Eight Models of Evaluation Use

Go Yoshizawa
Graduate School of Public Policy, University of Tokyo
g-yoshizawa@pp.u-tokyo.ac.jp

Abstract

A vast reservoir of knowledge and experience from utilization studies developed since the 1970s illuminates political and enlightenment functions in evaluation. This is virtually ignored by Japanese public agencies who report how they have used policy evaluation in a straightforward manner. As a bridgehead for applying the concept of evaluation use to evaluation studies and practices for public policy in Japan, this paper introduces independent/user-oriented, concrete/abstract (the process of use), substantive/symbolic (the purpose of use) and findings/process (the object of use) discriminators. With these discriminators and the appreciation of studies by policy analysts such as Weiss and Whiteman, this study illustrates eight different models of evaluation use with a novel, comprehensive matrix. Apart from the conventional typology of use including instrumental, conceptual and symbolic, this matrix gains a balanced perspective on the substantive and political nature of evaluation use.

Keywords

evaluation utilization, utilization studies, substantive use, symbolic use, politicization of evaluation

1. Introduction

How does the research and practice of evaluation help society? In Japan, for instance, the Basic Guideline for Policy Evaluation (Cabinet Decision, December 16, 2009) says “it is necessary for individual administrative agencies to adequately utilize the result of policy evaluation as important information in the design and planning of policy and appropriately reflect it on the policy concerned”. Based on this, each agency reports how they have used policy evaluation. However, such reports are rather simplistic and only few of them are academically conducted independent from government authorities. By the same token, there have been very few reflexive analyses of evaluation studies and practices in Japan. Evaluation activities are strategically directed toward decision-making and problem-solving and they acquire specific knowledge relevant to the political context (Alkin and Taut 2003). Therefore, we must discern the appropriate use of evaluation more carefully than that of other social knowledge and academic studies. This is why the social needs for evaluation utilization studies have increased in recent years (Alkin 2003). A trendy concept called reflexive governance (Voß, Bauknecht and Kemp 2006) emerges from modern social
science studies, according to which we should pursue an approach in ways that evaluation studies contribute to the social betterment.

For this purpose, the present paper focuses on and discusses the use of evaluation. To classify use, the next section first introduces two dichotomies relating to the process of use - independent or user-oriented, and concretely used or abstractly used. It then introduces discriminators of whether evaluation is used with a substantive intention or with a symbolic intention (the purpose of use), and of whether the evaluation findings are used or the evaluation process is used (the object of use). The last section compiles these four dichotomies in a matrix and develops the discussion.

2. Taxonomy of Use

2.1 Independent/User-oriented

Firstly, let us elaborate whether evaluators have independent credibility or stand on the user side in the process from evaluation to decision-making. This may underlie the conflict structure of Weiss-Patton debate in the late 1980s (Alkin 1990). Roughly following both claims, Patton looked for a consensual and rational decision-making by providing users with useful information whereas Weiss insists that decision-making in a pluralistic and contentious community is irrational and political - what evaluators must do in such a context is to generate exact and adequate evaluation information (Smith and Chircop 1989; Shulha and Cousins 1997). In this way, Weiss requires evaluators to have independent credibility while Patton claims that evaluators must do evaluations standing on the user side. This distinction is not just epistemological, but as Patton (1997) proposes ‘utilization-focused evaluation’, is also made in methodological terms. It closely relates to the dichotomy introduced in the next section, where we develop the discussion further.

2.2 Concrete/Abstract

Use can be categorized according to whether the evaluation findings are used in a concrete way (e.g. decisions on the continuation of a program are made following the evaluation findings) or whether the evaluation findings are used in an abstract way (e.g. mid- and long-term policies are planned implicitly referring to the evaluation findings). These two - ‘instrumental use’ in the former and ‘conceptual use’ in the latter - are historically established concepts.

Instrumental use refers to directly citing or documenting information as a basis for action such as decision-making or problem solving that would not have been made otherwise (Rich 1977, p.200; 1978, p.101; Beyer and Trice 1982, p.598; Landry, Amara and Lamari 2001, p.336). Researchers who introduce this term depict the engineering model in which social research will provide empirical and conclusive evidence that helps to solve a policy problem (Gouldner 1956; Crawford and Biderman 1969, pp.235-237). The assumption in this model is that decision-makers have a clearer idea of their goals and a map of acceptable alternatives and evaluators answer specific requests for information and knowledge in a straightforward manner. Slightly different variations of this model have been termed the ‘problem-solving’ (Havelock 1969, 1971; Rein and White 1977), the ‘decision-driven’ (Weiss 1977; Nelson et al. 1987), the ‘decisionistic’ (Floden and Weiner 1978), the ‘discrete functions’ (Robertson and Gandy 1983), and the ‘policy-driven’ (Hanney et al. 2003) models.

