of the squares of the deviations per program). The results

we obtained are displayed in Table 25.

TABLE 25

COMPARISON OF THE PREDICTIVE ABILITY OF
ALTERNATIVE MODEL COMBINATIONS

Model Combination	Average Deviation	Average Absolute Deviation	Average Squared Deviation
Alternative AA	-0.04	7.07	81.39
Alternative AI	0.73	6.93	78.65
Alternative IA	-0.05	7.24	78.11
Alternative II	0.67	7.37	81.28

The four model combinations studied do not differ from

each other significantly in terms of any of the three sta-

tistics computed. The average deviations of the "additive-

additive" and the "interactive-additive" combinations are

considerably smaller than those of the other two combina-

tions. In terms of average absolute deviation the "additive-

additive" and the "additive-interactive" combinations yield

the lowest values (the 7.07 average absolute deviation

figure corresponds to 37% of the mean of the actual LIS

scores). Finally, the average squared deviations of the

"additive-interactive" and the "interactive-additive" com-

binations are slightly lower than those of the other two.

However, the statistical differences between the model com-

binations are not large enough to warrant the selection of

any of the four model combinations as the one with the

highest predictive ability. If anything at all, the figures

shown in Table 25 perhaps eliminate the "interactive-inter-

active" combination from any further consideration.

Summary of the Analytical Findings

In the preceding parts of this section we primarily

dealt with empirical validation and comparison of the alter-

native models formulated in Chapter III. Here we conclude

the present chapter by reviewing the most important of our

major analytical findings and discussing their implications.

First, we showed that there is strong empirical support

for the overall conceptual model presented in Chapter III.

LIS is indeed a function of EFRT and IPOT and EFRT can be

explained empirically by ENVR, CAPB, and, in particular,

PSUP. Most of our hypotheses regarding these relationships

were confirmed by the Head Start data. We found that

(1) controling for IPOT, LIS increases as EFRT increases,

(2) controling for EFRT, LIS increases as IPOT increases,

(3) controling for ENVR or CAPB or both, EFRT increases as

PSUP increases, and (4) controling for ENVR, EFRT increases

as PSUP x CAPB increases. We disconfirmed the hypothesis that, controling for PSUP, EFRT increases as either ENVR or CAPB (or both) increase.

Second, we found that there is empirical support for both the additive and the interactive models proposed for the two relationships. We eliminated the hybrid model from further consideration and revised the additive model proposed for explaining EFRT. Although the closeness of the R^2 values of the additive and the interactive models in both instances were surprising, we were able to select the two additive models on the basis of the other four criteria.

Third, our attempt to choose between the four model combinations on the basis of their predictive ability failed to single out one of the explanatory paths as the "best". However, we were able to show that the "interactive-interactive" combination has relatively less predictive power than the other three.

Perhaps our most significant empirical finding in this section is that prediction of LIS on the basis of the two additive models requires data on only two explanatory variables: PSUP and IPOT. Since the two additive models overall were judged to be more powerful than the other model combinations, on the basis of the data examined here, we are in a position to predict the implementation success of the handicapped policy in a given Head Start program by obtaining an objective measure of the program's pre-implementation capabilities and an estimate of the director's

support of the policy before the start of the implementation. We believe that this is quite significant because it suggests two distinct strategies which can be followed by OCD for improving the overall success of the implementation nationwide.

The first strategy suggested by the "additive-additive" model combination is that OCD could carry out an objective assessment of the implementation capabilities of each Head Start program and formulate and execute a comprehensive action plan aimed at the installation/improvement of the capabilities missing at the local level.

The second strategy implied by the two selected models is directed at increasing PSUP. There is considerable evidence to suggest that OCD did not do a good job on "selling" the handicapped policy to the local programs. The data we studied indicates that support of the policy by the directorate of the local program is the most important determinant of the level of effort expended during the implementation. Many of the programs which were "turned off" by the manner OCD passed the policy down to the grantee level could have been "turned on" if OCD had demonstrated the significance of the policy and, in particular, provided better guidelines for the implementation.

CHAPTER VI

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

We started this study by observing that (1) implementation has rarely been studied and (2) there are no empirically verified theories about why some public policies are implemented successfully and others are not. We then formulated and tested a conceptual scheme which explains the implementation success of a single policy in 35 Head Start programs. We are now in a position to outline and discuss our major conclusions regarding, first, the implementation of the handicapped policy and, second, the study of policy implementation.

Handicapped Policy in Head Start

We do not know exactly why Head Start was asked to implement the handicapped policy. We suspect that OCD played an important part in the formulation of such a requirement. The issue that is not clear is why OCD wanted its "star" program to get involved with handicapped children in the first place.

Our understanding is that OCD had three reasons for favoring such a policy. First, given Head Start's prominent role in its overall program, OCD wanted to strengthen its

209

own position by increasing Head Start's chances of survival.
In an era when most of the old OEO programs were being
phased out, principals in OCD were justifiably worried about
the fate of Head Start, perhaps the most glamorous of all
previously OEO-sponsored programs. Recognizing that a
policy such as this one would reconfirm congressional com-
mitment to Head Start, OCD pushed for the adoption of the
handicapped policy.

Second, the handicapped policy made good sense in light
of OCD's ambition and long-term goal to become the single
federal agency with authority over all programs concerning
preschool children. Given that no federal agency had overall
authority over the programs for preschool handicapped, OCD
saw the policy as an opportunity to gain a strong foothold
in the federal "market" for preschool handicapped and a
stepping-stone for gaining broader authority for all handi-
capped and non-handicapped preschool children.

Third, the need for a nationwide program for preschool
handicapped had long been recognized. The special needs of
the preschool handicapped, especially those from disadvan-
taged families, were not being met and OCD recognized this
need for programs directed towards improving the lives of
these children and their parents. Since Head Start was the
largest comprehensive preschool program in the nation, it
was logical for OCD to extend Head Start's services to

include the handicapped.[1]

Thus, the handicapped policy was formulated and OCD began to take actions for implementing it nationwide. Was the first full-year implementation of the policy a success? The answer to this question depends on the way one interprets the policy. If one counts the children with only minor disabilities, such as those with visual or hearing problems which can be corrected with glasses or hearing aids, as handicapped the answer is "yes." Alternatively, if only the moderately or severely disabled are counted, Head Start has not met the ten per cent mandate.

OCD's interest in the actual percentage of handicapped in Head Start is understandable. The policy requires the Secretary of HEW to report back to the Congress on the numbers and disabling conditions of handicapped children enrolled in Head Start. Our interest, however, is in finding out why the programs enrolled the handicapped children they did. The question of numbers is purely a factual and a measurement issue. The "why" issue is more analytical/theoretical and the overall success of the policy can be improved only if satisfactory answers can be found to this issue.

Why were some programs successful in implementing the policy and others were not? We simplified this question by first replacing the term "success" with "effort": Why did

---

[1]We do not have any empirical evidence to support any of these claims. Therefore, they should be viewed mainly as our personal views and impressions on this issue.

some programs put a greater level of effort to the implementation and others did not? We know that capability differences among programs would account for a portion of the variation in success. The rest, we thought, would be mainly explained by level of effort and the data confirmed our expectations.

Why, then, did some programs put a greater level of effort to the implementation than others? The simplest explanation we can come up with is that programs in which the director supported the policy put a greater level of effort than those where the director did not support the policy.

The next obvious question, of course, is: Why did some directors support the policy and others did not? We found that the directors who strongly agreed with the policy in the first place and who had favorable attitudes towards the severely handicapped strongly supported the policy.

The next level of questions concern agreement with the policy and attitudes towards the severely handicapped. We do not know why some directors had positive and others negative attitudes towards serving the severely handicapped. We found only that the program's past experience with the severely handicapped did not affect the attitudes of the director. With respect to the directors' agreement with the policy, however, we found that the disagreements were mainly due to the manner in which the policy was operationalized and passed down to the local level. Lack of clear

definitions for the handicapping conditions and unavail-
ability of specific guidelines were two of the important
determinants of level of agreement. Two other factors that
were statistically associated with agreement were the direc-
tor's perception of (1) the program's capabilities to imple-
ment the mandate and (2) the level of support he can get from
the agencies in the community. Attitudes of the director
towards the severely handicapped were not strongly associated
with his agreement with the policy.

We have now reached the end of the causal chain. The
data we have at hand confirm the hypothesis that, holding
the implementation capabilities of the agencies constant,
increases in policy support will indirectly result in
improvement of implementation success through increases in
the level of effort. Assuming that we could not alter the
attitudes of program directors over the short-run, increases
in policy support can be generated through changes in the
director's level of agreement with the policy. Such changes
can be--or could have been--induced by:

- better and more frequent communication between OCD and the
  local programs about the policy and its implementation;

- involvement of representatives of local programs in the
  policy operationalization process;

- specification of how local programs can more objectively
  assess their own capabilities to implement the policy;

- preparation of comprehensive guidelines and "how-to-do"
  manuals about the implementation;

- clear specification of positive/negative rewards for success-
  ful/unsuccessful implementation;

- nationwide promotion of the policy to the agencies which provide services to the handicapped;

- provision of greater levels of training and technical assistance to the local program staff.

These are the actions that can lead to implementation success over the short-run. In the long-run, however, we cannot assume that neither the capabilities nor the attitudes will remain constant. Capabilities can be improved through capacity-building activities of OCD, or the local programs themselves, or both. Attitudes could change because of increased knowledge about and continuous involvement with the handicapped. In any event, OCD could have and still can increase the implementation success of the policy over both the short- and the long-run.

Study of Policy Implementation

Implementation cannot be studied in a complete vacuum. It is a component of the overall policy process and, as such, it should be studied in connection with the other components of the process, in particular with operationalization and evaluation.

We demonstrated how poor operationalization can lead to subsequent failures in implementation. Implementors expect a lot from those who operationalize the policy and operationalizers assume too many things about the implementors. The only ways this gap can be bridged are through (1) having one and the same agency both operationalize and implement the policy and (2) participation of the

implementors in the decision-making process that goes on
during the implementation. On the one hand it facilitates
the flow of information (especially about capabilities,
perceptions, attitudes, and expectation) from the implemen-
tors to the operationalizers and vice versa. On the other
hand it increases the implementors' level of commitment to
the policy. Implementors can more readily commit themselves
to the decisions which have been made in their presence and
with their participation.

