Policy Studies Review, Autumn 1987, Vol. 7, No. 1, 128-141

Charles J. Fox

BIASES IN PUBLIC POLICY IMPLEMENTATION EVALUATION

This essay will assess the main tendencies in the evaluation of public policy implementation. It will be suggested that the overall thrust of implementation studies are too negative. In turn, this implies and reinforces similar political beliefs that government can't do much good. With some notable exceptions, the sources of this negativism do not seem to be consciously antigovernment. Instances of blatantly ideological attacks on "fudge factories" are, at least in respectable academic publications, more rare than other sources of what this paper defines as bias. Sources of bias identified here are: 1) residual presuppositions favoring classical or rational-comprehensive policy implementation; 2) closely related, methodological and epistemological commitments to positivistic research protocols, often reinforced by disciplinary and academic reward structures; 3) the historical time frame within which most implementation studies have been conducted. The paper will conclude by recommending other directions, expanding our conceptualization of public policy in more optimistic, yet still critical, directions.

Before launching into the analysis, a few words of introduction are in order. First, this effort limits itself to the burgeoning sub-sub-field of public policy implementation. Many similar critiques of policy analysis and policy evaluation are well known (e.g., Beardsley, 1980; Diesing, 1982; Jennings, 1983; Mitroff, 1974; Mitroff & Kilmann, 1981; Mitroff et al., 1983; Rein, 1976; Weiss, 1983; for contrary views see McAdams, 1984; McCurdy & Cleary, 1984).

Second, the field of implementation studies is a relatively new one. Our understanding of implementation is not as mature as it is of policymaking processes. To an extent, this circumstance is itself the result of the persistence of rational-comprehensive assumptions in political science and public administration. Implementation of policies was long taken for granted. How a bill becomes a law was the focus, application of the law was supposed to be the automatic result of established agency procedures. Recognition of this knowledge gap can be virtually dated by the 1973 publication of Pressman and Wildavsky's *Implementation* and the subsequent outpouring of work from the "Berkely School" (Lee, 1978; not to slight earlier works of Bailey & Mosher, 1968; Derthick, 1972). Since then, "during the last decade or so, there has been a virtual explosion of research on the subject of implementation" (O'Toole & Montjoy, 1984).

Pressman and Wildavsky's work was primarily a case study of implementation of a federal program to reduce minority unemployment in Oakland, California, along with some middle level theoretical deductions derived from the experience. This seemed to set the mark toward which many subsequent case studies aspired (e.g., Koppes, 1983; Levin, 1985; Levy et al., 1974; Stiehm, 1981). Other influential works on implementation proceed the other way around; they present a conceptual scheme of implementation types and then place case studies within the categories (e.g., Edwards, 1980; Mazmanian & Sabatier, 1983; Nakamura & Smallwood, 1980; O'Toole &

Montjoy, 1984; Ripley & Franklin, 1986; Williams et al., 1982). In both sorts of literature, case studies are crucial. Indeed, in no other area of political inquiry of which I am aware, do case studies assume such importance.

RESIDUAL RATIONAL-COMPREHENSIVE PRESUPPOSITIONS

The rational-comprehensive or classical model of policy implementation is simple, clear, and elegant. Democratically elected officials make unambiguous policy choices. Policies are then handed over to an hierarchically structured agency. Specific instructions are formulated at the top of the pyramid and passed down the chain of command to the line personnel, who carry them out without discretion (Nakamura & Smallwood, 1980, pp. 7-9). Most observers know that this ideal is unachievable in reality. Still, as I will suggest, it is a backdrop ideal against which real world policy implementation is often measured. The further away from the ideal an implementation process is, the more likely will it be judged wanting.

It is difficult to be critical of such an appealing model. After all, shouldn't policy outcomes somehow fit policymakers' expectations? Aren't laws passed to influence society in some specific way? Shouldn't policy that doesn't work in expected ways be criticized and eliminated? Would we not abandon the fundamental standards of rationality itself if we did not expect policy implementation to produce predictable results? When posed this way, the common answer to these questions must surely be yes.

How, then, can commitment to rational-comprehensive implementation be thought of as an unwarranted bias? The point is well expressed by Dror (1986, p. 195): "Tendencies in the literature to overplay implementation difficulties can ... make implementation feasibility mapping into a conservative process inhibiting necessary policy innovations." Three aspects of residual commitments to rational-comprehensive implementation contribute to this "overplay": 1) unilateralism, 2) too high expectations, and 3) too narrow expectations.

