Policy Evaluation

Two Aspects of Evaluation

The notion of evaluating policy emerged in its modern form in the 1960s and 1970s and flourishes today in the virtually worldwide adoption of managerialist methods such as public sector management, auditing, and human resource management. Like so many other terms, definitions of evaluation abound, but Dye makes a useful, broad and general offering: policy evaluation is learning about the consequences of public policy. Parsons suggests that today there are two main aspects to policy evaluation:

- Evaluation of policy and its constituent programs
- Evaluation of people who work in the organizations responsible for implementation

The first, evaluation of policy and programs, is a form of rational analysis, the second a tool for human resource management.

Rational Evaluation and the Policy Cycle

Rational policy evaluation is intended to be a process of objective, systematic and empirical examination of the effects ongoing policies and public programs have on their targets in terms of the goals they are meant to achieve. Evaluation research addresses two dimensions: how a policy may be measured against the goals it sets out to attain (goal performance), and the actual impact of the policy. Evaluation techniques include cost-benefit and utility analyses, performance measurement and experimental evaluation.

Palumbo suggests that we may better understand the role of these techniques by relating the policy cycle to the information cycle:

![Figure 4.19: The policy cycle and the information cycle](source)

Source: Adapted from Palumbo (1987: figure 1.1)

‘Formative’ Evaluation during Implementation

Evaluation takes place when a policy or program is being implemented. It involves monitoring the way the program is being administered or managed - a ‘formative’ evaluation process which provides feedback to improve the implementation process itself. It is likely to include exploration of whether the program is reaching the target population, whether it is meeting design specifications and how resources are being managed. Various forms of management information systems (MIS), and financial management techniques are likely to be applied.

The usual emphasis is on performance measurement and the need to control public finances and attain higher levels of value for money, efficiency and effectiveness. This control approach brings with it a number of problems, specifically because control is not a neutral or purely technical technique. The performance indicators which are frequently used are heavily loaded with values, politics and power - who determines what is efficient and effective or what is value for money? Other issues include:

- Comparison - can one hospital be compared with another to determine its performance?
- Quantity v. non-quantitative dimensions - would more teachers, or better-trained teachers deliver more benefits?
- Fairness and equity - just because there’s more of something doesn’t mean it’s equitably distributed

Measures of performance in themselves mean nothing. The policy analyst needs to be concerned with the values and politics which have constructed the measures and the interpretation which is placed upon data. It has been argued that performance indicators merely increase the capacity of the state to control organizations and people, shifting power into the hands of auditors and accountants.

‘Summative’ Evaluation of Impacts

Summative evaluation generally involves post-implementation assessment of the net and gross effects of the program. This is essentially a comparative mode of inquiry:

<table>
<thead>
<tr>
<th>Comparison Point A</th>
<th>Comparison Point B</th>
</tr>
</thead>
<tbody>
<tr>
<td>before</td>
<td>after</td>
</tr>
<tr>
<td>impact on one group</td>
<td>impact on another group</td>
</tr>
<tr>
<td>what happened</td>
<td>what would have happened without the program</td>
</tr>
</tbody>
</table>

Like formative evaluation, summative evaluation is laden with values, beliefs, politics and ideology and the outcome is particularly dependent upon the interpretation placed on any
data or ‘evidence’. Often the evaluation process focuses on quantification and experimentation - trying out a policy on a control group and a test group. Experiments in the social context are inevitably subject to a vast array of complex and uncontrollable variables, so results will virtually always remain questionable and open to various interpretations.

Despite these limitations, quasi-experimentation by government in highly complex problems is something which continues to prove an attractive option. Dye points out that it is exceedingly costly for society to commit itself to large-scale problems and policies in education, welfare, housing, health and so on, without any real idea about what works.

**Evaluation and Human Resource Management**

While much evaluation literature focuses on the assessment of programs and policies, in a managerialist framework it also encompasses human resource management and the evaluation of people through the use of techniques such as:

- Performance pay schemes
- Personnel assessment and appraisal
- Organizational development strategies

Human resource management (HRM) is the latest in a long line of experiments with rational techniques in policy evaluation - but rather than seeking to improve the rationality of decision-making and the ‘system’, HRM is concerned with improving rationality by changing the motivation, culture and attitudes of people.