Instrumental use does not always occur according to the engineering model. There is another model to understand the instrumental use in policymaking and administration - technocratic model (Robertson and Gandy 1983; Wittrock 1991). This model assigns independent credibility to evaluators, rather than being oriented to decision-makers as in the engineering model. The basic premise is that evaluation is designed and used to promote
technical solutions to political problems for decision-making (Fischer 1990, p.18), and evaluation can provide definitive and direct solutions. From this perspective, policymakers simply follow evaluators, professional recommendations (Foster 1980) and evaluation findings are automatically to be reflected on policy. The correspondences to this model in a wider domain of social science are referred to as the ‘RD&D (research, development, and diffusion)’ (Havelock 1969, 1971), the ‘linear’ (Cherns 1972), the ‘knowledge-driven’ (Weiss 1977), the ‘empiricist’ (Bulmer 1982), and the ‘philosopher prince’ (Coleman 1991) models.

However, the role of evaluation in politico-administrative contexts is many-sided, subtle and more complex than those implied by the term of ‘instrumental use’ (Weiss 1979; Wittrock 1991). The reality is that decisions often take shape gradually, without the formality of agenda, deliberation, and choice (Weiss 1980); this is conceptual use, which refers to influencing thinking about issues in order to gradually and indirectly shape decisions aimed at future policymaking (Rich 1977, p.200; 1978, p.101; Weiss 1980, p.382). Reported studies on the use of social research (Caplan, Morrison and Stambaugh 1975; Patton et al. 1977; Rich 1977; Lindblom and Cohen 1979; Weiss 1980; Whiteman 1985b, 1997) provide much empirical support for the proposition that conceptual use is more prevalent than instrumental use.

Conceptual use is often referred to in the enlightenment model, as contrasted with the engineering model (Crawford and Biderman 1969, pp.240-241; Janowitz 1970, pp.243-259; Bulmer 1982). Noteworthy is that this contrast has a double sense - whether the way of use is instrumental or conceptual, and whether evaluators have independent credibility or stand on the user side (see Table 2). The enlightenment model assumes that evaluation does not so much solve problems as provide an intellectual setting of concepts, propositions, orientations, and empirical generalizations that inform policy. No single evaluation has much effect, but, over time, concepts become accepted (Weiss 1978). This model is also called the ‘milestone’ model in the sense that information trickles like water through porous rock and exerts indirect and accumulative effects on policy change (Thomas 1985, pp.99-100).

The conceptual use for which evaluators are oriented to users is called the classical bureaucratic model (Wittrock 1991). In this model there is a need for legally and administratively trained personnel who are competent in the application of relevant rule systems but who have little or no need for any kind of social-scientific knowledge. Here the objective of evaluation is not to eliminate chance and uncertainty and uncontrollable social processes such as competition and other non-regulated forms of social interaction, but to deal with them and to live with them - according to the modern theory of bureaucracy and planning which is essentially just an extension of microeconomic reasoning.

2.3 Substantive/Symbolic

The distinction between instrumental use and conceptual use illustrated in the above was recognized, despite in a less clear form, by some researchers (Crawford and Biderman 1969) as early as the late 1960s. Since the concept of ‘symbolic use’ emerged from contexts of the politicization of evaluation in the late 1970s, three basic categories such as instrumental, conceptual and symbolic use (however these are termed) were established and have been referred to date in evaluation (utilization) studies (Pelz 1978; Young and Comtois 1979; Leviton and Hughes 1981; Beyer and Trice 1982; Shulha and Cousins 1997; Landry, Amara and Lamari 2001; Alkin and Taut 2003; Rossi, Lipsey and Freeman 2004, p.411). Symbolic use occurs when users use the findings of an evaluation exercise to legitimize and sustain their views and pre-determined positions (Weiss 1984; Pelz 1978; Knorr 1977). In terms of the above independent/user-oriented distinction, this is obviously the latter type of use. Symbolic use seems most likely with a reasonably clear user's commitment to an evaluation exercise, and relatively high (economic, scientific etc) uncertainty about a wide range of impacts of evaluation.
Table 1 Use of Evaluation Findings