Unilateralism

One aspect of preferences, based on the classical conception of policy implementation, is unilateralism (Fox, 1981; Mosher, 1968, pp. 204-05). In implementation literature this is often referred to as the "top down" perspective (Barnett & Fudge, 1981; Edwards, 1980; Mazmanian & Sabatier, 1983; O'Toole & Montjoy, 1984; Rosenbaum, 1980). Contemporary unilateralists will usually pay lip service to the demise of the Wilsonian dichotomy between politics and administration. They will nonetheless cling to the view that democratically elected officials have the lead role in deciding policy (see, for example, Pressman & Wildavsky, 1984, p. 215). The chore of implementation analysts is to advise them how their wishes can best be fulfilled. Here is not the place to exhaustively critique these assumptions but some summary statements can be adduced. First, it is not at all clear that American democratic processes produce policy mandates of sufficient clarity that they can be the concrete expectations against which implementation ambiguities can be found wanting (see, for instance, Bachrach, 1967; Bachrach & Baratz, 1970).

A second premise of unilateralism is that democratically elected policymakers ought to have their way in implementation (Mazmanian & Sabatier,

1983; Pressman & Wildavsky, 1984) because only they can be held accountable to the people. If policies are unacceptable, elected officials can be thrown out of office. However, serious doubts exist as to whether the attentiveness of the American electorate is sufficient to make such an ideal credible. And, even if the electorate paid such close attention, it is questionable whether the binary choices with which they are faced in elections allows requisite degrees of policy specificity to make accountability a concrete reality.

If these objections have weight, policy implementation as a process of mutual adjustment between individuals, agencies, interest groups, and levels of government can, in comparison, be assessed as, at least in a pluralist sense, quite democratic indeed. If mutual adjustment, compromise, and accommodation are judged to be democratic in the legislative policymaking sphere, there is no reason to deny a similar accolade to the implementation sphere. If the policy expectations of some centralized authority lack the legitimacy assumed for them, then failure to fulfill the mandates as expected is not as serious as it would otherwise have been (an intelligent but inconclusive discussion is Pressman & Wildavsky, 1984, pp. 206-31).

Too High Expectations

Partly because of residual commitments to rational-comprehensive implementation, many observers are disappointed when governments have difficulty implementing programs which attempt to ameliorate deep-seated social problems. Taken as a whole, the case study literature is sufficient to drive would-be social engineers to dispair. Programs designed to address inadequate housing (Derthick, 1972), unemployment (Pressman & Wildavsky, 1984), and school achievement (Coleman et al., 1966) rarely approach expectation levels. Secondary literature identifies implementation shortcomings pretty well. Ripley and Franklin (1986) building on Lowi's (1972) policy classification point out the intractability of problems toward which "redistributive" policies are directed and the political barriers which even monumental efforts cannot overcome. But, as they also reveal, implementation evaluation is itself one of the barriers:

> The most common judgements about redistributive programs [is that they] ... are failures unless they have profoundly altered human behavior at the end of one year. Naturally, the alterations sought are likely to take decades and are very difficult to measure at any point in a long evolution.

High expectations on one side of the scale, expectations that programs can and should have profound and immediate impacts on complex social problems, create negativism and defeatism on the other side. This supports the partisan drum beat about "forty years of failed programs," and "throwing money at social problems," (Schwarz, 1983). As Dror (1986, p. 160) puts it: "experience of failures with naive planning approaches has become ... a barrier to broad and long-term policy perspectives."

Expectations Too Narrow

Closely related to the unrealistically high expectations are those which are too narrow (a similar argument about the narrowness of economists'

concentration on microeconomic monetary measures is Rhoades, 1985, p. 818). In these instances, rational-comprehensive implementation assumptions lead one to expect policy (causes) to lead to specific measurable benefits (effects). The types of implementation evaluation to which I refer here are overly negativistic because they fail to incorporate sufficiently broad perspectives of what are admittedly ambiguous policy impacts. In implementation literature the pessimism-optimism continuum is best expressed through the distinctions between outputs, outcomes, and impacts (Levy et al., 1974; Nakamura & Smallwood, 1980; a different distinction between process evaluation and impact evaluation is made by Nachmias, 1979; Palumbo & Sharp, 1980; outputs, outcomes, and impacts of the former are subsumed under "impacts" of the latter). Policy outputs are services actually rendered: number of students taught, miles of highway paved, etc. Policy outcomes are the consequences of programs. These are longer-term effects such as increased reading ability of students in targeted school districts or better health for the elderly as a result of Medicare. Outcomes are difficult to assess and measure, but measurements are nonetheless often attempted. Policy impacts refer to even longer-term results which might be quite remote and may be aggregated with the results of several programs. Output assessments are usually noncontroversial but are not regarded by most analysts as very informative. It is in the area of outcomes where pessimism is often found, whereas impact analysis can often be optimistic depending largely on the skill or inventiveness of the observer. Here again, case studies are the data base.

