

POLICY IMPACT EVALUATION

A report
submitted to: Mr. Michael Zak
PPC/E
U.S. Agency for International Development
Washington, D.C. 20523

By: Jerry B. Jenkins
Sequoia Institute
1822 21st Street, Suite 200
Sacramento, CA 95814

In completion
of: Contract No. OTR-0000-0-00-3363-00

Date of
submission: April 17, 1984

TABLE OF CONTENTS

	PAGE
I. RECONNOITERING A GOAL: POLICY IMPACT EVALUATION	1
A. Introduction	1
B. Policy Impact: Subject of Profuse Neglect	2
C. Policy Implementation: Missing Link for Policy Impact	4
D. Events Caused...Consequences Felt: Organizations, Implementation, Impact, and Learning	6
II. POLICY IMPLEMENTATION -- STILL ON ITS HEAD	9
A. The Real Problem With Street Level Policy Making	10
B. Means By Which The Problem Cannot Be Resolved	11
1. Act like the problem does not exist	11
2. Contribute to the problem	12
3. Redefine roles (obviating the seperation of powers)	13
4. Redefine more roles (obviating federalism)	18
5. Act like ambiguity is THE problem	21
6. Act like experimentation is THE answer	23
7. Act as if the policy situation provides the answer	31
C. The Real Costs of the Headstand	35
III. TOWARD STANDING POLICY IMPLEMENTATION ON ITS FEET	38
A. Explaining the Ascendance of Street Policy	39
B. Organizational Reification of Street Policy Ascendance	42
C. Reflection, Choice, and Organizational Change	46
IV. POLICY IMPACT: CONCEPTUALIZATION FOR EVALUATION	53
A. Implementation of What?	54
B. Implementation by Whom? Do They Agree on What?	56
C. Measurement for Evaluation	59
D. Policy Implementation Evaluation: Beyond Feasibility	75

I. RECONNOITERING A GOAL: POLICY IMPACT EVALUATION

A. INTRODUCTION

This report provides the United States Agency for International Development with a suggested methodology for evaluating the extent to which its development policies--what they say--are translated into what its representatives do in host countries.

At the outset, then, "policy impact evaluation" is differentiated from "impact evaluation". The latter is engaged in on a routine basis, yielding veritable mountains of information regarding consequences of whatever is done in the field. It does not, however, assess the correspondence between what and whatever--between what was to occur, as envisioned by policy producers, and whatever did occur, as a result of actions by policy implementers.\1

The interest here in policy impact evaluation does not, however, deny the existence of "whatever policy"--

\1References to policy producers, implementers, and evaluators appear throughout this paper, even as it is recognized that individuals engaged in the policy process will be doing some of all three activities. This usage simply reflects that primary responsibilities of different groups of individuals -- usually in accordance with organizational differentiations.

preimplementation policy so nebulous and vague that whatever happens in the field should not be unexpected. Instead, the methodology presented in the fourth chapter of this report would, if implemented, enable both the identification of such policies and their differentiation from others.

In sum, the methodology of Chapter IV enables the identification of which policies, and which components of the respective policies, can be expected to have the greatest (and least) impact on actions and results in the field. This accomplishment fulfills the most basic aim of the report. The experience of getting to this point, however, was scarcely anticipated.

B. POLICY IMPACT: SUBJECT OF PROFUSE NEGLECT

The report's initial strategy called for surveying, borrowing, and adapting. More specifically, techniques for evaluating policy impact were to have been extracted from the vast literature incorporating evaluations of domestic public policy. An obvious assumption of this strategy was that such a survey would yield the desired evaluation techniques. Then, the only task of the report requiring any degree of creativity was to have been the adaptation of those "domestic" techniques to the "foreign" assistance policy needs of the Agency. So much for strategy.

The very vastness of the "evaluation" literature accentuates

its neglect of "policy impact". The more extensive the search, and more fervent the effort to discern, the more does that neglect suggest conscious avoidance--and, the more does discovery of "policy impact evaluation" seem a needle-in-haystack pursuit of an ever-illusory objective.

The "policy impact" void in the literature of "evaluation" is even more remarkable in light of the extensive play given to the purported effects of policies, and of policy changes, in the mass media. Of course, reconciliation of this disparity in the treatment of policy impact is enabled by differentiating between "evaluation" and "attribution". The latter characterizes mass media assessments of policy impact.

Thus, one can expect any one or more representatives of the media to attribute a wide variety of effects (even incompatible effects) to any given policy or policy change. Thus, too, one can expect to find in media attributions a wealth of predispositions midst a methodological desert. At the other end of the spectrum, those steeped in the methodology of science, and engaged in evaluating almost everything, scarcely ever address policy impact.\2

\2The most distinct exception to this judgment is embodied in much of the work predicated on quasi-experimental designs. This exception, however, has extremely limited utility for evaluating the impact of most policies, as explained in the following chapter of this report.

Juxtaposing the disparate treatment of "policy impact" by mediaship and scholarship might call into question the very utility of the methods and techniques of the social sciences. While acknowledging that the latter are frequently misplaced, Chapter IV of this report embodies its own obvious answer. The root of that conclusion is summarized in a single word: "implementation".

C. POLICY IMPLEMENTATION: MISSING LINK FOR POLICY IMPACT

A missing variable in virtually all media assessments of public policies and their consequences is that of "policy implementation". Thus, it is not uncommon to find representatives of the media attributing effects to policies that have never been implemented. This failing might be at least partially explained by the neglect of implementation, until very recently, in the scholarship of the social sciences.

Upon publication of Implementation, in 1973, the significance of the subject could be ignored no longer. The change that has ensued is reflected in Aaron Wildavsky's preface to the second edition (in 1979) of that seminal work:

'Previewers' of the first edition took us to task for failing to integrate our findings into the vast literature on the subject. So we tried to prove a negative--namely, that at the time we wrote there was little such literature. Today that is no longer true; the literature is growing

so rapidly that the bibliography (replacing the appendix saying there was 'no there there') can only be suggestive rather than comprehensive. (Pressman and Wildavsky, xv)

The scant treatment of policy impact in the literature needs no further explanation. Without assessing the implementation of policies, evaluations of their impact could only be inadequate. Furthermore, the absence of the one legitimated the nonexistence of the other, and the neglect of policy impact could be viewed as responsible scholarship.

But this explanation is only good for the past. What about now, after a decade in which the burgeoning literature on implementation is going off the charts? Why is there still such a dearth of policy impact evaluations in the literature?

The answer rests in what has been discovered by the new focus on implementation: Not policy implementation, but policy ... and, then, implementation. In short, the void in the literature is an accurate reflection of the observable world. Descriptive social science can do no more; thus does practice shape scholarship. But if the products of scholarship suffer from the practices they describe, so too do the organizations whose behavior is characterized by those practices.

**D. EVENTS CAUSED...CONSEQUENCES FELT: ORGANIZATIONS,
IMPLEMENTATION, IMPACT, AND LEARNING**

If an organization formulates policy, but its implementation does not occur, how can the organization and its representatives learn? How can either policy or implementation be improved in the future?

As the closest domestic counterpart of A.I.D., the experience of the Economic Development Administration should be instructive.^{\3} Pressman and Wildavsky directly relate the deficiencies of EDA's public works programs in Oakland to the lack of organizational learning. Fortunately, they go beyond this descriptive malaise, offering an evaluative conclusion of prescriptive significance--after dispensing with other commonly alleged solutions to the problem ("false messiahs," they call them) such as:

bureaucracy and coordination--whose invocation only serves to obscure problems. Now it is time to bring in a real devil--the divorce of implementation from policy. Learning fails because events are caused and consequences are felt by different organizations. (Pressman and Wildavsky, 135)

With the benefit of this prescriptive cue, the methodology of Chapter IV is more than one of, and for, evaluation. Its use by

^{\3}An appreciation of the mutually instructive potential of the two organizations' experiences is manifest in a section of Pressman and Wildavsky's book entitled "Foreign Aid in America".

the Agency would direct the attention of both implementers and policy producers to their own divorce. By its employment, those who cause and those who experience consequences might come to share the same meanings and intend the same things.

The central premise of the suggested methodology is this: It must simultaneously provide a means for enhancing organizational knowledge and learning while evaluating policy impact. The circularity inherent in this premise is not only acknowledged, it is intended. Indeed, the circularity being promoted is no more than what is supposed to occur in the pursuit of organizational goals: formulation of policy; then implementation of that policy; then evaluation of that policy and its implementation; then reformulation of policy, etc.

Of course, the learning that is assumed for any reformulation of policy is critically dependent on the actual implementation of prior policy. There lies the rub. Anyone who doubts the divorce of implementation from policy must possess information that is either highly classified or imagined.

Chapter II comprises a literature assessment which illustrates the rupture between policy and its implementation. It contains much that revolves around, without ever quite focusing on, policy impact. With respect to the latter, it presents at least seven "false messiahs". Each is seen to direct (or misdirect) attention to different aspects of the policy process while sharing in the

failure to conceptualize "policy impact".

Chapter III, in turn, characterizes that failure and the divorce of implementation from policy as a "natural" byproduct of the growth of public organizations. It then makes a case for reconceptualizing public organizations in order to bring the causes and consequences of policy together--to an extent sufficient for the implementation of policy and the realization of policy impact.

Thus, the methodology proposed in Chapter IV can be employed for evaluation only as it affects organizational process--when policy impact is the subject of evaluation, anything less cannot suffice.

II. POLICY IMPLEMENTATION -- STILL ON ITS HEAD

When the study of public policy implementation was stood on its head 15 years ago, (Lipsky) it constituted a striking occurrence. Today, after that position has been further substantiated by a wave of implementation assessments (Bardach; Berman; David; Majone and Wildavsky; Pressman and Wildavsky; Rein, to name a few), and contradicted by none, "street level bureaucrat" has become part of the vernacular for depicting an individual who "makes" policy during the course of field implementation. (Williams, 17)

Now when the "study" of anything can be stood on its head, that constitutes prima facie evidence of never having been on its feet. Such a reversal indicates that "experts" have been both misunderstanding for themselves, and misrepresenting for others, the subject matter for which they have presumed expertise.

Unfortunately, the recognition of street level bureaucracy did not result from any improvement in conceptualization of the public policy process. Instead, policy making in the field had become so common as to demand its acknowledgment by students of public policy. In short, students of public policy have not become much smarter; street level policy production has simply become much more common.

A. THE REAL PROBLEM WITH STREET LEVEL POLICY MAKING

The importance of street level policy making for the present endeavor cannot be avoided. Briefly, if we do not know where policy is made, how can we discern its impact? Where do we begin our evaluation? And the answers to these questions might vary widely, among policies and projects. Inability to answer these questions has ramifications for the entire policy process.

Only as pre-implementation policy is known to influence its enactment in the field can evaluations of outcomes be meaningfully employed in any evaluations and reformulations of policy.

To the extent that policy is actually made in the field, prior policy is irrelevant to its consequences; yet, that prior policy might be modified, sustained or rejected on the basis of evaluations of outcomes on which it has had no influence. In other words, if we do not know the location(s) in which any given policy is made, we cannot know what we are evaluating--even when we have mountains of outcome data. This reflects a situation of grave inherent danger, in which individuals cannot behave responsibly because they will have been deprived of choice even as they think they have it.

It must be understood that the problem is not street level policy making; the problem, instead, is not knowing when, where, and to what extent, it is occurring. Likewise, no judgment is being made here as to the relative quality of street policy and

pre-implementation policy. The former might actually improve the consequences for intended beneficiaries, but this cannot be known in lieu of being able to discern the impact of pre-implementation policy.

B. MEANS BY WHICH THE PROBLEM CANNOT BE RESOLVED

1. Act like the problem does not exist. This is perhaps the most common response to the policy impact problem. Apart from reasons that might be self-serving, a void in the conceptualization of public policy is the most readily apparent reason for nontreatment of policy impact (possible reasons for the void itself are detailed throughout this chapter).