Connections between implementation and evaluation are
also important. The results of implementation are a set of
"policy outcomes" which, if not translated into "impacts"
do not show how much difference the policy has made in the
lives of people.[1] These impact assessments, at least in
theory, can be periodically fed back to the operational-
izers and the implementors for determining what changes, if
any, should be made in the goals or the processes of imple-
mentation.

With respect to implementation success we showed that
effort, potential, and policy support are its most important
determinants. All three of these factors are policy-specific,
which necessitates separate studies of implementation for
every single policy of interest. In particular, this implies
that knowledge of the overall organizational climate and

---

[1]We are using the term "people" instead of "members of the target
group" because a policy can have unintended effects on persons outside
the target group. A good example of this is the effects of the handi-
capped policy on the non-handicapped children enrolled in Head Start.

structure in an organization do not mean much in predicting implementation success. Different policies may be implemented with different levels of success within the same organization. This also implies that public administration and organizational behavior should be viewed as separate from policy implementation. These two fields generally address questions of overall performance in an organization. Policy implementation, on the other hand, deals with policy-specific and task-specific performance.

The three key explanatory variables, EFRT, IPOT, and PSUP, are quite easy to define but difficult to measure. Our concept of effort covers policy-specific and goal-directed activities. When the implementing agency is responsible for simultaneously executing several policies, sorting out all activities of the agency to identify those connected with a given policy is a quite complicated matter. Similarly, determination of the factors that should be considered in evaluating the policy-specific capabilities of the agency also introduces complex measurement problems. Finally, although we measured it only in terms of the director's level of support of the policy, policy support should be viewed as a concept referring to the power structure of the agency. Therefore its measurement may require, first, identification of all persons who have affected or can affect the important decisions made by the agency, and, second, assessment of the extent to which each of these individuals supports the policy in question. These measurement difficulties were perhaps

some of the most important reasons for lack of previous empirical work on implementation.

The final issue we need to address is the degree to which our empirical findings are generalizable to other settings. We believe that the overall conceptual model is quite generalizable, particularly due to the fact that none of the variables employed are generic only to the handicapped policy in Head Start. Our opinion is that the overall findings of the study are more applicable to settings in which (1) the policy studied is a marginal policy, (2) operationalization and implementation were carried out by separate agencies, (3) the policy is a human services-type policy implemented by more than one agency, and (4) the positive/negative rewards for successful/unsuccessful implementation were not spelled out clearly.

Recommendations

The two most important limitations of the study, in our opinion, are (1) the smallness of the size of the sample and (2) unavailability of more valid indicators for measuring especially IPOT and EFRT. Our first recommendation, therefore, is the replication of this study with a larger sample and with more reliable and valid indicators.

Secondly, we believe that the conceptual model we used in this dissertation can be used to investigate implementation success in cases where there are multiple implementation goals. Another interesting topic is the study of

the simultaneous implementation of several policies executed by the same agency.

Finally, we believe that empirical testing of a dynamic version of our conceptual model may lead to more powerful findings concerning policy implementation. In the remainder of this chapter we describe one such model.

A Dynamic Model of Policy Implementation

The conceptual model presented in Chapter III was formulated for explaining implementation success and level of effort during a single period of implementation. As it stands now, four of the variables--ENVR, PSUP, CAPB, and IPOT--refer to a single point in time which corresponds roughly to the start of the implementation. One variable, EFRT, refers to the time period between the start and the end of the period. The final variable, LIS, refers to the results attained at the end of the first period.

Implementation of most public policies extend beyond the first period. The period of implementation is usually a year and a year is usually insufficient for generating all the impacts the policy was originally designed for. It is therefore important to examine what may happen to the implementation of a policy beyond the first year.

Our views on a simple extension of the conceptual model are illustrated in Figure 6. This is a dynamic version of the static model we presented earlier. It includes only one additional variable, REWD, which refers to the rewards to
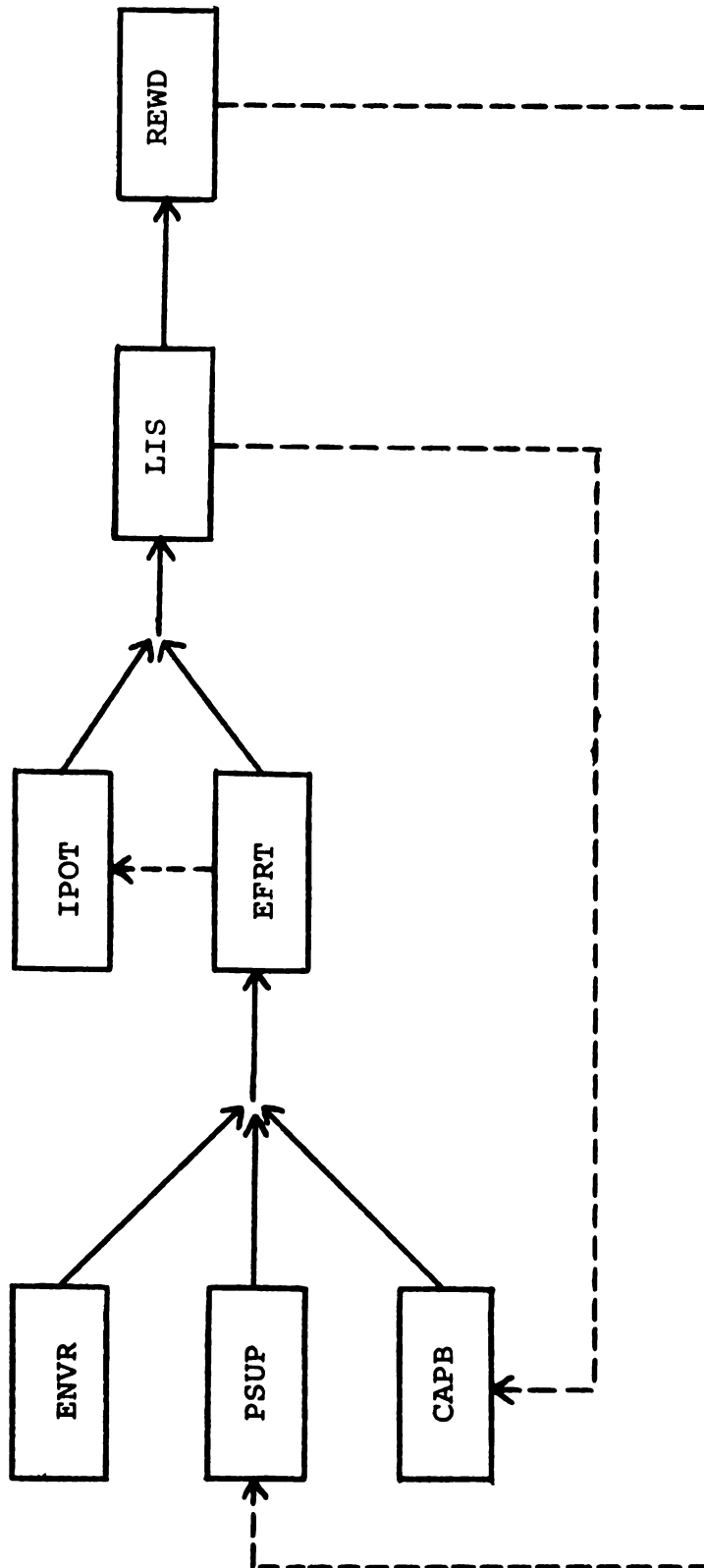
Figure 6. An extension of the conceptual model

the agency resulting from the success level attained at the end of the implementation period. The dotted lines show lagged relationships between the variables of the model. Given a set of initial values for ENVR, PSUP, CAPB, and IPOT the complete model in equation form can be stated in terms of the following six functions:

$$PSUP(t) = f[PSUP(t-1), REWD(t-1)]$$

$$CAPB(t) = f[CAPB(t-1), LIS(t-1)]$$

$$IPOT(t) = f[IPOT(t-1), EFRT(t-1)]$$

$$EFRT(t) = f[ENVR(t), PSUP(t), CAPB(t)]$$

$$LIS(t) = f[IPOT(t), EFRT(t)]$$

$$REWD(t) = f[LIS(t)]$$

where (t) and (t-1) denote successive implementation periods.

According to this model rewards are viewed solely as a function of the level of success achieved during the implementation. We realize that variables other than LIS, such as those depicting one or more characteristics of the reward allocation mechanism in operation, could affect these rewards. Thus, depending upon the policy being examined, it may be appropriate to add these to the model as exogenous variables affecting REWD.

The model includes three time-lagged relationships. First, it is hypothesized that the success level attained during a particular year influences the leadership group's perception of the agency's capability to implement the policy before the start of the following year. Second, it is claimed that actual rewards received at the end of a specific

implementation period affects the leaders' level of support of the policy during the following period. Finally, the model postulates that implementation efforts during a given period modify the potential of the agency to implement the policy during the next period.

Needless to say, the model requires further conceptualization before it can be tested empirically. We believe that further research along the lines of this model would prove to be quite fruitful in understanding and improving policy implementation.

APPENDICES

APPENDIX A

METHODOLOGY AND PROCEDURES

In this appendix we describe in detail the specific methodology and the procedures used in generating the data analyzed in the body of the dissertation. This description is organized along the following lines:

- research design used
- questionnaire design
- selection of programs for the second round visits
- pre-fieldwork activities
- fieldwork and data collection
- post-fieldwork activies.

Research Design

We have basically used a cross-sectional survey design for the second round visits. The data were collected at approximately one point in time (April-May, 1974) and the selected sample was intended to represent the larger population of all Head Start programs at that point in time. The period of data collection corresponded roughly to the final two months of full-year program operations.

The cross-sectional survey design, although it is mostly used to arrive at generalizations for the point in

222

time at which the data were collected, can also be used to
generate longitudinal data. This is accomplished either by
asking the respondent to report information about an earlier
point in time or by having the respondent describe a process
over time. In the case of the second round visits both of
these methods were used to arrive at estimates of longitu-
dinal data. In the first instance, the respondents to the
second round interviews were asked to provide us with data
about the previous year's enrollment of handicapped children,
the specific month they were notified of the new policy, and
their estimate of the implementing agency's capability to
execute the mandate. Secondly, a significant portion of the
interviews concentrated on collection of data about the spe-
cific processes instituted and the activities undertaken by
the implementing agency to get ready for or to facilitate
the implementation of the policy. Thus, the specific design
used for the second round visits was primarily cross-sec-
tional, with provisions for approximating some longitudinal
data within the overall framework of the cross-sectional
design.