It is the genre of outcome analysis that is, arguably, guilty of the charge of too narrow expectations based on residual classical implementation orientations. Three examples should suffice to make the point: 1) school desegregation and "white flight," 2) equalization of educational resources, and 3) urban fixed-rail mass transit systems.

School Desegregation and White Flight. It is, I believe, fair to summarize a prevailing view of desegregation and its implementation (outcomes) as follows:

- equal educational opportunity is an important national goal;
- separate provision of education for different races cannot be equal;
- schools must therefore be desegregated;
- other means failing, school desegregation must be accomplished by bussing;
- but, (outcome) bussing causes white flight and use of segregated private schools;
- therefore, the effort to desegregate has caused resegregation at higher ratios than had existed before (Pressman & Wildavsky, 1984, p. 221).

It follows that the policy implementation encountered such overwhelming resistance that it was a failure. Further, it can be implied that it should not have been attempted, or should now be abandoned.

The methodology of this outcome analysis has been disputed by some (Armor, 1978; Armor et al., 1977). But, even if we assume that the analysis is sound as it stands, a longer and broader impact analysis might put desegregation efforts in a different light. This is not the place to

perform such an analysis so a few suggestions must suffice to make the point. First, recent polling indicates a gradual and more positive shift in attitudes toward bussing. Blacks now approve of bussing, at least as a last resort by a 55 to 36 percent majority. On the other hand, whites oppose it by a 60 to 36 percent margin. Younger generation whites (18 to 30 years old) however, many of whom were exposed to bussing, opposed it by the narrow margin of 47 to 50 percent. These results seem to indicate that the further away from implementation inception one gets, the more favorably may the policy be viewed. Changes, which long-term impact analysis can pick up, may be generational.

Second, and more important, to consider bussing by itself may neglect important dimensions of it. Bussing, however unpopular initially, was a highly publicized exercise which underlined the nation's firm resolve to at least end *de jure* racial separation and the racism which it both presupposes and reinforces. Showing such resolve reinforced all of the other efforts with which it is associated, including desegregation of facilities, transportation, and the armed services, as well voting rights and affirmative action. More broadly yet, it is hard to imagine that the United States could do as much for human rights abroad as it has done without such highly publicized efforts.

Equalization of Educational Resources. There is evidence that at differing rates, states are slowly but surely equalizing educational resources across their districts. Influential policy analysis has disparaged these efforts. As Wildavsky has written "education cannot significantly improve the cognitive abilities of its most challenging children: and, when these children do make it through higher education, that education cannot guarantee to improve their life chances" (1979, p. 316). This is based on findings that inputs into educational institutions somehow, in the "black box," (Coleman, 1975, pp. 36-37) are absorbed or dissipated without producing measurable outcomes. Educational success seems determined by family background and schools can do very little about it (Coleman et al., 1966). The logic runs like this:

- the objective of improving educational resources is to increase cognitive development;
- this can be measured by tests;
- test scores do not indicate increased rates of cognitive development;
- therefore, the resources have been wasted.

Always tough minded, Wildavsky (1979, p. 317) puts it like this: "the greatest part of new expenditures goes into teachers' salaries. Why should anyone believe that paying those same teachers more to do more of the same will lead to appreciably better results?"

Again, rational comprehensive implementation, a given cause should produce a given effect, neglects many factors. The more broadly-cast net of impact analysis seems to be a necessary corrective. It is, of course, good to know that enhanced revinues do not produce better scores. Knowledge can produce better targeted initiatives. But is that all we should consider? (Diesing, 1982, pp. 206-212).

At the risk of being one of those bureaucratic folk that Wildavsky (1979, p. 317) charicatures as saying: "now that we have lost sight of our

objectives ... we must redouble our efforts," I think certain other factors should be considered:

- it may be argued that test scores do not in fact adequately assess cognitive development especially with regards to minority children;
- even if they do, other important types of child development may be enhanced (Bronfenbrenner, 1979);
- even if other types of development are equally intractable, why should blighted neighborhoods be further degraded, contributing to the downward spiral, by blighted schools?
- if local jurisdictions cannot pay their teachers at the going rate, why can't a more general level of government do so for reasons of equal pay for equal work?
- what other (private) uses of the resources are so important that educational use of them can be thought to be wasted?
- can it not be rationally argued that clean, well equipped schools are elements of public wealth and valuable unto themselves?