Thus, for example, one of the better introductions to "public policy analysis" never addresses the question of how we might discern the extent to which policy is actually reflected in programs and projects. As might be expected, "implementation" does not even appear in the book's index. (Dunn, 1981) It remains a useful text for addressing policy formulation and outcome evaluation questions, but by ignoring policy impact the extent to which the evaluated outcomes are those of pre-implementation policy is neglected entirely.

The book should be retitled: Public Policy -- A Partial Analysis. To be fair, all of the book's competitors should follow suit.

2. Contribute to the problem. Attempts by individuals to influence the implementation of policy normally reflect their positional interests. (Rein, 134) Hence, any enhancement of policy impact is coincidental with the primary interests of an individual's position in the policy process. For illustration, the following should suffice:

To repeat: the assessment of program design and implementation identifies the action needed to enhance program evaluability. The program manager is the person who must take action to alter the design and implementation of the program, giving precedence to evaluation requirements. The evaluator faces a very difficult task in attempting to bring about such changes because the organizational and 'political' context is not normally conducive to altering program design and controlling its implementation to facilitate the evaluation. (Rutman, 177)

Implementation's divorce from policy, then, is accompanied by no small degree of tension with evaluation. Between evaluators seeking implementation actions that will enhance "program evaluability", and producers calling for those that will maximize "policy impact", implementers are caught in a cross-fire that can itself account for the proliferation of street policy. Thus do positional interests compound the policy impact problem.

At the same time, however, the vigorous competition that transpires between implementers and evaluators over the content of policy can at least enhance the awareness of policy producers as to when and where street level policy is occurring.

On the other hand, differing positional interests provide no guarantee of such visible competition, and in some instances, evaluation objectives become policy (sidewalk policy?)...

...One celebrated example is the evaluation of the Head Start Program. Evaluators knew something about the measurement of cognitive competence but much less about social competence. As a result, they imposed the objective of the development of intelligence on the program that enabled them to evaluate the program with the tools they had on hand. (Rein, 135)

Acting as if such products of positional interests are a consequence of the U.S. Constitution, the General Accounting Office (GAO) has devised a model by which that document might be effectively negated.

3. Redefine roles (obviating the separation of powers).

There is no doubt that positional interests must be redefined as a part of the conceptualization required for standing public policy implementation, its impact, and their evaluation, on their feet. Such redefinition is directly pursued in the following chapters. Avoided, however, will be anything approximating the GAO's stab at redefining positional interests. Its suggestions for congressional oversight of public programs obliterates the very meaning of different positions (not to mention that of the Constitution).

Acting like policy ambiguity is the problem, and hence, anticipating the fifth means by which the policy impact problem cannot be resolved,

The GAO recommends that Congress monitor implementation as it occurs rather than checking a program for conformity with congressional intent after it has been implemented. This recommendation is based on an understanding that the implementation effort itself often changes the goals, resources, and strategies that Congress originally had intended. In most cases, Congress and other legislative bodies cannot state clear goals, specify administrative structures, or designate implementation strategies when laws are formulated. In fact, it may be important for Congress to avoid overspecifying 'legislative language or goals, when knowledge or political reality does not permit' (GAO, 1977, p.10) (Sorg, 142-144)

This statement is so exceptional as an exercise in obfuscation that it shall be referred to at several junctures in the next few pages. It is akin to the "laissez faire" whatever-is-implemented-is-policy orientation that is subsequently discussed, with one critical difference--legislators become implementers. Of course, they also become evaluators as they engage in ensuring conformity with (discovering?) legislative intent.

A picture and summary table of these suggestions for the legislative branch of government to become the executive as well are provided by Figure 2-1 and Table 2-1. Even though the constitutional separation of powers would go by the boards, these suggestions do at least manifest recognition of a policy impact problem. They would define the problem away by redefining governmental roles. Following these suggestions, the legislative specification of policy for implementation is continuous. The

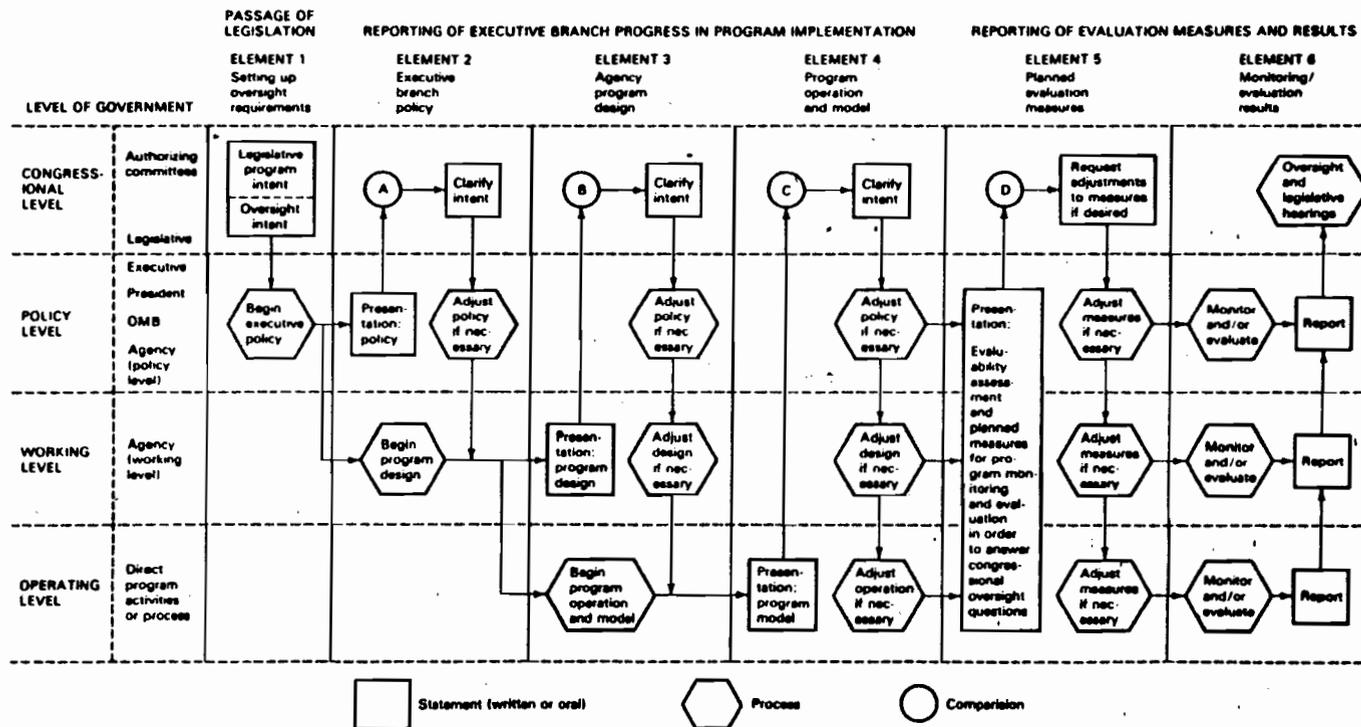
GAO would have it Casey Stengel's way--"it ain't over til its over". When it is over, then we will know what policy is (was), and we will know it had impact, or at least experienced it.

Table 2-1: Suggested Oversight Process Elements Portion of Oversight Feedback Loop (Sorg, 145)

<i>Element Number</i>	<i>Implementation of a Program</i>	<i>Feedback of Information</i>	<i>Congressional Response/Requirements</i>
1	n/a	n/a	Include a statement of legislative and oversight intent in the enabling act or accompanying reports
2	Formulation of executive-branch policy and strategy for carrying out the enabling act's intent	Presentation of executive-branch policy	(Point A) Clarify intent and request policy adjustments if desired
3	Planning, design, and development of an operating program by agency working level	Presentation of agency progress in program design	(Point B) Clarify intent and request policy and/or program design adjustments if desired
4	Establishment and initial execution of an operating program; model the actual program operation	Presentation of agency model of the operating program	(Point C) Clarify intent and request policy, program design, and/or program operation adjustments if desired
5	Perform evaluability assessment and develop planned evaluation measures	Presentation of evaluability assessment and planned evaluation measures	(Point D) Request adjustments to planned evaluation measures if desired
6	Conduct program evaluations and monitoring	Report results of the program evaluations and monitoring	Assess program results; amend, extend, or terminate enabling act; develop and include a new statement of legislative intent if appropriate

Source: From U.S. General Accounting Office, "Finding Out How Programs Are Working: Suggestions for Congressional Oversight," PAD-78-3, November 22, 1977, p. 23.

Figure 2-1: Flow Diagram of the Government Accounting Office's Suggested Oversight Process (Sorg, 143)



Source: U.S. General Accounting Office, "Finding Out How Programs Are Working: Suggestions for Congressional Oversight," PAD-78-3, November 22, 1977, p. 24.

Though the GAO model evokes greater concern for legislative control than for compliance with any initial intentions, its effect in the field would probably magnify the importance of compliance, which is, of course, "only one aspect of an effectively implemented policy". (Nagel, 1980, 32) The continuous monitoring of projects that would be entailed would be more likely to instill a "meeting the requirements" perspective among field personnel than it would a concern for maximizing the quality and performance of projects. (Goodrich, 49)

Furthermore, implementation of such a model would further decrease the burden of justification for passing new legislation, as legislators would know that "intent" could be discovered, or they could effectively rescind their votes, during the course of implementation. Even without such a model, "...Explaining policy failure in terms of the quality of the policy idea seems to work at least as well as explaining it in terms of social forces capturing and corrupting the implementation process." (Majone and Wildavsky, 182) With such a model, the quality of policy ideas might be expected to descend to new depths of impoverishment.

As if to add insult to injury, the GAO's suggestions would hike the cost of government by untold millions. In contrast, others are recommending that authority for the deobligation and reobligation of project funds be returned by the Congress to AID (per the pre-1977 status), that the appropriate Congressional

oversight of Agency performance is at the level of country/regional development and not at the level of specific projects, and even that the budget submission cycle should be two-year rather than annual (saving about 200 work-years of staff time). (President's Private Sector Survey on Cost Control, 61-64)

As a general rule, those who increase the cost of government should be prepared to demonstrate commensurate increases in beneficial results. Can GAO demonstrate that development efforts have been enhanced in any way since the Congress assumed authority for the deobligation and reobligation of project funds, at an annual cost of about two million dollars?

Disdain for federalism should be an expected corollary of disregard for the separation of powers. Fulfillments of this expectation inspire thoughts of Wonderland, for, as Alice said, they get "curiouser and curiouser". Simply consider the following assessment of the merits of categorical, relative to block, grants.