The use of a cross-sectional design, with provisions
for approximating longitudinal data, was necessitated by
the specific requirements of the project. Initially, the
overall design of the project called for two rounds of
visits to approximately twenty-five programs in each round
and did not rule out the possibility of visiting the same
twenty-five programs twice. However, during the initial

segments of the project it became evident that the project team needed first-hand, on-site information about Head Start and the initial stages of implementation of the handicapped policy before formulating the specific hypotheses that could be tested empirically. Consequently, the approach to Task III was revised early in the project to accomodate this need. Thus, the first round visits covered sixteen programs with an unstructured, exploratory data collection technique. This allowed the project team to increase the sample size for the second round by about fifty per cent (from twenty-five to thirty-seven) and to field a more relevant and comprehensive questionnaire than what would have been possible had the original design been implemented.

The cross-sectional survey design used in the study is clearly a non-experimental design and, therefore, suffers from many of the phenomena that typically limit the internal and the external validity of all non-experimental designs. Using Campbell and Stanley's terminology, the design of the second round visits can at best be identified as a "one-group pretest-posttest design."[1] And this would only be possible if we can assume that approximation of longitudinal data through the cross-sectional survey covers measurement of our primary dependent variable at two points in time: before the start and at the conclusion of the

[1]See: Donald T. Campbell and Julian C. Stanley, _Experimental and Quasi-Experimental Designs for Research_ (Chicago: Rand McNally and Co., 1963), pp. 7-12

program year during which the policy was implemented. Even this provision is not sufficient to raise the validity of the design to a level comparable to what one would achieve with the so-called "true" experimental designs.

At this point it would be appropriate to discuss two of the reasons for the unapplicability of a true or quasi-experimental design to the problem we have at hand.

First, it should be kept in mind that we are attempting to explain the difference between successful and unsuccessful implementors of a given policy, where all implementing agencies are mandated to implement the policy by law. It is, therefore, impossible to manipulate some variables and observe their effects on the dependent variable. By the same token, the circumstances at hand precludes prior selection of some programs as the "experimental" and others as the "control" group. Success can be observed only *ex post facto* and the programs can be divided into "successful" and "unsuccessful" groups after completion of the implementation of the policy.

Second, the research problem at hand is highly exploratory. There are no clear and well-formulated theories of implementation which suggest potentially explanatory variables which could be easily manipulated or varied in an experimental or quasi-experimental situation. Instead, we have a conceptual model with a set of independent variables. Relationships between each independent variable and the dependent variable and the dependence of the independent

variables upon one another have been postulated only at an exploratory level. If these relationships are not supported by the data at hand, we need to search for alternative explanations which can be verified. Use of an experimental or quasi-experimental design, however complex it might be, precludes exploration and limits the search for alternative explanations.

The basic superiority of the experimental designs over the others is that through experimental work the researcher can effectively control for other variables which can potentially influence the dependent variable. This often becomes an insurmountable task to achieve through non-experimental designs such as those based on survey research. In these instances the primary alternatives available to the researcher are: control of background factors by randomization, by specification, or by some combination of these two.[1] In the case of the present study, as we shall see below, the alternatives we have used are similar to these three.

Questionnaire Design

The instrument used in the second round visits to regular Head Start programs was developed over a two-month period following analysis of the data generated during the

---

[1]See: Robert T. Holt and John E. Turner, "The Methodology of Comparative Research," in The Methodology of Comparative Research, ed. by Robert T. Holt and John E. Turner (New York: The Free Press, 1970), pp. 1-20.

first round. The most critical issues to be explored during this phase of the project were identified, discussed, and finalized by the principals of the project team and the members of the Senior Consultant Group. It was agreed that three types of data should be collected in this round: data on the implementing agencies, data on children with specific handicaps, and data on classrooms within which handicapped children were being served. The final version of the questionnaire reflects this breakdown. It has four distinct parts:

Part I   : Identifying Information

Part II   : Program-Level Information

Part III : Child-Specific Information

Part IV   : Guide for Classroom Observations.

Data from the last two parts of the questionnaire served very specific purposes and are not used in this dissertation.

Design of the survey instrument was completed in five stages. The following is an overview of the activities undertaken in each stage.

First, the primary purposes of the second round visits were identified to be the following: (1) How well are Head Start programs presently serving the handicapped, in particcular the severely handicapped? (2) Why are some Head Start programs presently not serving the severely impaired? (3) What are the factors which account for the differences between programs in enrollment of severely handicapped children?

The answer to the first question requires collection of data on level of overall implementation success as well as on specific services provided to individual children within and outside the classroom. The second question involves in-depth study of programs with little or no severely handicapped children and their comparison with programs with higher enrollment figures. The third question requires generation of data on a series of potentially explanatory variables.

Second, the general structure of each part of the questionnaire was formulated and discussed. Key issues and potentially explanatory factors were identified. Tentative findings from the first round visits played an important role in this phase of questionnaire design.

Third, a draft questionnaire was prepared. This draft was reviewed and revised twice by a core group from the project team. The revised draft was reviewed by all project principals and some graduate students. Coding, data processing, and analysis implications of each question were discussed. This led to the finalized version of the questionnaire.

Fourth, the finalized version of the questionnaire was mailed to Washington, D.C., along with a statement describing how it will be used, for approvals by the Department of Health, Education, and Welfare and the Office of Management and Budget. No changes were suggested by the reviewers from the two federal agencies,[1] indicating that the questionnaire

---

[1] The OMB Approval Number was OMB-85-S-74-005.

and the data collection plan satisfied their minimum criteria of acceptability.

Fifth, implications of the survey instrument for training the interviewers, coding the responses and processing the data were identified and discussed.

Selection of Programs

A total of thirty-seven regular Head Start programs were originally selected for the second round of visits. One of these programs withdrew from the sample two days prior to the scheduled visit to that program and was not replaced due to logistical problems. Of the remaining thirty-six programs thirty-three were distributed among the ten OCD regions and three were Indian programs.

The sample was selected through a complicated scheme. Here we will describe the essential features of the procedures used in identifying the thirty-seven programs. These features include:

- sampling elements
- survey population
- sample size
- sampling frame
- stratification of programs
- screening criteria
- initial selection
- telephone interviews
- final sample

We will now discuss each feature in the order they appear above.

There were three types of *sampling elements* considered. First, regular Head Start programs, with or without handicapped children, comprised the primary unit of analysis. If a grantee agency had no delegates, the grantee was considered as a sampling element. If the grantee agency did not administer any Head Start center, i.e., if all centers were run by one or more delegate agencies, each delegate agency was given a sampling element status. Finally, if a grantee agency ran some centers and delegated others, both the grantee and the delegate agencies were considered as separate elements. This definition coincides exactly with the definition of a "respondent" used in the mail survey conducted in Task II.

Each handicapped child enrolled in Head Start was considered as a second type of sampling element. The third element consisted of each Head Start classroom with at least one handicapped child. Data from the children and the classroom samples are not studied here.

The *survey populations* were defined as the total population of 1973-74 Full-Year Head Start programs, the total population of handicapped children enrolled in these programs, at the time of the survey, and the population of Head Start classrooms with at least one handicapped child.

The *sample size* was determined on the basis of resource constraints and contractual obligations. Given that sixteen

programs were visited during the first round visits and that
a total of fifty regular Head Start programs were to be vis-
ited in Task III, the contract with HEW demanded completion
of visits to at least thirty-four programs. This figure was
increased to thirty-seven by adopting some cost-saving measures
such as visiting the selected programs in a round-robin fash-
ion instead of returning back to Syracuse after completion
of each visit. The contract also demanded that at least ten
of the visits were to be made to programs with no or very
few handicapped children.

The *sampling frame* consisted of the computerized list of
1,353 Full-Year Head Start programs which had responded to
the mail survey conducted during November-December 1973.
79 per cent of all grantees and delegate agencies responded
to this survey and at least one response was obtained from
83 per cent of all Head Start grantees.

The closeness of the sampling frame to the survey
population is quite important in judging the external valid-
ity of the findings based on a sample drawn from this sampl-
ing frame. For this reason we will briefly review here the
results of two special surveys conducted within Task II for
assessing the reliability of the responses to the mail survey
and for estimating the effects of nonresponse.

The test of the reliability of the Full-Year survey
was based on telephone interviews with a randomly selected
sample of respondents to the mail survey. This telephone
survey was conducted during the last half of January 1974,

approximately one month after the responses to the Full-Year survey were received. A simple random sample of one hundred programs were selected for telephone interviewing. After three attempts, there were thirty-three programs which did not respond to the survey and, of the remaining sixty-seven, three were incomplete and were dropped from the analysis. The usable sample of sixty-four responses were analyzed to assess the differences between the two surveys in terms of the reported number of handicapped children, their handicapping conditions, and the severity of their handicaps. The principal measure used to estimate the reliability of the Full-Year survey was the correlation coefficient between the total number of diagnosed handicapped children reported by the sampled programs to the Full-Year survey and the telephone survey. This coefficient was estimated at .972 which indicated that the two surveys did not differ statistically from one another in terms of the total number of diagnosed handicapped children.[1]

The second special survey in Task II was conducted in order to ascertain whether the nonrespondents to the mail survey differed significantly from the respondents in terms of enrollment of handicapped children. The survey was conducted in January 1974 and covered the 120 nonresponding

---

[1]For a detailed account of the two special surveys conducted within Task II see: "Appendix E: Verification of the Findings From the Survey of Full-Year 1973-74 Head Start Programs," in "The Status of Handicapped Children in Head Start – Final Report Based on Surveys of Summer 1973 and Full Year 1973-74 Head Start Programs."

grantee agencies. After three attempts responses were obtained from sixty-nine grantees. Responses from the grantees covered the handicapped enrollments in all centers administered by the grantee, including those operated by delegate agencies.

Two statistical tests were conducted to assess the significance of the differences between the reported handicapped enrollments of the respondents and the nonrespondents to the Full-Year survey. First, the difference in average percentage handicapped in the two groups was not found to be significant at the .01 level. Second, a Kolmogorov-Smirnov one-sample test revealed that the percentage distributions of handicapping conditions did not differ from one another at the .01 level of significance. This led to the following conclusion:

> ". . . On the basis of the above discussion, we have no reason to believe that the findings from the Full Year Survey would have changed had we been able to obtain a greater response rate. In other words, programs that did not respond to the Full Year Survey were not at all different, in terms of the variables used to measure the status of OCD's handicapped effort, from those that did respond to the mail survey."[1]

Thus, based on the evidence presented above, we can safely assume that the sampling frame used for selecting the sample of programs visited in the second round is a good approximation of the true survey population studied here.