Urban Fixed-Rail Mass Transit. As I have been arguing, implementation evaluation, by focusing on too narrow measurements of outcomes tends to discourage government projects. It contributes to what John Kenneth Galbraith has called private opulence and public squalor. As an example, I would like to consider an extreme case, building new mass transit systems. The scope of such projects rivals the building of the pyramids.

The implementation of the development of systems like the Bay Area Rapid Transit (BART) system or Washington D.C.'s Metro defies positive outcomes analysis (see the excellent study by Webber, 1980) and taxes even the most creative impact analysts. Building them will always be more costly than anticipated; they will be plagued by cost overruns. While being built they will disrupt life and hamper businesses, some times for more than a decade. After being built they will cost too much to operate. They will not come close to competing on a cost-effective basis with busses. They will be only marginally competitive with single driver subcompacts. They are less convenient than cars and, given the choice, cars will be preferred by the public. Contradicting presystem claims, they will not reduce traffic congestion. Building mass transit systems is a triumph of will over economic rationality.

Despite all this, someone up there likes them. They keep popping up on urban area agendas. Can this be totally irrational? Perhaps an impact analysis by a Lewis Mumford is needed. I offer the following conjectures:

- the majestic scope of such efforts lifts the spirit much like putting a man on the moon;
- they generate civic pride and, consequently, economic activity;
- in increasingly homogenized physical city-scapes--same fast food restaurants, same store chains, same freeways--new mass transit gives a city something unique;
- with increasingly compartmentalized life-styles, mass transit facilities are people-gathering places, windows on life, that satisfy that aspect of our natures that Aristotle called social.

These are conjectures and they could very well be wrong, but too narrowly conceived outcome analyses are, in the Freudian sense, anal retentive. I will assert that no matter how much busses can be proved cost-effective, a capitol city like Washington D.C. should have a Metro, just as it should have noncost-effective monuments and buildings with noncost-effective ornamentation. Implementation evaluation should not be allowed to negate aesthetics.

RESEARCH PROTOCOLS

We have been pursuing the position that evaluation of public policy implementation is, on the whole, too negative. This tends to support the conservative view that government is a negative influence on society; when it tries to do something, even with the best of motives, it invariably "screws up." In the previous section I suggested that part of this negativism was due to residual commitments to rational-comprehensive assumptions about how implementation ought to work. This section is closely related. As Tribe (1972, pp. 66-110) has argued, positivistic policy analysis and highly bureaucratic and rationalistic forms of public administration tend mutually to reinforce each other. In other words, the way we think about knowledge and the proper way to obtain it (epistemology) bolsters rational-comprehensive backdrop assumptions of how implementation ought to proceed. In this section I will first argue that the effects of logical positivism/empiricism are alive and well and outline what the basic assumptions are. Second, I will show specifically how these presuppositions affect policy evaluation and, hence, evaluation of implementation in ways I regard as excessively pessimistic.

Logical Positivism/Empiricism

In philosophy of science, logical positivism/empiricism (LP/E) no longer holds the high ground. The last two decades have witnessed a sustained attack on its basic premises (see, for instance, Fox, 1980; Gunnell, 1975; Kuhn, 1970). Jennings (1983) hits the mark exactly:

> ... if positivism and empiricism are now on the wane as epistemologies, the research procedures and protocols that have grown out of them remain firmly institutionalized in legislative and administrative requirements, funding structures, graduate curricula and professional reward systems. Positivism is the basic methodological orientation of the social sciences and will surely remain so for some time, even though it has now lost the secure philosophical warrant it once enjoyed.

LP/E is based on four canons: 1) the objective-subjective dichotomy; 2) the fact-value dichotomy; 3) the nomothetic deductive theory of explanation/prediction; and, 4) the faith that social science knowledge can be ahistorically cumulative. Full refutation of each of these cannons cannot be replicated here (for a full treatment see Fox, 1980). A few comments about each must suffice to show how they have been delegitimized.

The objective-subjective dichotomy presupposes an inside (subjectivity) outside (objectivity) bifurcation that cannot, in the light of either logic or perceptual theory (Merleau-Ponty, 1962) be sustained. Objects, at least

as we encounter them, cannot be known apart from our intentional projects. Subjectivity, on the other hand, is filled out by contact with, and also by transforming, objects. There is a dialectical relationship between the two (if they can be seen as separable at all) which confounds the dichotomy relied upon by LP/E.

The fact-value dichotomy is subsidiary to the objective-subjective one and suffers from its erosion. Facts do not just lie about, discrete and invariant, waiting to be discovered to confirm or disconfirm hypotheses. From an infinite number of possible facts available to one's perceptual apparatus, some are chosen precisely because they are *valuable* to our intentional projects. It follows that values of researcher/analyst can never be fully separated from so called "objective" or "factual" confirmations.