**Goals of Human Resource Management**

One textbook identifies the aim of HRM as being to change people so that they become:

- Committed
- Competent
- Cost-effective
- In sympathy with the aims of the organization

The HRM approach stresses that the goal is to improve ‘performance’ by developing a sense of commitment in each employee, rather than just compliance to hierarchical command or instruction. This increase in commitment is achieved through:

- recruiting the right kind of people
- an emphasis on training at all levels
- regular staff appraisal
- rewarding performance

This coupling of performance and reward has led to the linking of the carrots and sticks of training and appraisal to monetary incentives. Performance-related pay systems are designed to create a framework in which individuals and/or groups can be evaluated in
terms of the extent to which they have met defined and measured goals and objectives and contributed to the bottom line and profits. Considered in a public sector context, this raises questions as to whether money and self-interest are the best kind of motivator for public servants.

The emergence of HRM can also be seen as an extension of the power of the state to survey and control organizational processes and their target ‘customers’ and ‘clients’. Through HRM governments have increased their capacity for:

- information collection, storage and retrieval
- the supervision of people
- the monitoring and control of behaviour

From this perspective, evaluation techniques are essentially modes of altering ‘culture’ and behaviour, rather than ‘systems’ or organizational mechanisms. Evaluation has become surveillance. At the core of evaluation which is focused on ‘people’ rather than policies or programs is the belief that more control is needed over bureaucrats and professionals to ensure that, individually and collectively, the objectives which are defined by policy makers are implemented efficiently and effectively.

Alternative Frameworks for Evaluation

Mainstream evaluation has been framed by positivist assumptions about knowledge and methods. The fundamental predicate of evaluative techniques is that it is possible to obtain objective measurements and data which can then be fed into policy cycle stages (plan, monitor, control, report, evaluate) in an uncontroversial way.

However, the field of evaluation research is now characterized by an identity crisis and alternative models and approaches have sprung up which are more fragmented and alive to the political and value-based nature of the activity. Program or problem analysis is seen by many modern critics as essentially a political process, full of values rather than some kind of scientific quest for truth or an objective answer.

Three Approaches in ‘Post-Positivist’ Evaluation

Several new approaches advance a radically different view of what evaluation and policy analysis can and ought to be in a democratic society.

The ‘Multiplist’ Approach

Cook has argued that as there can be no ‘correct’ policy option or evaluation. Rather, we should seek to use multiple measures from several different approaches and methods. Modes of analysis should access a wide variety of options and information, with resulting sets of data compared and pitted against each other to see which one is superior or most useful.
Evaluation in this sense involves testing arguments and claims to knowledge, rather than advancing ‘the truth’ or ‘the correct solution’.

The ‘Design’ Approach

Multiple frameworks are also central to the ‘design’ approach advocated by Miller, Bobrow and Dryzek. This approach begins with the notion that policy-making is an activity which is about the pursuit of values or goals. Human beings ‘design’ the ‘reality’ which surrounds policy-making so therefore the idea that we can be objective is an erroneous basis on which to make any kind of evaluation. The design approach rejects the idea of analysis as neutral or the belief that policy problems can be studied in a positivistic way.

Dryzek argues that policy analysis must attend to the political process with which analysis and policy are involved. Analysis should focus on open communication, unrestricted participation and clear-eyed examination of the role of multiple values at work in a policy-making process located in a multiple reality. He proposes a recursive process of analysis by design which could include:

- addressing values
- capturing the context
- selecting appropriate analysis frameworks
- applying the appropriate approaches

The ‘Naturalistic’ Approach

A more radical model of evaluation has been developed by Guba and Lincoln, who describe four ‘generations’ of evaluation:

- Technical - evaluation as a technical exercise of measurement
- Descriptive - evaluation as a process for describing patterns, strengths and weaknesses of stated objectives
- Judgment - evaluation by measuring and judging efficiency and effectiveness against standards and objectives
- Responsive - evaluation focused on claims, concerns and issues put forward by members of policy stakeholding audiences

They suggest that we are now entering the responsive evaluation stage which will utilize processes such as illuminative evaluation, utilization-focused evaluation, adversarial evaluation and naturalistic evaluation. A collaborative process with stakeholders becomes the primary focus of inquiry and a multiplicity of constructions is examined in a political process which recognises that knowledge is the result of negotiation. Evaluation becomes a learning/teaching process and the role of the evaluator is to mediate and facilitate change.
References