The first major step in selecting the sample involved *stratification of programs* in terms of percentage of severely

---

[1]*Ibid.*, p. 18

handicapped enrolled and size. Percentage severely handi-
capped was chosen because it represented the major focus of
the second round visits. Program size, measured in terms of
the total enrollment, was selected as a stratifying variable
because it accurately reflects other dimensions such as
economies of scale and the urban-rural dichotomy.

Let us review, first, the stratification of programs
into clusters in terms of enrollments of severely handi-
capped children. One of the clusters was identified prior
to starting the stratification as a result of the contractual
requirement that we study at least ten programs with no handi-
capped children. This meant that it was necessary to divide
the remainder of the sample into at least two clusters if
we intended to make any meaningful comparisons between groups
of programs. Recognizing that dividing the remainder of the
sample into three clusters would result in a very low aver-
age of nine programs per cluster, it was decided to split
the remaining twenty-seven programs to only two clusters.
This decision was executed by first obtaining a frequency
distribution of all programs in terms of the percentage of
severely handicapped children enrolled. This distribution
was examined to arrive at a breakpoint which divides the programs
with handicapped children into two clusters. Of all poten-
tial breakpoints considered--which included the mean, the
median, the upper and lower quartiles, and a few "natural"
breakpoints in the frequency distribution--a natural
breakpoint corresponding to 3.5 per cent of total enrollment

was chosen as the dividing line between the two clusters.

Thus, according to this stratification criterion, the three

clusters we defined in the following manner:

Cluster I   : Programs in which the severely handicapped constitute 3.5 per cent or more of total enrollment

Cluster II   : Programs with handicapped children in which the severely handicapped constitute less than 3.5 per cent of total enrollment

Cluster III   : Programs reporting no handicapped children enrolled.

Stratification of programs in terms of their total

enrollment involved a similar process. In this instance,

initially two basic alternatives were explored. First, the

total population of Head Start programs was divided into

three groups containing equal numbers of programs. This

option was found unsatisfactory since it resulted in allo-

cation of a very large proportion of the sample to small

programs. The second alternative involved division of

programs into groups in a manner that would equate the

total number of children enrolled in each group. This sec-

ond option was also found unsatisfactory because one of the

groups included only very few extremely large programs. A

workable solution was found after a study of the interval

between the breakpoints which corresponded to these two

extremes. The solution consisted of selecting two natural

breakpoints within the two intervals. The resulting size

groups were as follows:

Small programs   :   1 - 120 enrolees

Medium programs   :   121 - 300 enrolees

Large programs   :   301 or more enrolees

Distribution of the 1,353 programs into the three

clusters and the three size groups is displayed in Table A.1.

TABLE A.1

DISTRIBUTION OF NUMBER OF HEAD START
PROGRAMS BY SIZE AND CLUSTER

Size	Cluster I	Cluster II	Cluster III	Total
Small	138	566	128	832
Medium	48	281	22	351
Large	15	144	11	170
Total	201	991	161	1,353

Source: Systems Research Incorporated, Head Start Project files,
Lansing, Michigan, 1974.

Selection of the thirty-seven programs was preceded by

the introduction of a constraint which dropped many of the

programs from the sampling frame. The need to obtain a

sample of programs with "adequate" number of children repre-

senting all handicapping conditions and severity levels was

the primary reason for the adoption of the two *screening*

criteria. The first criterion applied only to programs in

Cluster I and demanded representation of at least three of

the nine handicapping conditions, with at least two severely

handicapped children in each of the three handicapping condition categories. The second criterion applied only to programs in Cluster II and required representation of at least three handicapping conditions, with at least two minimally or moderately handicapped children in each of the three handicapping condition categories. These two criteria ensured that the field workers would be able to collect child-specific data at programs in Clusters I and II on at least three children with different handicapping conditions and that they would also have the option to select from at least two children within a given handicapping condition and severity level. These criteria were not at all applied to programs in Cluster III, since no child-specific or classroom observation data were to be collected from programs with no handicapped children.

The sample frame reduced by the two screening criteria is displayed in Table A.2.

Initial selection of the thirty-seven programs involved introduction of two additional constraints. First, it was felt that the programs selected from Clusters I and II should provide "minimal" representation of each handicapping condition. Second, it was agreed that during the selection process an attempt should be made to attain an "adequate" representation of programs across OCD regions.

The ten programs selected from Cluster III were drawn through a two stage process. First, a simple random sample of ten programs was selected. Second, the regional

TABLE A.2

DISTRIBUTION OF NUMBER OF HEAD START
PROGRAMS WHICH MET THE SCREENING
CRITERIA BY SIZE AND CLUSTER

Size	Cluster I	Cluster II	Cluster III	Total
Small	6	76	128	210
Medium	14	149	22	185
Large	6	100	11	117
Total	26	325	161	512

Source: Systems Research Incorporated, Head Start Project files,
Lansing, Michigan, 1974.

distribution of this sample was studied to observe if it dif-
fered significantly from the regional distribution of the pop-
ulation of 161 programs. In this instance the regional dis-
tribution of the first random sample selected was judged to
be "adequate" and, therefore, no replacements were made.

Selection of programs from Cluster II involved a dif-
ferent procedure. First, a quota of twelve programs was
established for Cluster I, leaving a total sample size of
fifteen for Cluster II. Second, a twenty-five per cent sim-
ple random sample was drawn from each size group. Programs
falling into this sample of eighty-one were studied individ-
ually in terms of their region and distribution of mildly/
moderately handicapped. Third, an initial sample of fifteen
programs was selected purposively from the reduced universe.

Selection of these programs was made collectively by three of
the project principals after a discussion of the advantages
and the disadvantages of including each of the eighty-one
programs into the sample of fifteen.

The quota of twelve for programs in Cluster I was dic-
tated by the need to obtain a sample of at least four severely
handicapped children for each handicapping condition. Given
that the minimum of four children with a particular handi-
capping condition were to be drawn from at least four differ-
ent programs, and given also that no more than four children
were to be studied at each program selected from Cluster I,
it was deduced that it would be necessary to select at least
twelve programs from Cluster I.

Selection of the twelve programs from Cluster I was
made on a purposive basis. Since there were only twenty-six
programs left after application of the two screening criteria,
there was no room for further reduction of the parent popu-
lation. The purposive selection process for Cluster I was
the same as the one used for Cluster II.

The next step in the selection process involved comple-
tion of *telephone interviews* with the directors of the thirty-
seven programs. These interviews had four specific purposes:
(1) to verify the questionnaire data on the numbers of mildly/
moderately and severely handicapped children, (2) to verify
the presence of children with specific handicaps selected to
the case-study sample, (3) to obtain the consent of the pro-
gram director for studying his program in the second round,

and (4) to schedule the visit and make the initial logistical arrangements. All telephone interviews were conducted by the project principals at Syracuse University. As a result of these interviews four programs were dropped from the sample because of scheduling difficulties. No program was dropped due to the absence from the program of previously reported handicapped children.

Each of the four programs dropped from the sample was replaced with one from the same cluster-size group. In addition, an attempt was made to select the substitute from the same OCD region and from among the programs with the same or similar combination of handicapping conditions. Interviews were also conducted with the directors of the four new programs.

The thirty-seven programs thus identified constituted the *final sample* of programs to be visited in the second round. One of the programs in the sample withdrew two days before the visit as a result of scheduling difficulties and unanticipated commitments. This program was not replaced with another one since most programs were drawing close to the end of their program year. Geographical distribution of the thirty-six programs in the final sample is displayed in Table A.3.

Pre-fieldwork Activities

Two types of pre-fieldwork activities deserve our attention here. The first one relates to the training of the

TABLE A.3

DISTRIBUTION OF THE SITES SELECTED FOR THE
SECOND ROUND VISITS BY OCD REGION,
PROGRAM SIZE,[a] AND CLUSTER

OCD Region	Cluster I	Cluster II	Cluster III
I. Boston	● Charlestown, Mass.(S) ● Bath, Maine(M) ● Portland, Maine(M)	● Keene,N.H.(S)	
II. New York	● Newark, N.J.(L)		● Bronx,N.Y.(S) ● Hudson,N.Y.(S)
III. Philadelphia	● Williamston,W.Va.(M)	● Charleston, W.Va.(S) ● Reading,Pa.(M)	● Scranton,Pa.(S)
IV. Atlanta	● Talladega,Ala.(S) ● Pikeville,Ky.(L)	● Columbus,Ga.(S) ● Boone,N.C.(S) ● Chattanooga,Tenn.(M) ● Jackson,Miss.(L) ● Monterey,Tenn.(L)	● Waycross,Ga.(M) ● Natchez,Miss.(L)
V. Chicago	● Waverly,Minn.(S)	● Lansing,Mich.(L)	● Pontiac,Mich(M) ● South Bend,Ind.(L)
VI. Dallas		● Forrest City,Ark.(M)	
VII. Kansas City	● Hastings,Nebr.(S)	● Remsen, Iowa (S)	

TABLE A.3 (continued)

OCD Region	Cluster I	Cluster II	Cluster III
VIII. Denver		● Canon City,Colo.(S)	
IX. San Francisco	● Visalia,Calif.(L)	● Reno,Nev.(S)	● Fresno,Calif.(S)
X. Seattle	● Spokane,Wash.(M) ● Seattle,Wash.(M)		
XI. Indian/Migrant		● Pine Ridge,S.Dak(M)	● Somerton,Ariz(S) ● Auburn,Wash.(S)

[a]Letters in parantheses indicate the size of the program. "S" stands for small, "M" for medium, and "L" for large.

Source: Systems Research Incorporated, Head Start Project files, Lansing, Michigan, 1974.

interviewer/observer group. The second pre-fieldwork activity concerns other preparatory activities.

Training of the interviewers for the second round visits differed sharply from the training of the participant observers for the first round of visits. Although our main concern is with the data generated through the second round visits, we will first describe the training of the observers for the first round since almost all of the interviewers who participated in the second round were also involved in conducting the first round visits.