The nomothetic-deductive model of explanation/prediction is difficult to sustain without the previous two canons. Laws, theories, hypotheses, and predictions are to be confirmed or disconfirmed by the "objective facts." If these are not independent of subjectively derived laws, theories, hypotheses or predictions, then confirmation and disconfirmation are not as secure as LP/E presumes. Self-fulfilling predictions are as likely as anything else. Independent objective status must be denied.

The belief that social science knowledge can be ahistorically cumulative is undermined by the problems of the previous three principles of LP/E. If knowledge cannot be purely objective or factual, and if values inevitably intrude in the process of inquiry, then the culture of the period must influence values and be incorporated into subjectivity. Social science truths, both because the researcher is historically situated and because the social facts that are the object of the research will vary over time, cannot, even in principle, be immutable, essential or universal. It cannot be the case that laws of social interaction can be built up over time to approach perfect knowledge. And it is probably a good thing that they cannot be, for perfect knowledge would mean that social interrelations would be perfectly predictable and hence subject to perfect control by whoever has the power to use such knowledge.

Lingering Influence of LP/E

Despite the demise of philosophical or epistemological warrant, policy research proceeds with an inflated self-assurance. Ripley (1985, pp. 9-11) is typical. To him good science means data, hypothesis testing, establishment of cause and effect, search for regularities, replication, and designing research based on the "belief that knowledge is cumulative." Nachmias' (1979, pp. 3-4) influential methods text similarly advises that:

> ... one method that can reduce the number of erroneous [policy] decisions is the formal scientific approach to knowledge. Viewed from the scientific perspective, policy evaluation research is the objective, systematic, empirical examination of the effects ongoing policies and public programs have on their targets in terms of the goals they are meant to achieve. (see also Langbein, 1980)

Juxtaposing these sentiments with discredited LP/E canons shows that "good science" need not be what these authors suggest. But so what? What harm do they do? How may they be regarded as distorting biases?

First, commitment to what many, along with Ripley, think of as "good science" has priority over the problematic for which inquiry has been called. Weiss (1978, p. 45) is on target: "researchers often choose and conceptualize problems in terms of the methodologies in which they are proficient. They do not pick the research method to suit the problem but almost unwittingly see that aspect of the problem to which their methodology applies" (see also, Ripley, 1985, p. 178). That is to say, the methodological tail is called upon to wag the problematic dog. Commitments to methods inspired by positivism delimits our capacity to properly analyze the impacts of policy implementation.

Second, in addition to the priority of methods over problematics, positivistic approaches unwittingly contribute to the negativism with which this paper is primarily concerned. In line with the four categories in Section 1, methods can reinforce negativism because they are associated with measures that are too broad to discover positive policy innovations or too narrow to see effects of policies marching together on a broad front.

A good example of the first is what Warwick and Pettigrew (1983, p. 344) identify as the "total effect" fallacy. As they point out: "the repeated preference of some federal agencies to fund policy oriented research with one type of method--large-scale sample surveys in cross-sectional designs ... insures that few significant results of nationwide social programs will be uncovered." With respect to evaluating implementation of the Head Start program, evaluators asked whether the total program had significant outcomes rather than asking which of the many local designs were working. In other words, evaluations should have been disaggregated. This would have required much more field work and involved qualitative or interpretive social science research designs, frowned upon by positivistically-inspired methodological protocols.

Research design delimitations and disciplinary parochialism generally make policy evaluation too narrow to capture policy-provoked social evolution.

Research done under the strictures of hypothesis testing must take hypotheses atomistically, one at a time (Dror, 1986, p. 230). The LP/E assumption of cumulative knowledge makes this acceptable because each researcher, working in a small particular corner of the vineyard produces knowledge that will later be aggregated. But aggregation in the desired sense rarely occurs. The Great Society programs, the war on poverty, were broadly based attempts to ameliorate adverse social conditions on a broad front. The programs, however, were evaluated piece meal. Does this mean we should not have made the effort? Conservatives point to the failures and answer yes: social engineering doesn't work. A nonpositivistic, broader-based analysis (e.g., Schwarz, 1983) shows that taken together, the programs have had a marked effect on reducing poverty. An even broader analysis would perhaps show that making the effort itself, no matter how impotent in its particulars, changed the very nature of our society and its attitudes. Networks of people, for instance, were formed in many communities the effects of which will not be fully known for decades. The almost spiritual aspects of the broad front of activism, the feeling that things could be accomplished, that pollution could be controlled, that racism could be overcome will, despite backlashes, have lasting social effects. Not all of these can be captured by positivistic methods, however otherwise useful.