4. Redefine more roles (obviating federalism). While recognizing that the costs which the categorical grant system "imposes in administrative inflexibility and fragmentation are real", a GAO official writes that the main problem with the block grant alternative is making them accountable. (Chelimsky, 118)

The first of three issues involved in "accountability for what?", she advises, is,

Accountability for achieving the national objective in funding the program; this issue focuses on (a) whether the program has followed the legislative intent (with regard to activities and target populations, if specified, for example) and (b) whether it has proven effective in meeting the broad national objective. (Chelimsky, 111)

Another observation tells us considerably more about what the "national objective" is likely to entail:

The block grant concept could force constituencies such as the disadvantaged and handicapped to compete with each other for federal funds. However, other constituencies--the traditionally strong ones in individual states--could quickly develop the same stranglehold on federal funds in block form as the national lobbies exercise on federal funds in categorical form. (Chelimsky, 104)

Though hardly enamored of national stakeholders (92-94), Chelimsky apparently prefers them to state and local stakeholders, because of "program integrity" which "would suffer under block grants unless a major effort at coordination were made by the national agency responsible..." (106) This appears to imply that program "integrity" has "consistency" or "uniformity" among state and local block grants as its primary objective. These, in turn, would conform to the "national objective" (and funding) on which national stakeholders have a "stranglehold". \4

\4If one goes beyond these judgments to consider some actual theory and evidence, then national stakeholders bear a disturbing resemblance to the "redistributive coalitions" associated with the decline of nations. (Olson)

One can only wonder at how the federalism embodied in the U.S. Constitution has come to be a lesser "national objective" than is the latest policy from national stakeholders. Former Budget Bureau Director, Charles Schultze, describes how policy emanates from the interplay of such groups:

The Elementary and Secondary Education Act of 1965...was enacted precisely because it was constructed to attract the support of three groups, each with different ends in view. Some saw it as the beginning of a large program of federal aid to public education. The parochial school interests saw it as the first step in providing...financial assistance for parochial school children. The third group saw it as an anti-poverty measure. If there had been any attempt to secure advance agreement on a set of long-run objectives, important elements of support for the bill would have been lost and its defeat assured. (Rutman, 53)

Furthermore, any expectation that a more distinct national objective, or discovery of legislative intent, is to be accomplished during the implementation of such legislation is unrealistic, for "Interventions by stakeholder groups at the program implementation stage have been equally powerful...." (Chelimsky, 93)

Were the GAO's oversight suggestions to be followed, why would anything qualitatively different be expected from the implementation period than from the legislative interval? Indeed, why would not such implementation monitoring by the Congress simply increase the compliance of final "implemented policy" with stakeholder interests as modified by field experience?

"Who's compliance?" becomes a relevant question when all of these considerations are taken into account, and members of the Congress appear to be less trustees than instructed delegates. So what would that make of domestic policy agencies? Control agents?... "A key factor in the field creating complexity and confusion has been a 'regulatory' attitude in the social agencies. This regulatory mentality equates control with compliance...." (Williams, 109)

The GAO compliance model is, in fact, a surrogate for a parliamentary form of government (but without a nationally elected leadership). It is also, of course, a mandate for greater centralized control. Its presumption is that such a model is necessary in order to ensure the realization of legislative "intent" by implementation. In fact, it would increase the compliance of the legislative formulation with the realities (as defined by stakeholders) of the world in which words might become deeds. In fact, also, we cannot be certain that the primary legislative "intent" is at variance with an intent to comply with the right (for reelection) combination of national stakeholder interests.

5. Act like ambiguity is THE problem. The problem of ambiguity in the expression of policy is reflected by most of the "means" that are discussed in this section. Some of the means

require decreased policy ambiguity, others desire it, and still others seem to just want to control it.

Ambiguity in legislative or administrative policy is most definitely a problem--especially for evaluation. The more ambiguous is policy, the more difficult it must be to discern policy impact.

Indeed, where implementation is frequently characterized as translating thought into action, the more common experience today might be thought attempting to catch up with action. This is the order of things that makes most sense of the GAO compliance model. Rather than ensuring compliance of implementation with what is expressed by policy, it seeks to ensure that whatever happens in the field will be acceptable to the Congress--after whatever occurs, we might come to know what is what.

However, putting thought and action in their preferred order does not provide immunity from other afflictions. Thus, to treat ambiguity as the problem for resolution without at the same time recognizing it as a result of other influences is akin to allowing a disease to spread while treating its symptoms. Policy ambiguity is necessarily perpetuated by failure to respond to its sources, and this is the case regardless of the energy that might be devoted to pursuing its alternatives. This is suggested by Alex Radian in a letter to Aaron Wildavsky:

....It is much like the error that public finance theorists have been making for years, if we could only find optimal tax structures policy makers will adopt them. Your notion of simplicity suffers from this weakness. In the tax field policy makers are forced to make increasingly complex policies because that helps to make the tax burden vague and because every pressure group demands its own special benefit. True it is more difficult to implement complex policies, but unless the rules of politics are changed it is not possible to make simple policies. (Wildavsky, 176)

Radian's observation points us in the right direction, but we risk falling into a parallel trap of "if politics, then ambiguity". In order to account for dramatic increases in policy ambiguity and its corollary, street level bureaucracy, we cannot simply attribute it to politics. If we were to try, it would be necessary to specify which rules of politics have changed, and/or how inputs into the political process have altered.

Furthermore, cannot the "stranglehold" on public policies by "stakeholders" be better understood as a consequence of rules against, rather than of, politics?^{\5} Of course, "if anti-politics, then ambiguity" is also insufficient to the task, though the perspective of experimentation, and perhaps the GAO, suggests a more redeeming role for anti-politics.

6. Act like experimentation is THE answer. Beginning with the assumption that experimentation is the "only available route

^{\5}Our conception of "politics" is acutely influenced by one of its most enduring tributes -- Bernard Crick's In Defence of Politics.

to cumulative knowledge...." (Campbell and Stanley, 1963, 3), then two other beliefs become entirely reasonable: (1) the "ambiguity revolution" is a result of too little experimentation and/or insufficient receptivity to experimental results, and (2) "responsible" government will pay for extensive experimentation, and maximize the use of its results in the formulation of public policy.

These appear to be the conclusions of the leading advocate of the "experimenting society". He expresses doubts that "political conditions" will ever make it possible to have "a dialectic of experimental arguments" in applied social science, since "the stakes and motivational structures of political domination place the powerful participants farther from the 'standards of appraisal and the incentive structure of the idealized scientific community.'..." (Campbell, 1982, 331-332) His reference to political "domination", rather than "life", tells us that his judgment is more severe than one which simply recognizes that people occupying different positions are likely to have different interests. Should there be any doubt, he adds that,

....An established power structure with the ability to employ applied social scientists, the machinery of social science, and control over the means of dissemination produces an unfair status quo bias in the mass production of belief assertions from the applied social sciences.... (Campbell, 1982, 335)

Putting all of these beliefs together, several dangers become readily apparent. First, the trap of "if politics, then ambiguity" has ensnared the social experimenter. Second, politics is posited as anti-truth (with experimentation, of course, being truth's best friend). Third, government is deemed "irresponsible", or worse, because it is neither sufficiently supportive nor responsive to quasi-experimentation in society. Fourth, the large-scale social experimentation that has occurred, such as the income maintenance experiments, cannot be explained. Fifth, scant appreciation is evident of the greatest advances in social relationships and public policies that have been achieved on this planet--all without benefit of any quasi-experimental design or previous experimental evidence. Sixth, revealed is a failure to recognize both the social experimenter's own positional interest, and the possibility that, with some transposition, "the stakes and motivational structures of social experimenters place their fervent belief in the procedures of an idealized scientific community in inherent conflict with the procedures of an open political system that cannot, by its very definition, exclude other interests and beliefs from its competition".

None of the preceding is intended to suggest that social experimentation is never appropriate or useful. When we know there is a distinct proposition upon which policy choice rests, and we know that it is amenable to testing by social

experimentation, and there is no reason to believe that target groups to be included in the test will be worse off from the conduct of the experiment than they would otherwise be, then the design and implementation of an experimental project has strong justification. The rub, of course, is in knowing all these things in order to provide that justification.

The social experimenter's answer is likely to be that we must engage in experimentation in order to acquire that knowledge, but this answer obviously begs the questions. The most critical problems of societies are--not coincidentally--the most "ill-structured...where the main difficulty lies in defining the nature of the problem rather than determining through selective experimental interventions the most effective reform to alleviate it." (Dunn, 1982, 300)

Thinking of policy changes, or "reforms", as experiments (Campbell, 1972) does not beg all of the preceding questions, though we are pointed back in the direction of acting as if ambiguity is the problem. On the one hand, this perspective only requires that one think experimentally in the evaluation of implemented policy. On the other hand, there will be few policies that warrant such thought.

The evaluation of policy-change-as-experiment requires discrete policy change and a discrete point in time to which the intervention of policy can be attributed. (Musheno, 1981) The

narrow range of policies that can be usefully evaluated by such interrupted time-series designs are obviously those which address social problems that are most easily defined and/or about which there already exists a high degree of consensus (and possibly knowledge) about the best means for goal attainment and problem resolution. Development policies quite obviously do not fall within this narrow range.

Instead, the subject matter of development policies must be included in anyone's list of ill-structured problems. This is true when attention is confined to the United States; when development policies are externally extended, all difficulties are compounded. For the moment, however, our focus is confined to the problem(s) of development, wherever and by whomever addressed. This focus is further restricted by our concern for those elements of development problems that make their understanding by experimentation and experimental evaluation a prohibitive endeavor.

The most fundamental problem with the extension of experimentation from physical science laboratories into the social arena is that individuals can choose their relationships with other individuals. These choices can change depending on the changes that individuals perceive in their environments, including the emergence of other individuals (the most dramatic of such changes being the birth of children). The feelings that are

associated with relationships among individuals are critical to their behavior, either individually or collectively.

Relationships among molecules simply do not possess such dimensions. Yes, those relationships can be more easily discerned by means of the laboratory's physical controls, but that is a necessary, not a sufficient, condition for their delineation. Furthermore, if molecules possessed the choice capacity of individual human beings, then physical controls of the laboratory would scarcely qualify as necessary, let alone sufficient, conditions for their understanding.

Yet, the controls of the laboratory have been glorified into sufficiency by social experimenters. Hence the devotion to randomization in the selection of participants in experiments and in their assignment to experimental and control groups. There is a point to the endeavor when it is applied to the simplest of life's problems. There is no point, and an exercise in rigorous futility, when applied to life's most difficult problems.

Policies that address the most trying of social problems specify goals that cannot be attained by any means other than by the interactions of numerous people. This means that existing relationships among individuals must somehow or other be altered in order for policy ends to be achieved. Reward structures must be changed, and, in order for this to occur, those that exist must be understood.

All of this suggests that experimental analyses, especially, and statistical analyses, generally, are doing a lot with the tools they have, but very little with respect to enhancing our understanding of the most serious social problems, and, therefore, with respect to contributing to improved public policy. Their central thrust is toward disaggregation of the observable world into as many discrete, and, hence, measurable parts as possible. Then, "controls" become all important, whether experimental or statistical. The disaggregation process loses what they subsequently attempt to find.

This is by no means intended to exclude quantitative analyses from attempts to better understand social problems, or from efforts to improve policy, its impact, and its evaluation. Rather, it is contended that ideas (theory) about social relationships must influence our representations and measurements of observables. (Chittick and Jenkins, 1976) Only by the imposition of such ideas can relationships among individuals be incorporated into quantitative analyses. They can never be discerned by disaggregation and control. Yet, they are what policy must impact in order for any of its goals to be realized. Relationships cannot be seen. They are not discrete. Thus, ideas about a continuous world must be imposed on its discrete extractions if quantitative analyses are to incorporate the

"relationship" variables inherent in development processes.
(Jenkins and Chittick, 1975)

One problem with the extension of experimentation into the social arena might be turned into a benefit by becoming a dual recommendation for development policy: expand individual experimentation while restricting social experimentation. If one adopts the view that "experimentation" is the "only route to cumulative progress", and couples this with a belief that the progress of intended beneficiaries is more desirable than that of social experimenters, then surely individual experimentation within projects must be promoted by policy at the same time that social experimentation is discouraged. Given the beliefs of the preceding sentence, maximization of the trial-and-error experiments (no matter how rudimentary their design) that are corollaries of individual decisions would be an imperative component of development policy and projects.

However, while development policy and its impact might be enhanced by this beneficiary-as-experimenter perspective, their evaluation by a policy-as-experiment design, would be more difficult. To engage intended beneficiaries in a process of ongoing decision-points is bound to involve more gradual implementation than are more government-determined (both host and donor) interventions.

A goal of beneficiary-as-experimenter policy would be that its implementation would never end!--that the learning process thus

encouraged would be sustained beyond a project's formal termination in the form of a positive exponential increase in the development of intended beneficiaries. Implementation of such policy, however, would obviate the policy-as-experiment perspective for evaluation, for a beneficiary-as-experimenter perspective in development policy, if it has impact, will have multiple intervention points (as many, theoretically, as there are intended beneficiaries).