Training of observers for the first round was spread over approximately two months in Fall, 1973. The training was heavily "process" oriented, concentrating on techniques of open-ended interviewing and participant observation. The observers were provided with considerable information about Head Start and the specific policy examined in the project. Periodic meetings were held in Syracuse and Boston. Towards the end of the two-month period a three-day meeting was held during which six members of the project team from SRI and the Syracuse University reviewed the final draft of the "observation guide" prepared for the first round.

The formal portion of the training was followed by a series of pilot visits to Head Start programs not in the first round sample. All fieldwork staff participated in these visits. Upon their return, a one-day review session was held to discuss the problems encountered and to provide additional guidance. Actual site visits to the sampled

programs were conducted by a fieldwork staff of sixteen per-
sons, including three professors of special education from
Syracuse University, one from Boston University, and one
from Indiana University. The rest of the fieldwork staff
consisted mainly of graduate students from Syracuse Univer-
sity.

The first round visits lasted two to five days per pro-
gram and were conducted by a team of two to three observers
per site. The visits involved semi-structured interviews
and informal contact with the Head Start director, the teach-
ing staff, parents, volunteers, and others involved with the
implementation of the handicapped policy. At least two
classroom observations were also completed at each site.

Training of the interviewers for the second round was
conducted in a much more structured manner and lasted about
a week. The first phase of the training included a two-day
formal training session during which the questionnaire and
related fieldwork documents were reviewed and discussed.
This portion of the training was conducted by two principals
of the project team and an outside consultant. This session
was followed, during the second phase, with three days of
classroom observations using Part IV of the questionnaire.
These observations were made in three carefully selected
preschool and/or special education settings which represented
a wide range of educational approaches. During these ses-
sions one of the project principals served as the "criterion
observer" and the consultant provided expert advice in

interpreting and analyzing the differences between the observers. A 75 to 80 per cent agreement was achieved on most items of the observation schedule.[1]

Other pre-fieldwork activities were conducted simultaneously with the training. These included preparation of a "program profile package" for each sampled program, assignment of sampled programs to the interviews, completion of scheduling and logistics arrangements, and development of contingency plans.

The "program profile package" contained all the data at the disposal of the project team about the sampled program, including the responses of the program to the mail survey. Programs were assigned to the interviewers in such a manner as to minimize transportation costs. Thus, each interview team was assigned all the programs in a geographical area so that visits could be made on a round-robin basis. All appointments and travel schedules were made at a central location at Syracuse University. In addition, a contingency plan was developed which outlined the actions to be taken in case one or more of the interviewers could not conduct the visits on the scheduled dates.

---

[1]For a discussion of methodology of and the findings from the child-specific and classroom observation portions of the second round visits see: "Final Report on Assessment of the Handicapped Effort in Experimental, Exemplary, and Selected Other Head Start Programs Serving the Handicapped" (report prepared by Division of Special Education and Rehabilitation, Syracuse University, October 1974).

Data Collection

Visits to the sampled programs in Cluster III lasted only a day and were conducted mostly on a one interviewer per program basis. Visits to the programs in Clusters I and II were about two to three days in duration and, except for five large programs, one interviewer collected the data from each selected program.

The principal respondent to the "Program-Level Data" section of the questionnaire was the Head Start director at the sampled program. The coordinator of the handicapped program at the selected agency provided answers to some of the factual questions. In all instances an attempt was made to obtain copies of documents supporting the responses to the factual questions.

The primary respondent to the child-specific questions was the principal teacher of the classrooms the selected child had been assigned to. Only the classrooms attended by the children selected to the case-study sample were observed by the interviewers.

The interviewers, in all instances, reviewed the responses recorded during the interviews prior to departing from the site. Questions with missing or unclear responses were reviewed a second time with the particular respondent.

Completion of the interviews for the program-level data portion of the questionnaire took at least three hours and this was only in the case of small programs in Cluster III. In other instances the interviews for this portion of the

questionnaire lasted an average of one day.

The interviewers, while in the field, periodically contacted the central office established at Syracuse University to report on their progress and to review special problems encountered.

The fieldwork was conducted over a six week period in April-May 1974. No major problems were encountered, except for the scheduling difficulty with one program which was later dropped from the sample.

Post-fieldwork Activities

Activities which followed completion of the fieldwork included finalization of the codebook, editing the collected data, reinterviews with respondents, coding of quantitative and anectodal data, and data processing.

The initial version of the codebook was prepared while the fielwork was still going on. It was completed after an initial review of the completed questionnaires. The final version of the codebook prepared by the project staff at Syracuse University included the codes for 1,035 variables. This included 383 program-level variables, 270 child-specific variables, and 382 classroom observation variables.

The data collected at the field were edited by a core staff from Syracuse University. The interviewers helped clarify their notes and, in some cases, telephone interviews were conducted with the respondents to obtain additional information.

The quantitative data, which totaled to 1,035 variables, were coded by three persons intimately familiar with the project who were supervised by a project principal. The project principal also conducted spot checks of the coded data.

Anectodal information were taken off the questionnaires and typed, question by question, on separate index cards. These responses were sorted and analyzed by two members of the project team with knowledge of and experience in procedures of qualitative analysis. These analyses were later used in illuminating the major differences between the programs in the three clusters.

Keypunching and processing of the coded data were conducted at Syracuse University. Analysis of the data by the Syracuse University team was mainly based on frequency distributions of the 1,035 variables, cross-tabulations between cluster groupings and selected program-level variables, frequency distributions of a few transformed variables, and a set of bivariate correlations. No multivariate techniques were used in analyzing the second round data. Results of these analyses are reported in the final report for Task III.

APPENDIX B

MEASUREMENT OF SECONDARY VARIABLES

In this appendix we examine the pool of indicators considered in developing the measures used for the six primary variables of the conceptual model. The term "secondary variables" refers to these indicators.[1]

We have decided to fully describe the process used in measuring the secondary variables for two basic reasons. First, the extreme length of the actual questionnaire used during the fieldwork made it impractical to include it in the dissertation as an appendix. The format we are using in this appendix gives us the opportunity to quote specific questions from the survey instrument and present the exact wording of the questions that are relevant to our inquiry. Secondly, we felt that there was a need to discuss the appropriateness of the empirical indicator used for each secondary variable since these indicators determine, to a large extent, the validity and the reliability of the measures used for the primary variables of the conceptual model.

The appendix is divided into six sections along the lines of the six primary variables. Within each section we

---

[1]They are "secondary" because none of these are included in the model directly and they are "variables" because they represent quantities which show variation across the sampled implementation agencies.

present and discuss each secondary variable in the following manner. First, we list the acronym and the full name of the secondary variable. We then state the full formal definition used for that variable. This is followed by a description of the specific indicator used for measuring the variable. The description ends with a discussion of the appropriateness of the empirical indicator for measuring that specific secondary variable.

Level of Implementation Success

This is our primary dependent variable and it is measured in terms of two secondary variables, PSEV and PMOD, which are defined below.

Variable 1 : PSEV—percentage severely handicapped

Definition : Number of severely handicapped children in the program expressed as a percentage of the total number of children enrolled.

Indicator : Responses of the program director to the following questions: *"Which of the handicapped children in your program would you consider to be severely impaired?"* and *"What is the total 1973-74 full year Head Start enrollment in the centers [administered by your agency]?"* were used in computing the percentage.

Discussion : The question on the number of severely handicapped children enrolled was preceded by questions on the total number of handicapped children, the breakdown of these children into ten handicapping conditions based on their most disabling condition, the persons or agencies that diagnosed the handicapping conditions of these children, and the number of additional children who are suspected to be handicapped but whose handicapping conditions have not been clearly diagnosed. The question on the number of severely handicapped, therefore, represents breakdown of the total population of handicapped children into two mutually exclusive categories:

those who have been diagnosed as severely handicapped and those who have been diagnosed as mildly or moderately handicapped. The program directors responded to the question, usually after examination of the agency's records, by indicating the number of severely handicapped children in each of the following ten categories of disabling conditions which are included in the original congressional mandate:

- number totally blind
- number severely visually impaired
- number totally deaf
- number severely hearing impaired
- number severely health or developmentally impaired
- number severely physically handicapped
- number severely speech impaired
- number severely seriously emotionally disturbed
- number severely mentally retarded
- number severely handicapped but with undifferential handicapping conditions.

Variable 2   : PMOD--percentage mildly or moderately handicapped

Definition   : Number of mildly or moderately handicapped children in the program expressed as a percentage of the total number of children enrolled.

Indicator   : Responses of the program director to the following three questions: *"How many handicapped children are currently enrolled in your program?"*, *"Which of the handicapped children in your program would you consider to be severely impaired?"*, and *"What is the total 1973-74 full year Head Start enrollment in the centers [administered by your agency]?"* were used in computing, first, the number of mildly or moderately handicapped as the difference between the total number of handicapped and the total number of severely handicapped and, second, the percentage mildly or moderately handicapped.

Discussion   : See the discussion of PSEV, above. The same categories of disabling conditions were used for the mildly or moderately handicapped.

Level of Effort

A total of eight variables were considered as separate dimensions of level of effort: ORGS, PLAN, COOR, MOBR, FUND, SIZE, TATA, and MFAC.

Variable 3 : ORGS--effort to organize

Definition : Effort to organize for the implementation of the policy.

Indicator : A three-point scale based on the responses of the program director to the following questions: *"Now I'd like to ask about how you have organized your staff with respect to the provision of services to handicapped children. Who is in charge of your program for handicapped children? Approximately how much of this person's time is devoted to the program for handicapped children?"* The three-point scale has the following values:

  0 = there is no person in charge of the implementation of the policy.

  1 = there is a person in charge of the implementation and he devotes less than twenty-five per cent of his time to the coordination/management of the implementation.

  2 = there is a person in charge of the implementation and he devotes twenty-five per cent or more of his time to the coordination/management of the implementation.

Discussion : The breakpoints for the scale were identified after an examination of the frequency distribution of percentage of time devoted to the policy by the person in charge. Programs with lower scale values are assumed to have spent less effort in coordinating/managing the implementation of the policy.

Variable 4 : PLAN--planning effort

Definition : Comprehensiveness of the plan prepared for implementing the handicapped policy.

Indicator : Number of policy-specific topics that are covered in the plan. This is based on the response of the program director to the question: *"Did your agency prepare a plan to recruit and provide*

*services to handicapped children? If yes, what
did the plan cover?"* The responses given were
coded in terms of coverage of the following
planning areas: recruitment; screening, testing,
diagnosis; involvement of parents; classroom pro-
grams; special materials and equipment; special
physical facilities; monitoring and evaluation;
others.