HISTORICAL TIME FRAME

Most of the implementation studies which have been cited as having an overall pessimistic effect, were published in the last 10 to 15 years. In this section I will suggest that pessimism about government capacities is related to the historical time frame within which it occurs.

Pressman and Wildavsky's trend-setting and epoch-marking study of the Oakland project was published in 1973. This was the time of the Watergate scandal, the winding down of a high profile policy failure in Southeast Asia and the end of the "American Century." Intractable economic difficulties, typified by stagflation, undermined confidence in the ability of government to regulate the economy to provide stable prices, high levels of employment and growth of G.N.P. and productivity. The post-World War II boom was over and things began to turn sour (Heilbroner, 1975).

It may be laid down as a general principle that everything works a little better when the economy is expanding and a lot worse when it is stagnating. When the pie is growing, people worry less about shares than when it is static. Affirmative action is more likely to conflict with seniority when RIFs become rampant. Job training programs work better when there are jobs to be filled. Educational aspirations will be higher when the rewards for it are more evident--when role models are working in jobs rather working the streets. Implementors at all levels of government are more cooperative when energy is not expended justifying their staffing levels or their very existence; new tasks will be more eagerly assumed during expansions, resisted more vigorously when existing services must be provided with diminishing resources. In short, most implementation short-comings were discovered during what, for Americans who had lived through the postwar boom, were, by comparison, hard times.

Here again, the positivist assumption that knowledge can be cumulative, makes negative findings about implementation seem ahistorically general when they may actually be more of a reflection of the times. The particular historicity is neglected and we are led to believe that the dysfunctions are permanent principles.

A second aspect of the particular times can be adduced; a peculiarity of Johnson's Great Society programs which form the evidentuary bulk of implementation case studies. As Weiss (1983, p. 216) has noted, every major program of the War on Poverty was accompanied by funds and requirements for program evaluation. Great Society programs are overrepresented in the case study literature. One way to get at this peculiarity, is through Dror's (1986, p. 98) concept of "fuzzy gambling," or "fuzzy betting." The somewhat audacious attempts of policy to influence the future must always be problematic because we cannot fully understand what it is we are trying to influence, nor the measures by which we try to do it. Many consequences of our attempts must be unanticipated. Nonetheless, as Raney (1976, p. 147) has written: "if the alternative ... is the resigned acceptance of human impotence and passivity in the face of human misery, then surely political engineering, with all its faults and failures, beats that." Dror proposes that we think of policymaking as a form of fuzzy betting that we can positively influence the flow of human events.

I suggest that it was precisely the above attitude that permeated Johnson era policy innovations. It was the legacy of the New Deal, where Johnson's political consciousness was formed. It is a pragmatic orientation

which calls for action whenever and wherever you can get it. As historians of the Roosevelt presidency have informed us, Roosevelt didn't know how to solve the problems of the Great Depression so he resolved, in a flurry of activity, to try something, try anything. So a myriad of programs were launched, redundancy and competitive overlap were encouraged and even designed. This was also true of Johnson's programs.

Implementation evaluation, picking over the separate pieces resulting from this flurry is, as has been argued above, ill-suited to pass favorably on them as a whole. And, what is important here, programs were launched on the upswing side of the American Century, but evaluated on the down side. Try anything you can get through Congress on one side; husband scarce resources on the other side. On the upside is Johnson telling the people that the Great Society is within reach; on the down side is President Ford offering Boy Scout and Farmers' Almanac hints on how to tighten our belts (the WIN program).

Another aspect of "fuzzy betting" merits mention. It requires long-term perspectives. "Inbuilt long lead times, as well as long implementation cycles, require policy time horizons much longer than usual in practice" (Dror, 1986, p. 159). Implementation problems discovered by evaluators are real enough, but they might be ameliorated over time. Institution and network building often take longer periods of time than implementation studies can be sustained. Organizational learning is also a long-term process (Hirschman, 1965). If commitments to programs wane before networks are allowed to mature, then implementation will suffer and our evaluations of them will necessarily be negative. From the point of view of "fuzzy betting," implementation studies conducted during the early stages of program implementation are also distorting. Working out relationships within a "government of strangers," (Heclo, 1977) "ain't pretty."

RECOMMENDATIONS FOR IMPROVING IMPLEMENTATION STUDIES

If the above analysis has validity, excessive negativism can be overcome by: 1) not judging the human process of actual implementation against unrealistic, mechanistic, rational-comprehensive expectations; 2) not concentrating our inquiry in positivistic modes; and 3) giving greater play to the historicity of policy incidents.