7. Act as if the policy situation provides the answer. Many of those who view the experimental perspective as futile, seek their own policy impact answer in the very source to which they attribute that futility. That is, the ill-structured situations which thwart the useful imposition of experimental designs in either implementation or evaluation are expected to provide critical cues for the most preferred answers.

....Since a policy's outcome depends on the interaction between strategies and constraints, policy makers should choose implementation strategies according to the situation's constraints (i.e., contingent conditions).
(Berman, 1980, 207)

There is merit in this perspective to the extent that it possesses or seeks to develop coherent ideas about contingent conditions and how to identify, and perhaps measure, them in any situation. Such a concern is suggested by its differentiation between "programmed" and "adaptive" implementation, and by the

claim that programmed implementation can be justified only "if all the following conditions in the policy situation hold":

- (1) The scope of change, implied by the policy, in the behavior of members of the implementing system is incremental.
- (2) The validity of the policy's technology (or theory) is relatively certain.
- (3) Members of the implementing system generally agree on the policy's goals and means.
- (4) The coordination structure of the implementing system is tightly coupled.
- (5) The environment of the implementing system is relatively stable. (Berman, 1980, 241)

Fulfillment of all these conditions reminds of the requirements for social experimentation and policy-change-as-experiment evaluation. It is not surprising, then, to find programmatic implementation linked with the "classic model of evaluation ... based on the paradigm of experimental design" as "paradigmatic partners". (Sharp, 100-103). It is also unsurprising to find an apparent preference for adaptive implementation strategies in most situations. the ideal of such implementation is:

The establishment of a process that allows policy to be modified, specified, and revised--in a word, adapted--according to the unfolding interaction of the policy with its institutional setting. Its outcomes would be neither automatic nor assured, and it would look more like a disorderly learning process than a predictable procedure. (Berman, 1980, 211)

While this might be preferable to programmatic implementation of most policies in most situations, it abdicates any position that might have been adopted for enhancing the impact of pre-implementation policies. Though policy outcomes are never automatic or assured, it is not too much to aspire to increased learning with decreased disorder. Coherent ideas about institutional settings should be useful in multiple policy situations, and be sufficient for prescribing preferred procedures and predicted outcomes. Without expectations, organizations cannot be demonstrably wrong, but neither can they learn.

Evaluation of adaptive implementation efforts must themselves be more disorderly than those which address programmed implementation endeavors. Thus is adaptive implementation linked with its own paradigmatic partner, "qualitative/process evaluation approaches". (Sharp, 106). While these employ myriad techniques (see Patton, 1980; 1982), they allow for both street level policy and evaluation in perhaps their purest form, as represented by "eolithic programs":

Those programs whose participants are guided by the principle of eolithism; i.e., they look around them to see what's available and then do whatever they can with whatever they find. What they do moves them toward emerging goals that are discovered as they grow out of the environment in which they find themselves, or are inherent in the materials available to them. (Patton, 1982, 114).

The evaluation of such programs is obviously goal-free, and

requires an evaluator to adopt a position of "eolithic craftsman" ... "working from the principle that each evaluation opportunity is unique--its purpose, function, and possibilities to be discovered, rather than imposed or preordained". (Patton, 1982, 117)

Acting as if there is nothing of value between the "discovered" and the "imposed", these programs constitute "pure" street policy and evaluation. Their impact can only be from, and never on, implementation. Their purity is that of non-policies, and, thus, "implementation" becomes everyday "behavior", or what individuals can be expected to do in the absence of policies and programs.

Ultimately, then, adaptive implementation and qualitative/process evaluation approaches may provide less in the way of policy impact answers than they do justification for policy abandonment. While introducing many important considerations for implementers and evaluators, their answer to the policy impact problem points in the direction of policy = situation policy = policy situation, and provides policy producers with too little for making sense of mountains of situational evidence.

The inventiveness of the General Accounting Office's suggested model for providing congressional oversight of public policies can now be appreciated. While programmed implementation would conform to the control function of the model, the GAO clearly recognizes

that programmatic implementation requires both greater knowledge and consensus than the Congress possesses. Hence, to promote congressional control in spite of these deficiencies, the GAO has devised a hybrid model that combines adaptive implementation with programmed oversight.

In seeking to enable the Congress to veto negatives where it is unable to promote positives, the GAO reveals a basic appreciation of the "policy situation" perspective. Indeed, it suggests an adaptive mode for implementation, except that positional interests are reversed. In the GAO model, legislators become implementers, and street level bureaucracy is raised to new heights.

In effect, the seventh means by which the policy impact problem cannot be resolved brings us full circle to the first of those means. The first denies (by omission) the existence of a problem. The seventh denies (by commission, or by commitment) the problem's existence--can there really be a problem when answers are expected to emerge from situations?

C. THE REAL COSTS OF THE HEADSTAND

The failure to conceptualize "policy impact" within the public policy process necessarily corresponds with the divorce between implementation and policy. This divorce, in turn, is better understood as a "confused relationship" than as a "clear

separation"; as "messy" rather than "clear". Thus has much begging of questions, misplaced measurement, and displaced policy been witnessed of the seven "false messiahs" considered in this chapter.

"Displaced policy" results from those "messiahs" who would simply assume away any distinction between policy and its implementation. Others were seen to echo pleas for decreased ambiguity in the language of policy, and to encourage their own specificity (misplaced measurement) in its stead--resulting, again, in an encouragement of displaced policy.

Having now considered the most prominent candidates for resolving the policy impact problem, and voting for none, the following pages launch an effort to create a "winning" candidate. Restating the problem should assist in focusing the endeavor: The existence of street level policy making is not the problem; the problem is in not knowing when, where, and to what extent, it is occurring. Without that knowledge, no consequences--be they good, bad, or indifferent--can be reasonably attributed to any policy (whether it be pre-implementation, eolithic, or somewhere in between).

In order to evaluate pre-implementation policy, the extent to which it is impacting policy-as-implemented must be determined. In lieu of this policy impact knowledge, it is impossible to make responsible decisions regarding pre-implementation policy on the

basis of any quantity and quality of outcome evaluations.

Hence, the real costs of failing to stand policy implementation on its feet are those of choices denied. Such denial poses an implicit, but no less real, threat to public policy of an open, competitive society. A void in the literature, then, is but a faint reflection of the real costs ensuing from the failure to resolve the policy impact problem. Those costs are insidious because indeterminate, born of a far more serious void--in public policy, itself.

Thus are dire implications attributed here to the failure to conceptualize policy impact for standing implementation on its feet. Thus, too, are we subject to allegations of excessive pessimism. But consider this: If the position of policy implementation--stood on its head--is perpetuated, can policy itself avoid the same fate? And if it cannot, how stands the public?

More positive possibilities are tendered in the following chapter. Then, the fourth chapter directs itself to realizing some of those possibilities.

III. TOWARD STANDING POLICY IMPLEMENTATION ON ITS FEET

The answers needed for standing both the practice and study of policy implementation on its feet cannot ensue from questions begged. The preceding chapter abounds in illustrations of this axiom. Rather than providing means by which the problem of street level policy making might be resolved, its "false messiahs" simply detail the various ways in which the practice and study of implementation might secure a more comfortable headstand.

The dearth of discernible effort to stand domestic policy implementation on its feet simply reflects the consistency that now exists between study and practice. Furthermore, the scant differentiation between those who study and those who practice evaluation and implementation promises reification of the present status quo into the indefinite future.

The remainder of this study is dedicated to an alternative, and far more positive, future. It begins with an assumption: In order for the practice (study) of policy implementation to assume an upright position, its headstand must be understood.

The first of the present chapter's three primary tasks, then, is to explain the ascendance of street policy and the "ambiguity revolution". The second objective is to enhance understanding regarding the probable future of this new status

quo. More specifically, its perpetuation (and reification) is seen to be a "natural" consequence of large organizations.

Resolution of the street policy problem, therefore, is viewed as requiring unexpected effort on the part of an organization's membership. Thirdly, some seminal suggestions from the literature are introduced as prerequisites for expecting the unexpected. Recognizing that positive behavioral changes in the human experience are typically preceded by changes in conceptualization, these recommendations entail changes in our thinking about public policy and the organizations chartered with bringing them to fruition.

Thus, the present chapter is devoted to conceptualization. The changes in thinking that it suggests are then brought to bear on behavioral change by the methodology of the concluding chapter.

A. EXPLAINING THE ASCENDANCE OF STREET POLICY

The street's ascendance in public policy is best understood as a consequence of rampant ambiguity, and not vice-versa. Therefore, our explanation of policy implementation's headstand must account for the ambiguity revolution.

In the preceding chapter, both "politics" and "anti-politics" were deemed insufficient for understanding the escalation of ambiguity in public policy. This is not to say that beliefs about

government and/or interests in the activities of government have no bearing on the ambiguities of public policy. Quite the contrary, but those beliefs and interests are not a necessary result of either politics or anti-politics. If anything, its the other way around.

Indeed, the explanation for the ambiguity revolution that is tendered here is directly predicated on changes in beliefs about what government is capable of accomplishing and/or should attempt to accomplish. In a nutshell, policy ambiguity is a function of the relationship between two things: What we attempt to do, and what we know how to do.

If we attempt to do substantially more than we have in the past, but our relevant knowledge has not grown commensurately, then policy is bound to become more ambiguous (assuming an open political process, of course; a Hitler, Stalin, or Khomeini can always generate highly specific policy on the basis of exceptionally private knowledge).

Though not exhaustive, several of the more obvious characteristics of the ambiguity revolution are specified below. Most of these are viewed as being both cause and consequence of the growth of governmental activities.

First, the dramatic increase in what government attempts to do (witnessed in the United States in just the past twenty years) has resulted from, and in, the incorporation of society's most "messy,

ill structured, or squishy" problems among those for which government has assumed some degree of responsibility. (Dunn, 1981, 195)

Second, since thought about how best to alleviate such problems invariably shares their discordant characteristics, and these are compounded by the various "answers" of stakeholder interests, legislated policy regarding them is typically nebulous, ambiguous and vague.

Third, the rapid emergence of professions in social implementation and evaluation both within and outside government has been necessitated by the dramatic increase in jobs for accomplishing some positive things in the problem areas addressed by legislated policy.

Fourth, the challenges for these "new" professionals far exceed those which government has presented in the past.

Fifth, as action has gotten ahead of thought in the production of policy, so have the demands on implementation and evaluation exceeded requisite knowledge. Overall knowledge (and ignorance) is probably roughly equal among the three professions.

Sixth, in the context of the preceding, increased demands for doing something, and something with positive results, yet, decreases the likelihood that thought can ever catch up with action.

In the United States, at least, government could hardly have brought about this state of affairs in lieu of non-government and the security orientations of private individuals. This is reflected in the drive to restrict the uncertainties and risks of competition, as by redistributive coalitions. (Olson) The success of their efforts amount to securing demand for their own production (supply). This contribution to the growth of governmental activities is complemented (?) by others who seek to secure supply for their own consumption (demand), as represented by the myriad stakeholder interests in public policy and its growth. (Chelimsky)

Then, government's organizational response to these non-governmental stimuli is sufficient in itself for perpetuating the after-the-fact circumstances of street policy ascendance.

B. ORGANIZATIONAL REIFICATION OF STREET POLICY ASCENDANCE

The complex social problems addressed by expanded governmental activities have guaranteed that more ignorance will be brought to bear on issues requiring more knowledge than on those requiring less. The organizational response to this dreary Catch-22 has entailed the multiplication of bureaus, divisions, and offices that will participate in the attempted resolution of any given question. This response is not without logic, for whatever the

ratio of ignorance to knowledge might be, the more specialized is an individual's area of expertise, the more can the individual be expected to "know" about the narrow range of subject matter falling within the parameters of that specialization.