Discussion : Scores based on this indicator range from zero
(for programs with no plan) to eight (for programs
with plans which cover all eight areas). Responses
of the program director were not verified by com-
paring the planning areas mentioned with the con-
tents of the written plan. It is assumed that
agencies with more comprehensive plans have spent
a greater amount of effort in planning the imple-
mentation.

Variable 5 : COOR--coordination effort

Definition : Coordination of the implementation within the
hierarchy of the agency administering the policy.

Indicator : A three-point scale based on the responses of the
program director to the following two questions:
*"With respect to handicapped children do you work
or coordinate your efforts in any way with other
Head Start programs?"* and *"With respect to handi-
capped children, what assistance, guidance, or
support, if any, have you received from the
regional office of OCD?"* The three-point scale
has the following values:

0 = no coordination of the handicapped effort
with other Head Start programs or the
regional office.

1 = coordination of the handicapped effort with
either the regional office or other Head
Start programs, but not both.

2 = coordination of the handicapped effort with
both the regional office and other Head
Start programs.

Discussion : It is assumed that higher levels of the scale rep-
resent a higher level of coordination effort. In
general, programs which coordinate their efforts
with other agencies within OCD learn more about
the services to the handicapped and become more
up-to-date on new developments regarding the imple-
mentation of the policy. Through coordination

they may also have a greater chance to pool their resources with those of other Head Start programs and arrange for services which would not otherwise be available within the constraints of any single program.

Variable 6 : MOBR--mobilization effort

Definition : Effort to mobilize the resources of the community.

Indicator : Number of community agencies providing special services for handicapped children in the program. It is based on the response of the program director to the question: *"What agencies are providing special services to the handicapped children in your program?"*

Discussion : Data were collected on up to four agencies providing special services. Thus, the indicator's range is between zero (no agencies providing special services) and four (four or more agencies providing special services). It is assumed that programs with greater utilization of the resources of the community have indirectly put in a higher level of effort to implement the policy.

Variable 7 : FUND--funding effort

Definition : Effort to secure funds for the implementation.

Indicator : Response of the program director to the question: *"Did your program receive any money, in addition to your regular OCD grant, for the purpose of serving handicapped children?"* Responses were coded as "yes" or "no."

Discussion : OCD did not provide any additional funds to local programs for serving the handicapped. However, some programs were able to obtain special funds from other agencies for the purpose of serving the handicapped. (This type of assistance is separate from the in-kind services provided by community agencies.) Thus, it is assumed that programs which have secured additional funding have put in a greater effort to implement the policy than those with no such additional funds.

Variable 8 : SIZE--staffing effort

Definition : Effort to increase the size of the implementation staff.

Indicator : Presence-absence of persons added to the Head

Start staff for the purpose of serving handicapped children, which is estimated on the basis of the response of the program director to the question: *"How many persons have been added to your staff this year for the primary purpose of serving or working with handicapped children?"*

Discussion : Hiring of new staff solely for the purpose of implementing the policy indicates a high level of commitment to the policy. The additional staff includes full-time or part-time staff who are either paid by Head Start or other agencies or who are working for the program on a voluntary basis.

Variable 9 : TATA—training and technical assistance effort

Definition : Level of policy-related training and technical assistance provided to the program staff.

Indicator : Number of distinct training and technical assistance activities participated by the staff of the program during the program year. This is estimated on the basis of the response of the program director to the question: *"Has any training or technical assistance been provided in your program this year for the primary purpose of serving handicapped children?"* and detailed descriptions of the nature of each training and technical assistance activity mentioned. The range of the indicator is from zero (for programs with a "no" answer to the question quoted above) to three (for programs with three or more training and technical assistance activities).

Discussion : Training and technical assistance activities include in-service or pre-service workshops, college conferences, and college-level courses. Participation in these activities is an indicator of the level of commitment the program has made to the upgrading of the policy-related skills of its staff and, in general, to the implementation of the policy.

Variable 10 : MFAC—effort to modify the facilities

Definition : Effort to modify the physical facilities in order to make them more appropriate for serving the handicapped.

Indicator : A dichotomy based on the response of the program director to the question: *"Have you made any changes in the physical facilities of your*

program this year for the purpose of serving handicapped children?"

Discussion : This variable refers to the efforts of the agency to upgrade the suitability of its physical facilities. The types of changes may include building ramps, enlarging passage areas, adding fixtures to toilets, etc. Although the policy does not mandate that all programs modify their facilities if they are not appropriate for serving all types of handicapped children, programs which have made such changes can be assumed to have made a greater commitment to serving the severely handicapped not only during the 1973-74 program year but in future years as well.

Implementation Potential

A total of five variables were considered as separate dimensions of implementation potential: EXPR, LEAD, KNOW, QUAL, and AFAC.

Variable 11 : EXPR--past policy experience

Definition : Previous experience of the implementing agency in executing the same or a similar policy.

Indicator : Responses of the program director to the following two questions: *"Did you have any handicapped children in your program last year?"* and *"Did you consider any of these children to be severely handicapped?"* coded as a dichotomy.

Discussion : Programs with severely handicapped children during the previous program year are viewed as more experienced to implement the policy than those with no such children. Data on the numbers of severely handicapped were not available consistently; therefore it was not possible to build an indicator at a higher level of measurement.

Variable 12 : LEAD--lead time

Definition : Lead time the implementing agency has had prior to start of the implementation.

Indicator : Response of the program director to the question: *"When did you find out about the requirement for Head Start programs to serve the handicapped?"* coded in terms of the number of months between

notification and the start of enrollment for the 1973-74 Full-Year program.

Discussion : Responses of the program director were not verified. Notification of the program about the policy could have come from the national or regional office of OCD or it could have come from the grantee agency if the program studied is a delegate to another agency. It is assumed that a longer lead time gives the program a greater chance to recruit severely handicapped children and to plan the delivery of services to these children.

Variable 13 : KNOW--knowledge of the target group

Definition : Knowledge of the size and the characteristics of the target group.

Indicator : Response of the program director to the question: *"In the area served by your program, do you know, or could you estimate, how many handicapped children there are who would be eligible for Head Start?"* The responses were coded as "yes-no."

Discussion : It is assumed that programs with some knowledge of the size of the handicapped population in their service areas would be in a better position to plan and control their recruitment efforts. With no knowledge of the size of the population of eligible handicapped children, the program would, most likely, have to follow a "first-come, first-served" type strategy rather than one which is based on search for children who can benefit from participation in the program most.

Variable 14 : QUAL--qualifications of staff.

Definition : Qualifications of the implementation staff.

Indicator : Presence-absence of persons on the program's staff who have completed at least one college-level course in special education. The number of staff with backgrounds in special education is estimated from the response of the program director to the question: *"How many persons on your staff have completed at least one course in special education at the college level?"*

Discussion : It is assumed that persons with backgrounds in special education are better qualified to serve the handicapped than those who have no such backgrounds. Responses to the question quoted above

were coded as "yes-no" because of the diffi-
culties faced in constructing an alternative
indicator which expressed the number of persons
with such backgrounds as a percentage of total
staff.

Variable 15 : AFAC—appropriateness of the facilities

Definition : Appropriateness of the physical facilities of the
program for serving the severely handicapped.

Indicator : A dichotomy based on the response of the program
director to the question: *"Do or would your pres-
ent physical facilities in any way make it diffi-
cult or impossible for you to serve certain kinds
of handicapped children?"*

Discussion : This variable concerns appropriateness of the
physical facilities for serving children with cer-
tain kinds of handicapped, such as the blind and
the physically handicapped. It is assumed that
programs with inadequate physical facilities would
be automatically excluding children with these
handicaps from participation in Head Start and
would, therefore, have a lower potential to serve
the severely handicapped.

Environmental Forces for Implementation

Only two variables were considered as separate dimen-

sions of environmental forces for implementation: COMM and

TARG.

Variable 16 : COMM—reaction from community agencies

Definition : Reaction of the agencies in the community to the
implementation of the policy.

Indicator : Response of the program director to the question:
*"Were other agencies in your community, who serve
handicapped children, aware of the requirement
for Head Start programs to serve handicapped
children? What did they think about it?"* which
was coded using the following five-point scale:
strongly agree, agree, neutral, disagree, strongly
disagree.

Discussion : Interviewers recorded the exact responses of the
program director to the question and the probes
which followed the question. Coding was done by

the interviewer on the basis of the actual
response given. The assigned code was reexam-
ined later by those editing the completed ques-
tionnaires. No direct contact was made with the
"other" community agencies. Therefore, the indi-
cator used illustrates the program director's
perception of the reaction from the competing
agencies in the same community.

Variable 17 : TARG--reaction from target group

Definition   : Reaction of the members of the target group to
                the implementation of the policy.

Indicator   : Response of the program director to the question:
                *"Were parents in your program aware of or informed
                about the requirement for Head Start to serve
                handicapped children? What did they think about
                it?"* which was coded using the following five-
                point scale: strongly agree, agree, neutral,
                disagree, strongly disagree.

Discussion   : Interviewers recorded the exact responses of the
                program director to the question and the probes
                which followed the question. Coding was done by
                the interviewer on the basis of the actual
                response given. The assigned code was examined
                later by those editing the completed question-
                naires. No direct contact was made with the
                parents of children enrolled in Head Start.
                Therefore, the indicator used illustrates the
                program director's perception of the parents'
                reaction to the mandate.

Policy Support

Two secondary variables were considered in construct-

ing the index for policy support: AGRM and ATTD.

Variable 18 : AGRM--agreement with the policy

Definition   : Level of agreement of the leadership of the imple-
                menting agency with the original intent of the
                policy.

Indicator   : Response of the program director to the question:
                *"How did you feel about the requirement to serve
                handicapped children?"* which was coded using the
                following five-point agreement scale: strongly
                agree, agree, neutral, disagree, strongly
                disagree.

Discussion	:	Interviewers recorded the exact responses of the program director to the question and the probes which followed the question. Coding was done by the interviewer on the basis of the actual response given. The assigned code was reexamined later by those editing the completed questionnaire.

Variable 19 : ATTD--attitudes towards the policy

Definition : Attitudes of the leadership of the implementing agency towards the policy.