1. We should continue to expect results from policy interventions. We should continue to be critical of agencies the bureaucratic processes of which prevent the job from getting done. But it should be expected that policy impacts are more like the ripples caused by throwing a pebble in a pond than a cue impacting a billiard ball. A more egalitarian assessment of the implementation players might also be helpful. We should not automatically assume that expectations developed at the pinnacle of an hierarchy are inordinately superior to expectations developed at lower levels. Multiple levels of accountability, not the unilateralist view, can be assumed as bargaining between the players occurs, as long as it occurs in good faith.

2. There can be no doubt about the fecundity of positivistic research on public policy. The excellent studies collected in Nachmias (1980), for instance, are as good as they can be. But, room must be made in the journals, and academic rewards should follow, and graduate curricula should incorporate, nonpositivistic research modes. Much more emphasis

on interpretive social science and policy analysis would be salutary. As Mitroff (1974) and associates (1981, 1983) have argued, we need multiple ways of learning and knowing to permit inclusion of appropriate tacit and subjective knowledge (Dror, 1986, p. 148; see also, Dallmayer & McCarthy, 1977; Jennings, 1983, p. 28; Rabinow & Sullivan, 1979).

3. We should avoid the pitfalls of assuming that implementation dysfunctions are self-contained. We should look for changes in regime norms or *geist* changes as possible contributing factors. Also, we should be wary of believing that a cumulative social science can gather up universally valid principles of implementation and/or its defects. It is more likely that organizational learning will transcend one sort of dysfunction only to encounter other sorts.

REFERENCES

Armor, D.J. (1978). *White flight and the future of school desegregation*. Santa Monica, CA: Rand Corporation.

Armor, D.J., Pettigrew, T.F., & Wilson, J.Q. (1977). On bussing: An exchange. *The Public Interest, 28*.

Bachrach, P. (1967). *The theory of democratic elitism*. Boston: Little Brown.

Bachrach, P., & Baratz, M. (1970). *Power and poverty*. New York: Oxford University Press.

Bailey, S.K., & Mosher, E.K. (1968). *ESEA: The Office of Education administers a law*. Syracuse, NY: Syracuse University Press.

Barnett, S., & Fudge, C. (Eds.). (1981). *Policy and action*. London: Methuen.

Beardsley, P. (1980). *Redefining rigor: Ideology and statistics in political inquiry*. Beverly Hills, CA: Sage.

Bronfenbrenner, U. (1979). Head Start, a retrospective view: The founders. In E. Zigler & J. Valentine (Eds.), *Project Head Start: A legacy of the war on poverty*. New York: Free Press.

Coleman, J.S. (1975). Problems of conceptualization and measurement in studying policy impacts. In K.M. Dolbeare (Ed.), *Public policy evaluation*. Beverly Hills, CA: Sage.

Coleman, J.S., et al. (1966). *Equality of educational opportunity*. U.S. Department of Health, Education, and Welfare, Office of Education, Washington, DC: U.S. Government Printing Office.

Dallmayer, F.R., & McCarthy, T.A. (Eds.). (1977). *Understanding and social inquiry*. Notre Dame, IN: University of Notre Dame Press.

Derthick, M. (1972). *New towns in town*. Washington, DC: Urban Institute.

Diesing, P. (1982). *Science and ideology in the policy sciences*. New York: Aldine.

Edwards, G.C., III. (1980). *Implementing public policy*. Washington, DC: C.Q. Press.

Fox, C.J. (1980, September). The existential-phenomenological alternative to dichotomous thought. *Western Political Quarterly, 33*, 357-379.

Fox, C.J. (1981). Civil service reform and ethical accountability. *Public Personnel Management, 10*(1).

Gunnell, J.G. (1975). *Philosophy, Science and Political Inquiry*. Morristown, NJ: General Learning Press.

Heclo, H. (1977). *A government of strangers: Executive politics in Washington*. Washington, DC: Brookings.

Heilbroner, R.L. (1975). *An inquiry into the human prospect*. New York: Norton.

Hirschman, A.O. (1965). *Journeys toward progress*. Garden City, NY: Doubleday Anchor.

Jennings, B. (1983). Interpretive social science and policy analysis. In D. Callahan & B. Jennings (Eds.), *Ethics, the social sciences and policy analysis*. New York: Plenum.

Koppes, C.R. (1983). *JPL and the American space program: A history of the Jet Propulsion Laboratory*. New Haven, CT: Yale University Press.