The principal problem with this organizational response, however, is that it tends to become as "messy, ill structured, and squishy" as the problems addressed. At some point, organizational differentiation can be expected to result in more people talking past one another than are talking with one another. After all, if specialized knowledge is easily communicated to non-specialists, it ipso facto is not specialized knowledge.

While such an organizational response can succeed in burying problems, resolution by those means is dubious. Indeed, the response seems more appropriate for granting policy production and implementation the final decree in their divorce proceedings than for accomplishing a reconciliation. And, as indicated in the preceding chapter, implementation's divorce from policy may be matched by its relationship with evaluation.

The organizational differentiation between implementation and evaluation is matched, and reinforced, by academic differentiation and specialization. Contributors to the burgeoning body of

implementation literature rarely contribute to that of evaluation, and vice-versa.\6

The core of the tension between implementation and evaluation may be simple--their practitioners have different interests--but this is too simple. a designation of "positional" interests in the process of public policy tells us substantially more (Rein, 134) What individuals will want to know, and therefore the questions they will ask, is very much a function of their respective positions, even when they are purportedly addressing the same subject.

Associated with the increasing differentiation of positions and specialization of their occupants, then, is a proliferation of "important" questions--for surely no one, occupying any position, can be asking unimportant questions. To ask how all of these questions (and some ostensible answers) can be important is to ask how all of these positions can be important, and it is simply not in the self-interest of occupants to ask.

Thus, both training and position of organizational membership buttress the ascendance of street policy, but not because anyone intends it. The sins associated with street policy and the denial of choice are those of omission, not commission. To the extent

\6This conclusion is confirmed by comparing the index of names in any book on evaluation with that of any book on implementation.

that an organization's personnel talk past one another, there can be no alternative to street policy--it is a "natural" outcome, and will occur even when a sizable majority of those personnel intend otherwise.

The preceding discussion does not warrant a conclusion that policy producers, implementers, and evaluators do not, or should not, have specialized skills. Quite the contrary. However, we do know that differentiation and specialization can be carried too far. And we know one other thing--no one has the skill, or the measuring instrument, for determining the point at which "too far" has been reached.

The methodology of this study's concluding chapter will be more fully appreciated with the above-stated conclusions in mind. Any organization that adopts this methodology will accept changes in its processes. Conversely, any methodology for policy impact evaluation that does not require changes in organizational processes is doomed to failure. At the same time, any such methodology will reify the "natural" outcome of street policy ascendant.

Thus, the methodology outlined in the next chapter will be seen to run counter to the exclusivity of knowledge and its use that is associated with increasing differentiation and specialization. Indeed, its adoption would result in the creation of new information and expectations for the benefit of policy

producers, implementers, and evaluators--individually and collectively. Implementation of the proposed methodology would expand an organization's body of information about itself and its efforts that would be interpreted and understood in the same way by the wide variety of its personnel.

Such "knowledge in common" (as distinguished from "common knowledge") would be generated by the methodology's implementation regardless of the adopting organization's location on a differentiation-specialization continuum. It can only be adopted, however, by organizations that have not already traveled so far along the continuum that they cannot "decide on what to decide." (Schattschneider, 1960)

C. REFLECTION, CHOICE, AND ORGANIZATIONAL CHANGE

Direct implications of the failure to usefully conceptualize "policy impact" have been asserted throughout this paper. More recently, in this chapter, the failure to conceptualize "policy impact" has been understood as a result to be expected of organizations--as a "natural" consequence. This conclusion echoes two seminal pleas for new conceptualization--one of these addresses public administration/organization (Ostrom, 1973), and the other regards "democracy". (Schattschneider, 1960)

Positive correspondence between those earlier appeals and the present study is hardly surprising. After all, both public

administration and democracy concern, among other things, how people do organize, and might organize, to get things done-- including the translation of beliefs and ideas into actions (i.e., policy implementation).

Ostrom demonstrates the virtual absence of good ideas (theory) for justifying adherence either to hierarchic models of conventional public administration or to expectations that those models might be useful for guiding the actions of public officials and accounting for what actually transpires. (Ostrom, 1973)

His appeal for new conceptualization in public administration echoes the "intellectual crisis" afflicting the study, understanding, and, hence, survivability of democratic politics:

The intellectuals have done very little to get us out of the theoretical trap created by the disparity between the demands made on the public by the common definition of democracy and the capacity of the public to meet these demands. (Schattschneider, 1960, 136)

The demands made on the public by inadequate conceptions of politics and democracy ("government by the people," for example) are as unrealistic as those made on public administrators by hierarchic conceptions of the operation of organizations. The former give us an upside down triangle with 200 million or so of the public at its base (top), whereas the latter depicts the reverse.

Simply putting these two triangles (conceptions) together establishes the impossibility of both being useful

representations. In fact, as Ostrom and Schattschneider forcefully demonstrate, the two conceptions are irreconcilable. As such, if they are intended to be simultaneously operative, then that intent is misconceived and therefore dangerous.

Neither of these scholars denies that there are some circumstances in which a centralized, hierachic government might be appropriate. Thus, Ostrom concludes that "...Hobbes was right in fashioning a constitution appropriate to a garrison state capable of functioning with reasonable effectiveness in a world plagued by recurrent warfare...." (p.130) But he also contends that,

...Hamilton and Madison were right in presuming that societies of men are capable of establishing good government by reflection and choice where a system of constitutional rule can be enforced in a political system characterized by substantial fragmentation of authority and overlapping jurisdiction. Such a constitutional system is capable of maintaining democratic administration as a general form of public administration which stand in contradiction to bureaucratic administration. (Ostrom, 1973, 130-1)

Furthermore, in a world that does not require a garrison state, and in which we aspire to a democratic polity, both Ostrom and Schattschneider would perceive hierarchic, centralized government as thwarting those very aspirations--whether all of "the people" or but one of the people are presumed to be at the "top" of the hierarchy. In either case, too much is expected of

the top. The following admonition applies no less to public officials than to the electorate:

...The power of the people in a democracy depends on the importance of the decisions made by the electorate, not on the number of decisions they make...The unforgivable sin of democratic politics is to dissipate the power of the public by putting it to trivial uses. (Schattschneider, 1960, 140)

To the extent that the rate of growth in what government attempts to do exceeds the rate of growth in knowledge of how to do it, then the likelihood that the public and its representatives are put to trivial uses is compounded. And, as the sheer volume of decisions grow, the truly important decisions are increasingly jeopardized.

In such a context, there are increased incentives for evaluating only those things that are most readily evaluated by the tools at our disposal, or to create greater specificity in "policy for evaluation" than ever existed at any time during the conflictful debate, logrolling and compromise associated with its formulation. By either of these responses, what is ultimately reported as an evaluation of policy may bear only the dimmest reflection of what producers, or anyone, had in mind. To the extent that such evaluations are actually used by producers--and this is as unclear as is the impact of policy (Weiss)--a new category might be created for the accommodation of "evaluation policy".

The danger inherent in all of this is not that policy and/or outcomes might be made worse (indeed, at least in the short run, they might be made better). The danger is that no one will know what is being decided. This is worse even than being denied important choices by a multiplication of minor ones. Thousands of labor hours and untold reams of paper can be exhausted in the course of choosing among unknowns. This is not to say that there will not be veritable mountains of data regarding each of the unknowns (now called "options", or "alternatives", perhaps), and, indeed, like shadows on cave walls, "something" is there. That "something" recalls another definition: "Political science: a mountain of data surrounding a vacuum." (Schattschneider, 1969, 8)

At risk is the "reflection and choice" which Hamilton deemed essential to political life. With so much to reflect upon, reflection about any thing must decline, and, with that, quality of choice is endangered by quantity. In this context, it is remarkable that no matter how conflictful and vituperative debates over the size of government become, they only rarely reflect a concern for the quality of choice. While quantity varies, it is as if quality is assumed to be a constant. Upon reflection, this state of affairs makes little sense, but, then, who has time to reflect?

Though Ostrom does not wish to characterize these issues in terms of "centralization versus decentralization", the central

thrust of his answer to the preceding question would be: only a multiplicity of choosers can be expected to take time to reflect. Thus,

The industry characteristics of multi-organizational arrangements functioning in a public service economy can only be realized where diverse public agencies are able to develop different economies of scale in response to varying communities of interest. Overlapping jurisdictions and fragmentation of authority thus are necessary conditions for public services industries, other than fully integrated monopolies, to exist. Centralization cannot be conceived as the converse of decentralization in the sense that we speak of centralization versus decentralization. In responding to problems of diverse economies of scale, elements of centralization and decentralization must exist simultaneously among several jurisdictions with concurrent authority. (Ostrom, 1973, 73)

This assessment identifies the central genius of this country's political mathematics--dividends produced by multipliers. The latter are too frequently expounded upon while ignoring the former. Perhaps this is to be expected from those whose platforms are defined by positional interests, as in the following:

...The 'separation of powers' means that evaluations and other analyses will almost always be of an adversary nature, and frequently duplicated, for the executive branch and the Congress. It also throws particular doubt for the United States on the viability of the 'helping' evaluation, since evaluatees will always feel that they are being judged by someone. (Levine, 34)

Thus is the multiplier of the separation of powers acknowledged with no apparent appreciation of its dividends. Ironically, when evaluations of virtually everything and anything

are among this country's most abundant products, but they neglect policy impact, it is difficult to not contemplate the possibility that powers are not sufficiently separate. We do not exclude the potential for being helped by adversaries; we cannot know without listening, and we cannot listen if they do not exist.

Further delay in the conceptualization of policy impact assures the expansion of vacuums, no matter how high the mountains of data surrounding them. It is impossible to determine the costs of this neglect because we cannot know where we would be without it. In addition, where costs are manifest and should be attributed to this neglect, odds are they won't. After all, how can costs be attributed to the indeterminate...to a vacuum? This means that when these costs are attributed they will be to the wrong sources, and thus compound themselves.

Today's latent costs are invariably tomorrow's manifest costs. In this instance, waiting for the manifest entails costs of its own. The next chapter recommends both not waiting and means by which the time "not waiting" might be best spent.

IV. POLICY IMPACT: CONCEPTUALIZATION FOR EVALUATION

A conclusion to be drawn from the preceding chapters is that what we look at typically depends on how we can look at it. How we can look at things, in turn, depends on what we have to look with. Many of the things people are looking with (the techniques they employ) might be useful for evaluating policy impact, but not in lieu of its conceptualization within the policy process. The most critical component, is also the most neglected. The absence of conceptualization renders otherwise useful techniques into detriments to understanding of those things to which they are applied.

"Absence" of conceptualization is used advisedly. The major problems to be found in that small part of the literature that might be viewed as looking at policy impact are more easily attributed to no conceptualization than to bad conceptualization (unless one employs a rule that the former always qualifies for the latter).

In this chapter, "policy impact" is conceptualized in such a way that it can be distinguished from other elements of the policy process. Then, the utility of that conceptualization is illustrated by operational demonstration. First, however, semantic confusion born of the divorce between implementation and policy must be recognized, dealt with, and resolved.

A. IMPLEMENTATION OF WHAT?

We are told that, "Evaluation of implementation is distinguishable from implementation feasibility analysis." That is, "Evaluation of implementation occurs during and/or after the implementation of policy." In contrast, "implementation feasibility analysis" occurs "prior to policy adoption and focuses on predicting obstacles to implementation and recommending changes in policy design to avoid or overcome them." (Sorg, 153)

The complete text of the article from which these quotations are taken reveals that the author employs the term "policy" to refer to projects. Such interchangeability of terms is a disconcerting characteristic common to the literature of both evaluation and implementation. In fact, as the preceding chapters have established, policy implementation is demonstrated neither in practice nor in the literature.