Indicator : Attitudes of the program director toward serving severely handicapped children in Head Start measured as a total attitude score based on the responses to the following question:

"Please indicate whether you strongly agree, agree, don't know, disagree, or strongly disagree with each of the following statements:

"1. Head Start is not the right kind of setting for serving severely handicapped children.

"2. Head Start has enough to do in running a program for children who are not handicapped and should not be asked to take on the responsibility of serving severely handicapped children.

"3. Severely handicapped children will benefit from being in the same classroom with children who are not handicapped.

"4. Children who are not handicapped will benefit from being in the same classroom with severely handicapped children.

"5. Other agencies can serve severely handicapped children better than Head Start can."

Discussion : This indicator is based on a Likert-type scale with a range of possible total score between five and thirty. The items used in the scale were identified by project team members after a lengthy discussion of alternative items. No item analysis was conducted due to the expenses involved in conducting a pilot study which uses similar respondents. Reliability of the scale, therefore, has not been tested prior to data collection.

Self-Evaluation of Capability

Measurement of this variable was accomplished in the following manner:

Variable 20	:	CAPB--self-evaluation of capability
Definition	:	Capability of the implementing agency to execute the policy as evaluated by the leadership of the agency.
Indicator	:	The specific types of severe handicapping conditions the program could accomodate in the opinion of the program director. The <u>number</u> of such conditions was estimated on the basis of the responses of the program director to the following question: *"Taking into account the present capabilities of your staff, the physical resources and budget of your program, and the resources available to you from other agencies in the community, what kinds of severely handicapped children do you think you could serve?"* The list of handicapped conditions include the ten severely disabling conditions identified earlier.
Discussion	:	The emphasis of this indicator is on the types of severe handicapping conditions as opposed to the number of children with such conditions. It is assumed, therefore, that a program director who believes his program can accommodate a greater number of different handicapping conditions has greater confidence in his agency than a program director with self-perceived capabilities to serve fewer different disabling conditions. Thus, this variable, as operationalized here, can take on integer values between zero and ten.

SUPPLEMENTARY TABLES

TABLE C. 1

KNOWLEDGE OF THE SIZE OF THE HANDICAPPED
POPULATION BY IMPLEMENTATION SUCCESS

Knowledge of the Size of the Handicapped Population	Level of Implementation Success						Total	
	Low		Moderate		High			
	#	%	#	%	#	%	#	%
No Knowledge	10	90.0	13	92.9	4	40.0	27	77.1
Some Knowledge	1	9.1	1	7.1	6	60.0	8	22.9
Total	11	100.0	14	100.0	10	100.0	35	100.0

TABLE C. 2

QUALIFICATIONS OF STAFF BY
IMPLEMENTATION SUCCESS

Qualifications of the Implementation Staff	Level of Implementation Success						Total	
	Low		Moderate		High			
	#	%	#	%	#	%	#	%
No Persons with Special Education Background	5	45.5	2	14.3	0	0.0	7	20.0
Some Persons with Special Education Background	6	54.5	12	85.7	10	100.0	28	80.0
Total	11	100.0	14	100.0	10	100.0	35	100.0

TABLE C.3

APPROPRIATENESS OF FACILITIES BY
IMPLEMENTATION SUCCESS

Pre-Implementa-tion Status of Physical Facilities	Level of Implementation Success						Total	
	Low		Moderate		High			
	#	%	#	%	#	%	#	%
Not Appropriate	7	63.6	11	78.6	8	80.0	26	74.3
Appropriate	4	36.4	3	21.4	2	20.0	9	25.7
Total	11	100.0	14	100.0	10	100.0	35	100.0

TABLE C.4

REACTION FROM COMPETING COMMUNITY AGENCIES[a]
BY IMPLEMENTATION SUCCESS

Reaction from Community Agencies	Level of Implementation Success						Total	
	Low		Moderate		High			
	#	%	#	%	#	%	#	%
Strongly Oppose	2	18.2	0	0.0	0	0.0	2	5.7
Oppose	0	0.0	1	7.1	0	0.0	1	2.9
Neutral	7	63.6	4	28.6	4	40.0	15	42.9
Support	2	18.2	5	35.7	2	20.0	9	25.7
Strongly Support	0	0.0	4	28.6	4	40.0	8	22.9
Total	11	100.0	14	100.0	10	100.0	35	100.0

[a]This is measured in terms of the Head Start director's percep-tion of the reaction from other agencies in the community which also serve handicapped children.

TABLE C.5

REACTION FROM MEMBERS OF THE TARGET
GROUP[a] BY IMPLEMENTATION SUCCESS

Reaction from Target Group	Level of Implementation Success						Total	
	Low		Moderate		High			
	#	%	#	%	#	%	#	%
Strongly Oppose	1	9.1	0	0.0	0	0.0	1	2.9
Oppose	0	0.0	2	14.3	0	0.0	2	5.7
Neutral	9	81.8	4	28.6	4	40.0	17	48.6
Support	1	9.1	6	42.9	3	30.0	10	28.6
Strongly Support	0	0.0	2	14.3	3	30.0	5	14.3
Total	11	100.0	14	100.1	10	100.0	35	100.1

[a]This is measured in terms of the Head Start director's perception of the reaction from parents of children enrolled in Head Start.

TABLE C.6

REACTION FROM MEMBERS OF THE TARGET GROUP BY
REACTION FROM COMPETING COMMUNITY AGENCIES

Reaction from Target Group	Reaction from Community Agencies						Total	
	Strongly Oppose or Oppose		Neutral		Support or Strongly Support			
	#	%	#	%	#	%	#	%
Strongly Oppose or Oppose	2	66.6	1	5.9	0	0.0	3	8.6
Neutral	1	33.3	9	52.9	5	33.3	15	42.9
Support or Strongly Support	0	0.0	7	41.2	10	66.6	17	48.6
Total	3	99.9	17	100.0	15	99.9	35	100.1

TABLE C.7

REACTION FROM COMPETING COMMUNITY AGENCIES
BY AGREEMENT WITH THE POLICY

Reaction from Community Agencies	Director's Agreement with the Policy						Total	
	Strongly Disagree or Disagree		Neutral		Agree or Strongly Agree			
	#	%	#	%	#	%	#	%
Strongly Oppose or Oppose	2	25.0	1	16.7	0	0.0	3	8.6
Neutral	3	37.5	5	83.3	7	33.3	15	42.9
Support or Strongly Support	3	37.5	0	0.0	14	66.7	17	48.6
Total	8	100.0	6	100.0	21	100.0	35	100.1

TABLE C.8

REACTION FROM MEMBERS OF THE TARGET GROUP
BY AGREEMENT WITH THE POLICY

Reaction from Target Group	Director's Agreement with the Policy						Total	
	Strongly Disagree or Disagree		Neutral		Agree or Strongly Agree			
	#	%	#	%	#	%	#	%
Strongly Oppose or Oppose	1	12.5	1	16.7	1	4.8	3	8.6
Neutral	5	62.5	4	66.7	8	38.1	17	48.6
Support or Strongly Support	2	25.0	1	16.7	12	57.1	15	42.9
Total	8	100.0	6	100.1	21	100.0	35	100.1

TABLE C.9

PREVIOUS EXPERIENCE IN SERVING THE SEVERELY
HANDICAPPED BY ATTITUDES TOWARDS SERVING
THE SEVERELY HANDICAPPED

Previous Experience	Total Attitude Score								Total	
	10 or less		11-15		16-20		21-25			
	#	%	#	%	#	%	#	%	#	%
No Previous Experience	3	60.0	5	62.5	4	44.4	4	30.8	16	45.7
Some Previous Experience	2	40.0	3	37.5	5	55.6	9	69.2	19	54.3
Total	5	100.0	8	100.0	9	100.0	13	100.0	35	100.0

TABLE C.10

SELF-EVALUATION OF CAPABILITY BY
AGREEMENT WITH THE POLICY

Number of Handicapping Conditions	Director's Agreement with the Policy						Total	
	Strongly Disagree or Disagree		Neutral		Agree or Strongly Agree			
	#	%	#	%	#	%	#	%
0	3	37.5	1	16.7	1	4.8	5	14.3
1 - 2	3	37.5	2	33.3	1	4.8	6	17.1
3 - 4	1	12.5	0	0.0	2	9.5	3	8.6
5 - 6	0	0.0	2	33.3	4	19.0	6	17.1
7 - 8	0	0.0	0	0.0	8	38.1	8	22.9
9 -10	1	12.5	1	16.7	5	23.8	7	20.0
Total	8	100.0	6	100.0	21	99.9	35	100.0

TABLE C.11

EFFORT TO ORGANIZE BY IMPLEMENTATION
SUCCESS

Percentage of Coordinator's Time Devoted to the Implementation	Level of Implementation Success						Total	
	Low		Moderate		High			
	#	%	#	%	#	%	#	%
0	7	63.6	1	7.1	0	0.0	8	22.9
1- 25	4	36.4	7	50.0	3	30.0	14	40.0
26-100	0	0.0	6	42.9	7	70.0	13	37.1
Total	11	100.0	14	100.0	10	100.0	35	100.0

TABLE C.12

PLANNING EFFORT BY IMPLEMENTATION
SUCCESS

Number of Topics Covered by the Plan[a]	Level of Implementation Success						Total	
	Low		Moderate		High			
	#	%	#	%	#	%	#	%
0	5	45.5	2	14.3	1	10.0	8	22.9
1 - 4	3	27.3	3	21.4	1	10.0	7	20.0
5 - 8	3	27.3	9	64.3	8	80.0	20	57.1
Total	11	100.1	14	100.0	10	100.0	35	100.0

[a]The eight policy-specific planning topics include the following: (1) recruitment; (2) screening, testing, diagnosis; (3) involvement of parents; (4) classroom programs; (5) special materials and equipment; (6) special physical facilities; (7) monitoring and evaluation of the implementation; and (8) others.

TABLE C.13

COORDINATION EFFORT BY IMPLEMENTATION
SUCCESS

Level of Coordination Effort[a]	Level of Implementation Success						Total	
	Low		Moderate		High			
	#	%	#	%	#	%	#	%
0	0	0.0	2	14.3	0	0.0	2	5.7
1	6	54.5	4	28.6	3	30.0	13	37.1
2	5	45.5	8	57.1	7	70.0	20	57.1
Total	11	100.0	14	100.0	10	100.0	35	99.9

[a]See Appendix B for definitions of the codes that appear in this column.