<table>
<thead>
<tr>
<th>Concrete use</th>
<th>Substantive use</th>
<th>Symbolic use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract use</td>
<td>Instrumental use</td>
<td><strong>Persuasive model</strong></td>
</tr>
<tr>
<td></td>
<td>Conceptual use</td>
<td><strong>Legitimizing model</strong></td>
</tr>
</tbody>
</table>

Source: Author’s own

Table 2 Models of Instrumental Use and Conceptual Use

<table>
<thead>
<tr>
<th>Instrumental use</th>
<th>Conceptual use</th>
<th>Independent</th>
<th>User-oriented</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Technocratic model</td>
<td>Engineering model</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Enlightenment model</td>
<td>Classical bureaucratic model</td>
<td></td>
</tr>
</tbody>
</table>

Source: Adapted from Wittrock (1991, Table 1)

The reason why the above three categories have not been critically reviewed may be because evaluation studies and evaluation utilization studies have developed diverging from knowledge utilization studies and knowledge policy studies. For this the implication of symbolic use has been rarely reflected by researchers. An attempt at theory construction in which symbolic use is placed as a contrary concept of instrumental/conceptual use was developed chiefly by Whiteman, a policy scientist. This paper reevaluates his achievement by referring to other relevant discussions.

Instrumental use and conceptual use illustrated in the previous section are based on the premise that there is an established division of labor between evaluation practitioners and users in scientific evaluation and the subsequent rational decision-making. This corresponds to the rational-objectivist approach proposed by van der Knaap and others (van der Knaap 1995; Albaek 1998; Dahler-Larsen 2000). However, evaluation is not a pure scientific endeavor but rather on a boundary between science and politics (Weiss 1973). Put it differently, users attribute scientific authority to their own decision and judgment but at the same time they maintain their political, administrative and legal control in the decision and judgment. This is the subjectivist-argumentative approach in evaluation and policymaking (van der Knaap 1995; Albaek 1998; Dahler-Larsen 2000), which describes the reality in which decision has to be made among conflicting values. Comparing to the premise of instrumental/conceptual use that the evaluation and decision-making process is objective and rational, symbolic use premises that the process is subjective and political. Accordingly, as Whiteman and others (Eaton 1962; Whiteman 1982, 1985a) pointed out, more persuasive would be to place symbolic use in another dimension and classify both instrumental use and conceptual use together as **substantive use**. Then the present study defines ‘persuasive model’ to correspond to instrumental use, and ‘legitimizing model’ to correspond to conceptual use.

As indicated in Table 1, substantive use includes instrumental use and conceptual use, and symbolic use includes persuasive model and legitimizing model.

**Persuasive model** occurs when evaluation becomes ammunition in political debate for whichever side finds its conclusions expedient and supportive. In concrete terms, it advocates issues, persuades people to act, neutralizes opponents, convinces waverers, and bolsters supporters (Weiss 1978, p.32; Johnson 1998, p.94). Since the chief function of such evaluation exercises is persuasion, users are perhaps more subject to the tendency to have the findings exploited for propagandistic aims (Merton and Lerner 1951). In this sense it is also labeled as ‘political’ (Boeckmann 1976; Weiss 1979). If users aim to facilitate persuasion, evaluation can be conducted in a way to support
their favorite position rather than to provide credible evidence, and the use can serve the detriment of social betterment. A tendency not to persuade the public based on evidence but to compromise the evidence to facilitate persuasion is called the paradox of persuasive use (Henry 2000).

**Legitimizing model** is supporting or refuting predetermined positions and legitimizing decisions and resolutions based on political considerations such as political ideology, electoral hopes, coalition expediency or personal idiosyncrasies (Weiss 1977, p.15; Vedung 1997, p.275; Rossi, Lipsey and Freeman 2004, p.411). Evaluation findings are used selectively and often distortedly to publicly support a decision that has been taken on different grounds or that simply represents an opinion the decision-maker already holds (Knorr 1977). Unlike persuasive model, which helps users pressure other actors by providing evaluation findings directly, this model refers to a case in which users indirectly and implicitly use evaluation findings to legitimate and protect their own position (cf. Luukkonen-Gronow 1987). In other words, the difference is whether to explicitly persuade others to accept the users’ position or to implicitly guide others to make it acceptable.