Perhaps the frequency with which projects and/or programs are graced with the term "policy" reflects wishful thinking. Unfortunately, such casual use of language only camouflages the vacuum that is policy implementation, policy impact, and their evaluation. The use and acceptance of that language thereby thwarts the development of alternatives to the vacuum.

It might be recalled that additional semantic distinctions were necessary in order for this study to progress beyond its

first page. The differentiation there, between "impact evaluation" and "policy impact evaluation" is reiterated here and more fully elaborated.

"Impact evaluation" may, or may not, involve evaluation of policy impact. To the extent that policy indeed is influencing projects and their outcomes, then an impact evaluation is a policy impact evaluation. To the extent there is no such influence, then results of the two evaluations are independent of one another.

An "impact evaluation" necessarily involves an "outcome evaluation", where the latter estimates the consequences of whatever is done in the field setting. These outcome estimates are compared with those that were expected upon initiation of project implementation. The difference between these expectations and results, in the field, is both necessary and sufficient for an impact evaluation, even when completely independent of policy.

By the present usage, "policy impact evaluation" includes "impact evaluation" if, and only if, the essential components of "policy impact evaluation" have been fulfilled, and the results of those evaluations demonstrate that a project (whose consequences are the subject of an impact evaluation) has been discernibly influenced by policy.

"Policy impact evaluations" must respond first to origins-- the source, and at least partial cause, of what follows. "Impact

evaluations" must first respond to endings--the results, and at least partial effects, of that which preceded.

B. IMPLEMENTATION BY WHOM? DO THEY AGREE ON WHAT?

Having provided these distinctions, the most essential component of policy impact evaluation may now be specified. It will be seen to be first in three respects. First, because it is most narrow in focus. First, because it is closest to the origin of policy content. And first, because it occupies a position of primacy, or substantive importance, relative to other components of policy impact evaluation. Indeed, a case is made that evaluations of the other components are valueless (for policy impact evaluation) when the results of the first component's evaluation are negative.

What can be so important? It is so basic that the literature of both evaluation and implementation is imbued with acknowledgement of its significance. The degree of concern for "compliance" is likely to be inversely related to it. It is ubiquitous where things are good, and scarce where things are bad. The Kremlin and Washington definitely do not have it, though at times they have appeared to achieve its opposite.

Lest we lapse into a "who was that masked man?" exercise, let "positive interdependence" be recognized without further ado. To the extent that an organization (or any other collectivity) is

characterized by positive interdependence among its members, the benefits (costs) that are directly experienced by some of its members will be at least perceived by other members as benefiting (costing) themselves. In contrast, negative interdependence of a collectivity is marked by the perception of benefit by some of its members when others suffer loss (or by a perception of cost when others experience gains). Thus, the more negatively interdependent a collectivity, the more prevalent will be zero-sum orientations (i.e., whatever I gain, you lose) among its members.

Obviously, there will be some "zero-sum individuals" within any organization. However, they are less likely to experience success, and, therefore, will be fewer in number, within an organization characterized by high positive interdependence. Such an organization is likely to have a left hand that not only "knows" what the right is doing, but knows that it would be doing the same thing were it in the other's place.

This expression of an organization with high positive interdependence is obviously that of an ideal model, never to be realized by human beings. At the same time, it denotes the critical importance of an organization's positive interdependence for translating ideas into action. Evaluating the extent to which it exists, then, can tell us a great deal about both the likelihood of such translation occurring and the utility of further extending the evaluation of policy impact.

Indeed, answering the preceding sections question-- implementation of what?--can only now be answered as we prefer. "Policy," not projects or programs, was our answer there. That answer, however, is incomplete, for what is the content of "policy"? what does it mean? We mean no more, nor less, by "policy," or by any given policy, than do the members of whatever organization(s) is responsible for its formulation and for bringing it to fruition.

Therefore, to the extent there is shared understanding among an organization's membership regarding the meaning of a given policy, it is likely to be implemented and impact outcomes in the field. In these circumstances, the utility of extending the evaluation of policy impact to the field, both pre- and post-implementation, is clearly positive.

Conversely, to the extent there is disagreement among an organization's personnel regarding what any given policy is, there should be little concern for evaluating its impact (whatever impact of the amorphous might mean).

Agreement about the meanings of an organization's policies, then, is assumed to be positively associated with the positive interdependence extant among its membership. The greater (lesser) is one, the greater (lesser) is the other. At the same time, of course, variance is expected among any given organization's policies as to the shared understanding of their respective

meanings by its membership. This can be accounted for by differences in quality of writing alone. More typically, however, such variance results from the differences in knowledge (or ignorance) of subject matter addressed by the multiplicity of policies. For any given policy, though, impact should be expected only if there is considerable agreement regarding its meaning(s) among organization members.

To this point, considerable conceptualization has occurred in order to identify what to measure and why. Though scarcely exhausting all relevant considerations, it should be sufficient for taking the next step.

C. MEASUREMENT FOR EVALUATION

The methodology outlined here employs questionnaires for completion by Agency personnel engaged in policy production, project identification and/or implementation, and in the evaluation of Project Identification Documents (PIDs).

Completed questionnaires provide the basic input data, by individual and by project, regarding what we shall call "policy implementation probability". The questionnaires entail policy evaluation with respect to each identified project. Existing evaluations of PIDs regarding their implementation feasibility would continue, and would be unaffected in any direct way by adoption of the new evaluation instruments. Their emphases are

quite different, and their consequences might be even more disparate (as contemplated in the next section of this chapter).

A policy evaluation questionnaire associated with each PID feasibility evaluation would enable both more specific and complete meaning to be given to the respective policies of the Agency. In brief, respondents to each questionnaire are asked to identify to which policies a PID is responsive, and to which policies it should be responsive.

Various possible designs of the policy evaluation questionnaire are not discussed in this study. Should this paper's basic conceptualization and operationalization of policy impact and its evaluation be positively received, the development of questionnaire designs would be the most appropriate next step. Nonetheless, several things should be said about their intent and potential use. \7

Whatever policy questionnaire might be employed, it should be jointly approved by the Office of Evaluation and the Office of Policy Development and Program Review, at minimum.

\7Should a policy evaluation process like this be adopted, "modus operandi" techniques (Smith) should be seriously considered for incorporation in the questionnaire design. (Scriven) Among other things, they offer the possibility of identifying non-private evaluations. The results of these evaluations and of their subsequent analyses should be shared, but the inputs should be those of individuals.

Furthermore, the policy evaluation instrument should be incorporated into the normal PID writing and review process. Any movement of a Project Identification Document from one individual to another would be complemented by a completed policy evaluation questionnaire sent to the Office of Evaluation.

In order to provide a more complete picture of intentions, both as to elements for inclusion in a policy evaluation questionnaire and regarding analyses of resulting responses, Tables 4-1, 4-2, and 4-3, are offered. These tables are presented in the order by which one of the questionnaire's two principal dimensions--the relative importance that the respective policies of the Agency should have on an identified project--is translated into disparity and interdependence values.

The input data for these three tables, then, do not reflect individual judgments with respect to which policies a PID is responsive. This simplification is joined by others in order for the logic of the policy evaluation methodology to be more easily discerned.

Even quite arbitrary simplifications are introduced into the analysis of these pages. For example, a maximum of twenty-four policy documents is employed, thus enabling the number 25 to be assigned to those policy documents not referred to by a given

evaluator. In addition, the number of evaluators is limited to three.

A risk run by such simplification of the policy evaluation process is that the richness of information which can be derived by its adoption will not be envisioned. In hopes of counteracting

Table 4-1: Relative Importance Which Policy Documents should have on Identified Project, by Individual Evaluator (where 1 = most important, ..., 25 = no reference to a policy referred to by another evaluator)

POLICY DOCUMENT	EVALUATOR		
	1	2	3
Institutional Development	1	4	3
Recurrent Cost	25	2	2
Water & Sanitation	2	3	1
Private Sector Development	25	1	25

that possibility, the reader is encouraged to recognize that the following tables would be coupled with a parallel set of information provided by responses regarding the other principal dimension of the questionnaire, namely, the policies to which a given PID is responsive.

These pages, then, constitute but a preliminary introduction to an enterprise for which one's abacus must soon make way for an

IBM. Just to illustrate the possibilities that exist in Table 4-1 that are not readily evident in those rankings of policy importance, consider the possible divergence that can underlie even identical rankings.

Suppose, for example, that the project in question (presented by a PID) offers a combination water-sewage project for a small community. Table 4-1 reveals that both Evaluator 2 and 3 rank recurrent cost policy second in importance as to the influence it should have on the project. This identical ranking accounts for the zero value in the 2/w3 column of Table 4-2.

Now, what does this tell us? Perhaps that each thinks recurrent cost policy should have considerable influence on the water-sewage project? Not necessarily, for there may be a considerable distance between Evaluator's 3's first- and second-ranked policies, whereas Evaluator 2 might have been torn as to the preferred order of her top-ranked policies.

Well, if the "identical" ranking given to recurrent cost policy by Evaluators 2 and 3 does not necessarily indicate that they attribute the same magnitude of importance to it, then might it not at least indicate a shared belief regarding the kind, or nature, of influence the policy should have on the water-sewage project? The answer, of course, must be an emphatic "no".

Each policy of virtually any large organization is comprised

of many elements, and a policy document is typically negotiated as well as written. The component elements of such policy reflect disparate interests, of varying degrees of intensity, of an organization's membership. "Success" is achieved in the production of a policy document when none of its component elements are incompatible with one another. When those elements complement one another, that success is strengthened. And, to the extent that a policy's component elements complement, without contradicting, other policy documents, that success may be the subject of considerable acclaim.

With regard to our two evaluators and their ranking of recurrent cost policy vis-a-vis the water-sewage project, they could have in mind quite different (even inconsistent) elements of the policy. If that were the case, then the zero-difference in their rankings of recurrent cost policy denoted by Table 4-2 hides more than it reveals. Indeed, that difference would be deceptive rather than informative. A first impulse in responding to such an eventuality might be to engage in in-depth interviews with each of the evaluators in order to discern those differences not reflected in the tabulated questionnaire results. That impulse is not only deemed wrong, it highlights an assumption of the proposed methodology.

It is assumed that any difference between evaluators regarding the desirability of any given policy's influence on a

project affects their differences with respect to the desirability of any other policy's influence on that project. Hence, where

Table 4-2: Differences Between Evaluators in Rankings of Importance which Policy Documents should have on Identified Project, by Policy Document, by Dyad--derived from rankings in Table 4-1.

POLICY DOCUMENT	PAIRS OF EVALUATORS (Dyads)			Differences
	1/w2	1/w3	2/w3	
Institutional Development	3	2	1	6
Recurrent Cost	23	23	0	46
Water & Sanitation	1	1	2	4
Private Sector Development	24	0	24	48
Sum of Differences	51	26	27	104
*Possible Sum of Differences	90	69	90	
Dyadic Disparity Value	.5567	.3768	.3000	

*[(25)(R)] - $\sum r$, where R=number of policies ranked by either member of dyad, and r=lowest possible sum of ranks, given R.

there are differences regarding one policy, a zero-difference in ranking of any other policy is assumed to be a consequence of

different reasons. A corollary assumption is this: a zero-difference in policy ranking by any pair of evaluators with regard to the desired influence of that policy on any given project indicates shared beliefs about the policy and its principal objectives only if all other policies ranked by the respective evaluators with regard to that project are identical in rank.

With the preceding assumptions in mind, a review of Table 4-2 has Evaluators 2 and 3 giving recurrent cost policy the same rank, but for quite different reasons. A summary measure of this conclusion is provided by dividing the sum of their differences in policy ranking by the maximum sum of their differences that was possible given the number of policies that either ranked. Dividing a pair of evaluators' actual differences by their possible differences (maximum) yields the dyadic disparity values that occupy the bottom row of Table 4-2.