TABLE C.14

EFFORT TO MOBILIZE RESOURCES BY
IMPLEMENTATION SUCCESS

Level of Mobilization Effort[a]	Level of Implementation Success						Total	
	Low		Moderate		High			
	#	%	#	%	#	%	#	%
0	7	63.6	0	0.0	0	0.0	7	20.0
1	0	0.0	1	7.1	0	0.0	1	2.9
2	4	36.4	11	78.6	3	30.0	18	51.4
3	0	0.0	0	0.0	6	60.0	6	17.1
4	0	0.0	2	14.3	1	10.0	3	8.6
Total	11	100.0	14	100.0	10	100.0	35	100.0

[a]See Appendix B for definitions of the codes that appear in this column.

TABLE C.15

EFFORT TO OBTAIN ADDITIONAL FUNDS
BY IMPLEMENTATION SUCCESS

Program Received Additional Funds?	Level of Implementation Success						Total	
	Low		Moderate		High			
	#	%	#	%	#	%	#	%
No	10	90.9	6	42.9	4	40.0	20	57.1
Yes	1	9.1	8	57.1	6	60.0	15	42.9
Total	11	100.0	14	100.0	10	100.0	35	100.0

TABLE C.16

EFFORT TO ADD NEW STAFF BY
IMPLEMENTATION SUCCESS

Staffing of the Implementation	Level of Implementation Success						Total	
	Low		Moderate		High			
	#	%	#	%	#	%	#	%
No New Staff	10	90.9	7	50.0	4	40.0	21	60.0
Some New Staff	1	9.1	7	50.0	6	60.0	14	40.0
Total	11	100.0	14	100.0	10	100.0	35	100.0

TABLE C.17

TRAINING AND TECHNICAL ASSISTANCE EFFORT
BY IMPLEMENTATION SUCCESS

Number of Train-ing and Technical Assistance Activities	Level of Implementation Success						Total	
	Low		Moderate		High			
	#	%	#	%	#	%	#	%
0	2	18.2	0	0.0	0	0.0	2	5.7
1	2	18.2	0	0.0	1	10.0	3	8.6
2	2	18.2	1	7.1	1	10.0	4	11.4
3 or more	5	45.5	13	92.9	8	80.0	26	74.3
Total	11	100.1	14	100.0	10	100.0	35	100.0

TABLE C.18

MODIFICATION OF FACILITIES BY
IMPLEMENTATION SUCCESS

Status of Physical Facilities	Level of Implementation Success						Total	
	Low		Moderate		High			
	#	%	#	%	#	%	#	%
Not Modified	11	100.0	12	85.7	4	40.0	27	77.1
Modified	0	0.0	2	14.3	6	60.0	8	22.9
Total	11	100.0	14	100.0	10	100.0	35	100.0

LIST OF REFERENCES

LIST OF REFERENCES

Babbie, Earl R. <u>Survey Research Methods</u>. Belmont, California: Wadsworth Publishing Co. Inc., 1973.

Bailey, Stephen K., and Mosher, Edith, K. <u>ESEA: The Office of Education Administers a Law</u>. Syracuse, N.Y.: Syracuse University Press, 1968.

Braybrooke, David and Lindblom, Charles E. <u>A Strategy of Decision</u>. New York: The Free Press, 1963.

Bunker, Douglas R. "Policy Sciences Perspectives on Implementation Processes." <u>Policy Sciences</u>, III, 1972.

Campbell, Donald T., and Stanley, Julian C. <u>Experimental and Quasi-Experimental Designs for Research</u>. Chicago: Rand McNally and Co., 1963.

Caro, Francis G., ed. <u>Readings in Evaluation Research</u>. New York: Russell Sage Foundation, 1971.

Cummings, L. L., and Schwab, Donald P. <u>Performance in Organizations</u>. Glenview, Illinois: Scott, Foresman and Company, 1973.

Derthick, Martha. <u>New Towns In-Town</u>. Washington, D.C.: The Urban Institute, 1972.

Dror, Yehezkel. <u>Public Policymaking Reexamined</u>. Scranton, Pennsylvania: Chandler Publishing Company, 1968.

Dye, Thomas R. <u>Understanding Public Policy</u>. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1972.

Easton, David. <u>A Framework for Political Analysis</u>. Englewood, Cliffs, N.J.: Prentice-Hall, Inc., 1965.

Gergen, K. "Assessing the Leverage Points in the Process of Policy Formulation." <u>The Study of Policy Formation</u>. Edited by R. A. Bauer and K. H. Gergen. New York: The Free Press, 1968.

Glaser, Barney G., and Strauss, Anselm L. <u>The Discovery of Grounded Theory</u>. Chicago: Aldine Publishing Co., 1967.

Haveman, Robert H., and Margolis, Julius. Public Expenditures and Policy Analysis. Chicago: Markham Publishing Co., 1970.

Head Start Project Team, Syracuse University. "Interim Report on Assessment of the Handicapped Effort in Experimental and Selected Other Head Start Programs Serving the Handicapped." Report submitted to HEW by Syracuse University, Syracuse, N.Y., 1974.

_____. "Final Report on Assessment of the Handicapped Effort in Experimental and Selected Other Head Start Programs." Report submitted to HEW by Syracuse University, Syracuse, N.Y., 1974.

_____. "Statement on Policy Recommendations to the Office of Child Development." Statement prepared by the members of the SRI-Syracuse University Project Senior Consultant Group, Syracuse, N.Y., 1974.

Heclo, Hugo H. "Policy Analysis." British Journal of Political Science, II (January, 1972).

Holt, Robert T., and Turner, John E., eds. The Methodology of Comparative Research. New York: The Free Press, 1970.

Johnson, J. Econometric Methods, 2nd Edition. New York: McGraw-Hill, 1972.

Jones, Charles O. An Introduction to the Study of Public Policy. Belmont, California: Wadsworth, 1970.

Kaplan, Abraham. The Conduct of Inquiry. San Francisco: Chandler Publishing Co., 1964.

Laswell, Harold and Kaplan, Abraham, Power and Society. New Haven, Connecticut: Yale University Press, 1950.

LaVor, Martin L. "Economic Opportunity Amendments of 1972, Public Law 92-424." Exceptional Children. (November, 1972).

Lawler, Edward E. Motivation in Work Organizations. Monterey, California: Brooks/Cole Publishing Company, 1973.

_____. Pay and Organizational Effectiveness. New York: McGraw-Hill Book Co., 1971.

Lindblom, Charles E. The Policy-Making Process. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1968.

Marshall, Dale R. "Implementation of Federal Poverty and Welfare Policy: A Review Essay." Policy Studies Journal. II (Spring, 1974).

Merton, Robert K. On Theoretical Sociology. New York: The Free Press, 1967.

Porter, Lyman W., and Lawler, Edward E. Managerial Attitudes and Performance. New York: Dorsey-Irwin, 1968.

Pressman, Jeffrey L., and Wildavsky, Aaron. Implementation. Berkeley, California: University of California Press, 1973.

Ranney, Austin, ed. Political Science and Public Policy. Chicago: Markham Publishing Co., 1968.

Reynolds, James F. "Policy Science: A Conceptual and Methodological Analysis." Policy Sciences. VI (March, 1975).

Robinson, James. Congress and Foreign Policy-Making. Homewood, Illinois: Dorsey Press, 1962.

Rossi, Peter H., and Williams, Walter, eds. Evaluating Social Programs. New York: Seminar Press, 1972.

Rourke, Francis E. Bureaucracy, Politics and Public Policy. Boston: Little, Brown and Co., 1969.

Selznick, Philip. TVA and the Grass Roots. Berkeley, California: University of California Press, 1949.

Sharkansky, Ira, ed. Policy Analysis in Political Science. Chicago: Markham Publishing Co., 1970.

Siegel, Sidney. Nonparametric Statistics. New York: McGraw-Hill, 1956.

Smith, Marshall S., and Bissell, Joan S. "Report Analysis: The Impact of Head Start." Harvard Education Review. 40 (Winter, 1970).

Smith, Thomas B. "Policy Roles: An Analysis of Policy Formulators and Policy Implementors." Policy Sciences, IV, 1973.

_____. "The Policy Implementation Process." Policy Sciences, IV, 1973.

Stinchcombe, Arthur L. Constructing Social Theories. New York: Harcourt, Brace and World, Inc., 1968.

Suchman, Edward A. Evaluative Research. New York: Russell Sage Foundation, 1971.

Sundquist, James L., and Davis, David W. Making Federalism Work. Washington, D.C.: The Brookings Institution, 1969.

Systems Research Incorporated. "Response to RFP 53-73-HEW-OS-- Evaluation of the Handicapped Effort in the Head Start Program." Proposal submitted to HEW by Systems Research Incorporated, Lansing, Michigan, 1973.

_____. "The Status of Handicapped Children in Head Start--Final Report Based on Surveys of Summer 1973 and Full Year 1973-74

Head Start Programs." Report submitted to HEW by Systems
Research Incorporated, Lansing, Michigan, 1974.

U.S. Department of Health, Education and Welfare. Head Start Child
Development Programs: A Manual of Policies and Instructions.
Washington, D.C.: Office of Child Development, 1967.

_____. Head Start Services to Handicapped Children--First Annual
Report of the U.S. Department of Health, Education and Welfare
to the Congress of the United States on Services Provided to
Handicapped Children in Project Head Start. Washington, D.C.:
Office of Child Development, 1973.

_____. OCD Notice N-30-331-1 on Head Start Services to Handicapped
Children. Washington, D.C.: Office of Child Development, 1973.

_____. OCD Notice N-30-334-1 on Program Options for Project Head
Start. Washington, D.C.: Office of Child Development, 1972.

_____. OCD Notice N-30-364-1 on Head Start Program Performance
Standards. Washington, D.C.: Office of Child Development,
1973.

_____. Project Head Start 1969-70: A Descriptive Report of Programs
and Participants. Washington, D.C.: Office of Child Develop-
ment, 1972.

_____. Project Head-Start Statistical Fact Sheet--Fiscal Year 1972.
Washington, D.C.: Office of Child Development, 1973.

Vroom, Victor H. Work and Motivation. New York: John Wiley and Sons,
Inc., 1964.

Weiss, Carol H. Evaluation Research. Englewood Cliffs, N.J.: Prentice-
Hall Inc., 1972.

Wholey, Joseph S., et al. Federal Evaluation Policy. Washington, D.C.:
The Urban Institute, 1971.