Kuhn, T. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.

Langbein, L.I. (1980). *Discovering whether programs work: A guide to statistical methods for program evaluation*. Santa Monica, CA: Goodyear.

Lee, K. (1978). Review of Eugene Bardach, the implementation game. *Policy Sciences, 10,* 225-226.

Levin, M.A., & Ferman, B. (1985). *The political hand: Policy implementation and youth employment programs*. New York: Pergamon.

Levy, F.S., Meltzner, A.J., & Wildavsky, A. (1974). *Urban outcomes*. Berkely: University of California Press.

Lowi, T.J. (1972, July-August). Four systems of policy, politics, and choice. *Public Administration Review, 32,* 298-310.

Mazmanian, D.A., & Sabatier, P.A. (1983). *Implementation and public policy*. Glenview, IL: Scott Foresman.

McAdams, J. (1984, September). The anti-policy analysts. *Policy Studies Journal, 13*(1), 91-101.

McCurdy, H.E., & Cleary, R.E. (1984, November-December). A call for appropriate methods. *Public Administration Review, 44*(4), 533-554.

Merleau-Ponty, M. (1962). *The phenomenology of perception*. London: Routledge and Kegan Paul.

Mitroff, I. (1974). *The subjective side of science*. New York: Elsevier.

Mitroff, I., & Kilmann, R.H. (1981, March). The four fold way of knowing: The varieties of social science experience. *Theory and Society, 10,* 227-248.

Mitroff, I., Quinton, H., & Mason, R.O. (1983, June). Beyond contradictions and consistency: A design for a dialectic policy system. *Theory and Decision, 15,* 107-120.

Mosher, F.C. (1968). *Democracy and the public service*. New York: Oxford University Press.

Nachmias, D. (1979). *Public policy evaluation: Approaches and methods*. New York: St. Martin's.

Nachmias, D. (Ed.). (1980). *The practice of policy evaluation*. New York: St. Martin's.

Nakamura, R.T., & Smallwood, F. (1980). *The politics of policy implementation*. New York: St. Martin's.

O'Toole, L.J., Jr., & Montjoy, R.S. (1984, November-December). Interorganizational policy implementation: A theoretical perspective. *Public Administration Review, 44*(6), 491-503.

Palumbo, D., & Sharp, E. (1980). Process versus impact evaluation of community corrections. In D. Nachmias (Ed.), *The practice of policy evaluation*. New York: St. Martin's.

Pressman, J.L., & Wildavsky, A. (1984). *Implementation* (3rd ed.). Berkely: University of California Press.

Rabinow, P., & Sullivan, W.L. (Eds.). (1979). *Interpretive social science: A reader*. Berkely: University of California Press.

Raney, A. (1976, March). The divine science: Political engineering in American culture. *American Political Science Review, 70,* 140-148.

Rein, M. (1976). *Social science and public policy*. New York: Penguin.

Rhoades, S.E. (1985, November-December). Do economists overemphasize monetary benefits? *Public Administration Review, 45*(6).

Ripley, R.B. (1985). *Policy analysis in political science*. Chicago: Nelson Hall.

Ripley, R.B., & Franklin, G.A. (1986). *Policy implementation and bureaucracy* (2nd ed.). Chicago: Dorsey.

Schwarz, J.E. (1983). *America's hidden success: A reassessment of twenty years of public policy*. New York: Norton.

Stiehm, J.H. (1981). *Bring me men and women: Mandated change at the U.S. Air Force Academy*. Berkely: University of California Press.

Tribe, L.H. (1972). Policy science: Analysis or ideology. *Philosophy and Public Affairs, 2,* 66-110.

Warwick, D.P., & Pettigrew, T.F. (1983). Toward ethical guidelines for social science research in public policy. In D. Callahan & B. Jennings (Eds.), *Ethics, the social sciences and policy analysis*. New York: Plenum.

Webber, M.M. (1980). The BART experience. In D. Nachmias (Ed.), *The practice of policy evaluation*. New York: St. Martin's.

Weiss, C.H. (1978). Improving the linkage between social research and public policy. In L.E. Lynn, Jr. (Ed.), *Knowledge and policy: The uncertain connection*. Washington DC: National Academy of Sciences.

Weiss, C.H. (1983). Ideology, interests, and information: The basis of policy positions. In D. Callahan & B. Jennings (Eds.), *Ethics, the social sciences and policy analysis*. New York: Plenum.

Wildavsky, A. (1979). *Speaking truth to power*. Boston: Little Brown.

Williams, W., et al. (1982). *Studying implementation: Methodological and administrative issues*. Chatham, NJ: Chatham House.