Dyadic disparity values, ranging between 0 and 1, advise as to the likelihood that any given pair of evaluators will share the same understanding of any given policy. The greater is any dyadic disparity value, the less the likelihood that the associated pair of evaluators mean the same thing about one or more policies. Of course, were the methodology suggested here to be implemented, any susceptibility to error of these judgments would decline with the introduction of additional policy evaluations, centered on other projects.

With no more than the single-case evidence of this chapter's illustration, it can be concluded that with respect to at least four agency policies, an understanding of their respective intentions and relative importance is less likely to be shared by Evaluators 1 and 2 than by any other pair of evaluators. In addition, since their dyadic disparity value is in excess of .5000, the understanding of either will be more conflictful than coincident with the understanding of the other.

It should be clear that whether a given understanding of policy is "right" or "wrong" is irrelevant to the proposed methodology for evaluating policy evaluations. It is recognized, for example, that an omniscient observer might deem each of any pair of evaluations to be 100% wrong, even where there is strong disagreement between them (as with Evaluators 1 and 2).

Similarly, there is no necessary relationship between the disparity values for pairs of evaluators regarding which policies should influence a given project and the likelihood that they will be in agreement with respect to supporting project approval. The same is true of the interdependence of an individual's policy evaluations with those of others (provided in the bottom row of Table 4-3). Thus, in the example of this chapter, Evaluator 2 might be the only dissenting voice regarding project approval even while having the lowest disparity with another evaluator (.3000

with Evaluator 3), and even though Evaluator 1 evinces the lowest interdependence with the group.

Table 4-3: Dyadic Disparity Values (in cells); Interdependence Values (at margin)--derived from pairwise difference in ranking of policy document importance for subject PID, as displayed in cells of Table 4-2.

EVALUATOR	EVALUATOR		
	1	2	3
1	---	.5667	.3786
2	.5667	---	.3000
3	.3768	.3000	---
Interdependence	.4718	.4334	.3384

*Equals average (mean) dyadic disparity; .4145 for 3-member group.

Getting from evaluations of policies to those of projects in a way that makes sense, i.e., with some degree of logic, would entail introducing information from the questionnaire that taps its other principal dimension. Unless it is known, for any given evaluator, (1) the policies to which a project is deemed responsive, as well as, (2) the policies which should influence that project, then any expectation regarding that evaluator's project support will be deficient. Rather than introducing another set of tables for the "is" that parallels the "should", the latter can be drawn upon in conjunction with possible "is"

evaluations for illustrating the bridge from policy evaluation to project evaluation.

It might be found, for example, that Evaluator 2's ranking of recurrent cost policy--regarding the influence it should have on the water-sewage project--was predicated on her being quite taken with the point of fn. 10, p. 10, of that policy document. As a result, she felt that since the project entailed piped water to all houses in the community, a minimum consideration should have been to provide occupants with the option of piped water with a user's fee or subsidized community stand pipes. Instead, the identified project was providing subsidization for all piped water. The negative disposition of Evaluator 2 toward the project might thus be understood as a clash between that individual's belief about what "ought to be", and what "is", allowed for by the identified project.

In addition, it might be found that Evaluator 3 (whose ranking of recurrent cost policy on the "should" dimension is identical to that of Evaluator 2) simply felt the overall problem addressed by recurrent cost policy to be extremely significant. Since the PID took into consideration the host government's positive response regarding the maintenance of previous capital investments, for Evaluator 3 there was complementarity of the PID's "is" and the evaluator's "ought to be".

Thus can the two evaluators (2 and 3) be on opposite sides

regarding project approval because of the differing emphases they accord a policy which they rank identically as to the impact it should have on the project. the priority given to private sector development by Evaluator 2, in conjunction with its being ignored by Evaluator 3, would probably make additional sense of both their differing emphases regarding recurrent costs and their offsetting votes as to project approval. The importance of the suggested methodology, however, is not predicated on its capacity for making these determinations. Instead, its importance is that of an instrument whereby an organization's members are more likely to make those determinations for themselves while coming to "share the same meanings and intend the same things". (p. 7, above)

The language just expressed, of understanding in common, was introduced in this paper in response to the divorce of implementation from policy. A principal point of the methodology suggested here is that attempts to reconcile these parties are apt to be futile without first addressing their respective understandings of themselves. Since policy is logically prior to implementation in the evaluation of policy impact, this chapter's illustration of methodology has focused on policy producers.

A corollary, but no less important, point of this chapter's methodology is that impact of any given policy should only be expected if there is considerable agreement as to its meaning among its producers. Where such agreement is lacking, then

outward concern for their divorce from implementation is likely to perpetuate the divorce by depriving themselves of the inward attention necessary for resolving their own confusion. There is divorce among both individuals and organizations that is accounted for by simple confusion.

Divorce can and does result from no more, nor less, than confusion about wants. It is not, then, simply the flip-side of marriage. The latter requires both clarity regarding wants, and their ordering by priorities. It is akin to political decisions, as Schattschneider recognized some time ago. Such decisions require that what people want more "becomes the enemy of what they want less. Politics is therefore something like choosing a wife, rather than shopping in a five-and-ten-cent store."

(Schattschneider, 1960, 68)

It should be clear, then, that the inward attention demanded by this chapter's methodology is born of a high regard for marriage. And the latter is seen to be more than reconciliation. Reconciliation of confused parties has more to do with deciding not to decide than with anything else. Deciding not to decide, in turn, has more to do with five-and-ten-cent stores and potpourri than with policy, its impact, and their evaluation.

When the results of the proposed methodology's inward attention incorporate the policy evaluations of both field and producer personnel, the organization provides itself with

expectations that it could not otherwise have. These expectations reflect policy implementation probabilities, or the likelihood of impact in the field of any given policy. How these expectations can be derived from the positive interdependence values of Table 4-3 is explicated by Table 4-4.

Table 4-4: Interdependencies among Policy Impact Evaluations of PDPR, by Interdependencies among Policy Impact Evaluations of all Personnel (field and PDPR)--with expectations in cells

		Interdependence Among PDPR Evaluations	
		High	Low
Interdependence Among PDPR and Field Evaluations	<u>High</u>	High Policy Impact Anticipated	Low Field Interdependence (idiosyncratic policy)
	<u>Low</u>	High Field Interdependence & Low Policy Impact (Street Policy)	High Field Interdependence Possible & Low Policy Impact (Street Policy)

Given that two of Table 4-4's cells are apt to result in street policy, the preceding chapter's conceptualization for understanding the phenomenon is buttressed by methodology and measurement. Furthermore, since a third cell's expectation of idiosyncratic policy implementation may well result in considerable street policy, the odds against policy impact are

here starkly portrayed. But no one ever said that marriage is easy, while a multitude have appreciated the work it requires. Without work--in lieu of a commitment to something like the methodology proposed in this study--policy impact should not, because it cannot, be imputed to the results of an organization's myriad evaluations of its own activities.

Implicit throughout this paper has been a judgment that the evaluation of policy impact can be realized only by introducing new processes whereby the likelihood of that impact can be both estimated and substantially enhanced. Thus, our methodology has had simultaneous goals: Maximizing the evaluation of policy impact by enabling recognition of where, when, and to what extent, that impact is occurring.

These goals are explicit in Table 4-4. A primary objective, of course, is to see an increasing proportion of policy impact evaluations falling in its upper left-hand cell. This can occur over time either by improvements in policy documents, intra-Agency education efforts (as would be strongly emphasized with respect to cases, and associated personnel, in the lower left-hand cell), or by changes in project designs that result in a stronger reflection of policy documents. Implementation of the proposed methodology enables knowledge of where organizational resources can best be invested.

The upper right-hand cell (low producer interdependence/high total interdependence) must be read as total interdependence (producer and field) being greater than that among producers separately. It is dubious that the "high" positive interdependence of the combined policy evaluations could in fact be very high when that of either of its component groups is low. For cases falling in this cell, work on the policy documents referred to by their associated evaluations would be of the highest priority.

Development of the methodology suggested by this chapter enables an adopting organization to acquire a new information base (reflected in Table 4-4) that provides it with a different perspective in subsequent evaluations of both field implementation and outcomes. Thus, these two remaining essential steps for completing the evaluation of policy impact are not only contingent upon the fulfillment of the initial necessary component to which these pages have been devoted, they will themselves be assessed through a new lens provided by the first step in policy impact evaluation.

For example, evaluations of field implementation of cases falling in the upper left-hand cell should be the most positive. If they are, but subsequent outcome evaluations yield mediocre results, this would be both more significant and troublesome than

would similar outcome evaluations for cases of any of the other three cells of Table 4-4.

Thus, the proposed methodology not only provides an organization with distinct pre-implementation expectations of when, where, and how much street policy it will incur, it allows it to differentially evaluate subsequent implementation and outcome evaluations. The latter, in turn, might then have meaning that is otherwise denied by omission; meaning that is clearly relevant to the words of pre-implementation policy. The methodology of this chapter can be thereby understood as a means by which an organization can increase its capacity for making important choices. Furthermore, the conceptualization that underlies our methodology is a positive influence for the creation of important choices. This potential is briefly explored in concluding the present chapter.

D. POLICY IMPLEMENTATION EVALUATION: BEYOND FEASIBILITY

The policy evaluation methodology resulting in policy implementation probabilities, or estimates of the likelihood of policy impact, has an analogue in the "implementation feasibility analysis" that is normally associated with the evaluation of projects prior to their adoption. However, the analogy obscures more than it reveals.

Not only does project "implementation feasibility analysis"

fail to produce any intersubjective expectations regarding the likelihood of a project design being implemented in the field, it embodies a bias toward doing again what has been done before.

"Implementation feasibility analysis" elevates the importance of getting a project to "work" (of getting from its beginning to its end). As a consequence, it puts a premium on field experience--what has worked, and what has not--with stress on worked, or not worked, rather than on what did or did not work.

At the same time, the task of "implementation feasibility analysis" is hardly unimportant. After all, even with maximum policy implementation, it can have no impact if the design for implementation cannot work. The point here, then, is to recognize it not as a rival of evaluations for determining policy implementation probabilities, but as a complementary endeavor for translating high probabilities of policy implementation into cases of high policy impact in the field.

In short, the two analyses simply address different things and ask different questions; both are essential to policy impact and its evaluation.

However, while implementation feasibility analysis is probably, in and of itself, no more biased in its perspective than is the analysis pertinent to policy implementation probabilities, there is a critical difference in their effects.

The neglect of policy impact by large organizations transforms

the bias of implementation feasibility analysis into one that is insidious for organizations and their realization of goals. A legitimate interest in making things work becomes over-emphasized by the omission of other serious considerations.

Thus do organizations tend toward the path of least resistance, with an "art of the probable" rather than an "art of the possible" orientation. Rather than encouraging new ideas and efforts for their realization, this orientation emphasizes tactical decision-making, to the detriment of strategic considerations. In contrast, policy evaluation places thought where it belongs--before action. The failure to succeed in this tends to constrict thought to the very narrow confines of assessing whether previous actions can be replicated in different circumstances. Again, such considerations are hardly unimportant, but to so confine our attention is to deny ourselves important choices.

Perhaps the most stark contrast between the two analyses can be summarized in a single question: "Did they believe in it?" This question is at the core of this paper's methodology for policy evaluation. To the extent that "they" do not share the same beliefs about the meanings of policies, then the implementation and impact of those policies should interest no one. "Their" beliefs are an essential ingredient in the determination of policy implementation probabilities. Why would

not the same be true of project implementation, and hence be incorporated in implementation feasibility analysis? After all, it is virtually axiomatic that people who believe in something are far more likely to bring it to fruition than are those who do not. Yet in all the discussions of what has and has not worked, try to recall when last was heard, "Did they believe in it?" Ever?

Furthermore, the methodology for assessing the positive interdependence of beliefs--of policy evaluations--need not be confined to Agency personnel. Though only those personnel (producers and implementers) are essential for the generation of policy impact expectations, the Agency might learn more about its activities within the realm of the possible by broadening its base of policy evaluators.

Thus, contractors engaged in field implementation would be an obvious addition for evaluating the policy dimensions of those projects for which their services are contracted. In some instances, it might even be desirable to gain the participation of host country officials who are most directly involved with a given project.

Finally, the Agency might engage outside policy "experts" whose evaluations of one or more of the Agency's policy documents are known to be widely divergent. Among other things, the Agency would learn of possible evaluations that its personnel are not engendering. When the policy evaluations of those outside the

Agency are entered into the equation, an appearance of wide disparities, and low interdependence, among the evaluations of Agency personnel might be discovered. At the very least, then, such possible outliers would benefit statistical analyses of interdependence scores by their provision of different anchoring data points. (Beardsley, 166-170) In addition, the directionality of changes in interdependence over time would be thereby facilitated.

The art of the possible need not be difficult, but when requiring change in large organizations, it can be excruciating. However, the methodology recommended by this study, and the changes in organizational processes that it would entail, escapes most of the costs associated with organizational change. It does this by recommending not changes in what is being done, but by occupying the void born of current neglect.

This study sees us denying ourselves important choices because of what we fail to do. Thus have voids in the literature been found consistent with vacuums in practice. Thus, too, do we deny the best that we, by commission, can offer due to critical omissions (including the neglect of omissions). This paper has been dedicated to changing that circumstance.

In its attempt to enhance the quality of choice, this study is intended to present, with regard to its own recommendations, precisely that--not a non-decision for transfer to file cabinets

or worse, but a quality choice for coming to know policy impact and thereby improving the content and consequences of public policy.

REFERENCES

- Bardach, Eugene. The Implementation Game: What Happens After a Bill Becomes Law. Cambridge: MIT Press, 1977.
- Beardsley, Philip L. Redefining Rigor: Ideology and Statistics in Political Inquiry. Beverly Hills: Sage Publications, 1980.
- Berman, Paul. "Thinking About Programmed and Adaptive Implementation: Matching Strategies to Situations." In Why Policies Succeed or Fail, edited by Helen M. Ingram and Dean E. Mann. Beverly Hills: Sage Publications, 1980. 205-227
- Campbell, Donald T. "Experiments as Arguments." Knowledge: Creation, Diffusion, Utilization. 3(3), March 1982. 327-337
- Campbell, Donald T., and Stanley, Julian C. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally & Company, 1963.
- Campbell, Donald T. "Reforms as Experiments." In Handbook of Evaluation Research. Volume 1. Edited by Elmer L. Struening and Marcia Buttentag. Beverly Hills: Sage Publications 1975. 71-100
- Chelimsky, Eleanor. "Making Block Grants Accountable." In Evaluation in Change: Meeting New Government Needs, edited by Lois-ellen Data. Beverly Hills: Sage Publications, 1981.
- Chittick, William O., and Jenkins, Jerry. "Reconceptualizing the Sources of Foreign Policy Behavior." In In Search of Global Patterns, edited by James N. Rosenau. New York: The Free Press, 1976. 281-291.
- Critchfield, Richard. "Science and the Villager: The Last Sleeper Wakes." Foreign Affairs. 61(1), Fall 1982. 14-41.
- Dallmayr, Fred R. "Critical Theory and Public Policy." Policy Studies Journal. 9(4), Special Issue Number 2, 1980-81. 522-534.
- David, Jane L. "Local Uses of Title I Evaluations." Educational Evaluation and Policy. January-February, 1981. 27-39.
- Dunn, William N. "Reforms as Arguments." Knowledge: Creation, Diffusion, Utilization. 3(3), March 1982. 293-326.
- Dunn, William N. Public Policy Analysis. Englewood Cliffs: Prentice Hall, Inc., 1981.

- Goodell, Grace. "Conservatism and Foreign Aid." Policy Review. Volume 19, Winter 1982. 111-131.
- Goodrich, James A. "Optimizing under CETA: Program Design, Implementation Problems, and Local Agencies." In Implementing Public Policy, edited by Dennis J. Palumbo and Marvin A Harder. Lexington: D.C. Heath and Company, 1981. 45-54.
- Jenkins, Jerry, and Chittick, William O. "Reconceptualizing Foreign Policy Behavior: The Problem of Discrete Events in a Continuous World." In Foreign Policy Analysis, edited by Richard L. Merritt. Lexington: D.C. Heath and Company, 1975. 79-92.
- Just, Marion R., and Murray, Charles A. "Community Participation and Development Projects in Thai Villages." In Theory Construction and Data Analysis in The Behavioral Sciences, edited by Samuel Shye. San Francisco: Josey-Bass Publishers, 1978. 112-133.
- Kiser, Larry L., and Elinor Ostrom. "The Three Worlds of Action: A Metatheoretical Synthesis of Institutional Approaches." In Strategies of Political Inquiry, edited by Elinor Ostrom. Beverly Hills: Sage Publications, 1982. 179-222.
- Krauss, Melvyn. In "On the Brink: Third World Economies in Crisis." Manhattan Report on Economic Policy. 3(1), January 1983. Passim.
- Kress, Guenther, and Koehler, Gustav, and Springer, J. Fred. "Policy Drift: An Evaluation of the California Business Enterprise Program." In Implementing Public Policy, edited by Dennis J. Palumbo and Marvin A. Harder. Lexington: D.C. Heath and Company, 1981. 19-28.
- Levine, Robert A. "Program Evaluation and Policy Analysis in Western Nations: An Overview." In Evaluation Research and Practice: Comparative and International Perspectives, edited by Robert A. Levine, Marian A. Solomon, Gerd-Michael Hellstern and Hellmuth Wollman. Beverly Hills: Sage Publications, 1981. 27-60.
- Lipsky, Michael. "Standing the Study of Public Policy Implementation on Its Head." In American Politics and Public Policy, edited by W. Dean Burnham and Martha Wagner Weinberg. Cambridge: MIT Press, 1978. Chapter 16.
- Majone, Giandomenico, and Wildavsky, Aaron. "Implementation as Evolution." In Implementation, edited by Jeffrey L. Pressman and Aaron Wildavsky. Berkeley: University of California Press, 1979. Second edition, expanded. 177-197.

- Milward, H. Brinton. "Policy Entrepreneurship and Bureaucratic Demand Creation." In Why Policies Succeed or Fail, edited by Helen M. Ingram and Dean E. Mann. Beverly Hills: Sage Publications, 1980. 255-277.
- Munger, Michael C. The Costs of Protectionism: Estimates of the Hidden Tax of Trade Restraint. St. Louis: Center for the Study of American Business, Washington University. Working Paper Number 80, July 1983.
- Musheno, Michael C. "On the Hazards of Selecting Intervention Points: Time-Series Analysis of Mandated Policies." In Implementing Public Policy, edited by Dennis J. Palumbo and Marvin A. Harder. Lexington: D.C. Heath and Company, 1981. 77-89.
- Nagel, Stuart S. Policy Evaluation: Making Optimum Decisions. New York: Praeger Publishers, 1982.
- Nagel, Stuart S. The Policy-Studies Handbook. Lexington: D.C. Heath and Company, 1980.
- Nairn, Ronald C. Wealth of Nations in Crisis. Houston, Texas: Bayland Publishing, 1979.
- O'Brien, David M. "Crosscutting Policies, Uncertain Compliance, and Why Policies Often Cannot Succeed or Fail." In Why Policies Succeed or Fail, edited by Helen M. Ingram and Dean E. Mann. Beverly Hills: Sage Publications, 1980. 83-106.
- Olson, Mancur. The Rise and Decline of Nations. New Haven: Yale University Press, 1982.
- Ostrom, Elinor. "Institutional Arrangements and the Measurement of Policy Consequences in Urban Areas: Applications to Evaluating Police Performance." Urban Affairs Quarterly. June 1971. 447-475.
- Ostrom, Vincent. The Intellectual Crisis in American Public Administration. University, Alabama: The University of Alabama Press, 1973.
- Ostrom, Vincent, and Ostrom, Elinor. "A Behavioral Approach to the Study of Intergovernmental Relations." The Annals of the American Academy of Political and Social Sciences. 359, May 1965. 137-146.
- Ostrom, Vincent, and Ostrom, Elinor. "Public Choices and Public Goods." In Alternatives for Delivering Public Services: Toward Improved Performance, edited by E.S. Savas. Boulder, Colorado: Westview Press, 1977. 7-49.
- Palumbo, Dennis J., and Harder, Marvin A., eds. Implementing Public Policy. Lexington: D.C. Heath and Company, 1981.

- Patton, Michael Quinn. Qualitative Evaluation Methods. Beverly Hills: Sage Publications, 1980.
- Patton, Michael Quinn. Practical Evaluation. Beverly Hills: Sage Publications, 1982.
- President's Private Sector Survey on Cost Control. Task Force Report on the Department of STATE/AID/USIA. Washington, D.C., May 1983.
- Pressman, Jeffrey L., and Wildavsky, Aaron. Implementation. Berkeley: University of California Press, 1979. Second edition, expanded.
- Rein, Martin. "Comprehensive Program Evaluation." In Evaluation Research and Practice: Comparative and International Perspectives, edited by Robert A. Levine, Marian A. Solomon, Gerd-Michael Hellstern and Hellmuth Wollmann. Beverly Hills: Sage Publications, 1981. 132-148.
- Rutman, Leonard. Planning Useful Evaluations: Evaluability Assessment. Beverly Hills: Sage Publications, 1980.
- Schattschneider, E.E. The Semisovereign People: A Realist's View of Democracy in America. New York: Holt, Rinehart and Winston, 1960.
- Schattschneider, E.E. Two Hundred Million Americans in Search of a Government. New York: Holt, Rinehart and Winston, Inc., 1969.
- Schultz, Theodore W. Investing in People: The Economics of Population Quality. Berkeley: University of California Press, 1981.
- Scriven, Michael. Evaluation Thesaurus. Inverness, California: Edgepress, 1981.
- Sharp, Elaine B. "Models of Implementation and Policy Evaluation: Choice and Its Implications." In Implementing Public Policy, edited by Dennis J. Palumbo and Marvin A. Harder. Lexington: D.C. Heath and Company, 1981. 99-115.
- Smith, Nick L. "Metaphors for Evaluation." In Metaphors for Evaluation, edited by Nick L. Smith. Beverly Hills: Sage Publications, 1981. 51-65.
- Sorg, James D. "Pursuing Policy Optimization by Evaluating Implementation: Notes on the State of the Art." In Implementing Public Policy, edited by Dennis J. Palumbo and Marvin A. Harder. Lexington: D.C. Heath and Company, 1981. 139-154.

Stipak, Brian. "Using Clients to Evaluate Programs." Computers, Environment, and Urban Systems. 5(3/4); reprinted in Evaluation Studies Review Annual, 7, edited by Ernest R. House Sandra Mathison, James A. Pearsol and Hallie Prescoll. Beverly Hills: Sage Publications, 1982. 585-602.

Thomas, John Clayton. "'Patching Up' Evaluation Designs: The Case for Process Evaluation." In Implementing Public Policy, edited by Dennis J. Palumbo and Marvin A. Harder. Lexington: D.C. Heath and Company, 1981. 91-98.

Wells, Miriam J. "Success in Whose Terms? Evaluations of a Cooperative Farm." Human Organization. 40(3), 1981. 239-246.

Weiss, Carol H. "Measuring the Use of Evaluation." In Utilizing Evaluation: Concepts and Measurement Techniques, edited by James A. Ciarlo. Beverly Hills: Sage Publications, 1981. 17-33.

Wildavsky, Aaron. "Implementation in Context." In Implementation, by Jeffrey L. Pressman and Aaron Wildavsky. Berkeley: University of California Press, 1979. Second edition, expanded. 163-176.

Williams, Walter. The Implementation Perspective. Berkeley: University of California Press, 1980.

85