### 2.4 Findings/Process

All the aforementioned models of use can be grouped together as **findings use**, in that evaluation is understood as an outcome of the interests, needs, and calculations of particular actors (Patton 1997, pp.63-85). However, the process of the evaluation exercise can be used even if there are no outcomes of the exercise (Patton 1988). This is **process use**, which can be clearly distinguished from findings use because the objects of use are different. The concept emerged from evaluation utilization studies in the late 1980s (Shulha and Cousins 1997), and the taxonomy in which findings use and process use are contrasted becomes common (Smith 1988, pp.10-11; Kirkhart 2000, pp.9-10). Findings use occurs at summative (judgment-oriented) or formative (improvement-oriented) evaluation (Scriven 1972). On the other hand, process use is sometimes associated with evaluation in a constructivist fashion, which has been called ‘fourth generation evaluation’ (Guba and Lincoln 1989), in which evaluation and use are integrated (Finne, Levin and Nilssen 1995). Likewise, Dahler-Larsen (2000) illustrates that process use signals the integration of evaluation into routine regimes of organizational guidance and control as the third approach to evaluation, following the rational-objectivist and the subjectivist-argumentative. It should be noted here, however, that as we discuss below, process use has impacts on the development of personal competence or the substantive improvement of the quality of a project, so it does not always contribute to organizational management.

Patton, who introduced and has strenuously disseminated the term ‘process use’, appears to delimit process use as a substantive one. This may be related to a habit that he has carefully avoided using the term ‘politics’ in his discourse since the Weiss-Patton debate. By contrast, this study applies the term in a broader sense so as to include symbolic implications. Adopting Vedung’s usage (1997, pp.274-276) and applying Whiteman’s substantive-symbolic distinction (1982), as shown in Table 3, this study terms the substantive process use ‘interactive model’, as distinct from the symbolic process use hereafter referred to as ‘tactical model’.

**Interactive model** brings about individual changes in thinking and behavior, and program or organizational changes in the evaluation process (Patton 1997, p.90). It occurs when users as individuals and groups involved in the

<table>
<thead>
<tr>
<th>Process use</th>
<th>Substantive use</th>
<th>Symbolic use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interactive model</td>
<td>Substantive use</td>
<td>Symbolic use</td>
</tr>
<tr>
<td>Tactical model</td>
<td>Substantive use</td>
<td>Symbolic use</td>
</tr>
</tbody>
</table>

Source: Author’s own
Table 4 Eight Models of Evaluation Use

<table>
<thead>
<tr>
<th></th>
<th>Substantive use</th>
<th>Symbolic use</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Independent</td>
<td>User-oriented</td>
</tr>
<tr>
<td>Findings use</td>
<td>Technocratic</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Enlightenment</td>
<td>Classical</td>
</tr>
<tr>
<td></td>
<td>bureaucratic</td>
<td></td>
</tr>
<tr>
<td>Process use</td>
<td>Interactive</td>
<td></td>
</tr>
</tbody>
</table>

Source: Author’s own

exercise learn by interpreting, understanding, and making sense of their experiences from the exercise process itself. The users apply evaluation-informed knowledge in conjunction with further research-based data and other forms of background like common sense, conventional wisdom, intuition, and the users' own first-hand experiences. Users’ participation in the exercise can prompt both the clarification of standard operating procedures and their revision. This interactive model includes roles such as organizational learning, developing networks, extending communication, strengthening the project, and boosting morale (Forss, Rebien and Carlsson 2002; Patton 1997, 1998). Results of a survey on evaluation use suggest the perceived value of process use in that majority of respondents think highly of organizational development and empowerment or self-determination of individuals through internalizing evaluation processes (Preskill and Caracelli 1997).

There is a case where decision-makers use the evaluation process to signal to those concerned that something is being done about the problem. Important here is not the final outcome but the fact that the evaluation exercise is appointed and underway. Such use is called tactical model. For example, it refers to gaining time to postpone or neglect decisions and measures that should be taken (Knorr 1977, p.171; Vedung 1997, p.276). Decision-makers also use evaluation to deflect criticism, to try to avoid responsibility for unpopular policy outcomes, or to enhance the prestige of the agency involved (Weiss 1979, 1984; cf. Merton and Lerner 1951). In this mode, an evaluation exercise is also used to assume the pose of objective, scientific research (Suchman 1967, p.143). In any case, an empirical utilization study shows that tactical model has rarely occurred (Nilsson 1992).

2.5 Summary and Discussion

Individual classifications presented thus far are summarized in Table 4. Categories distinguishing between independent and user-oriented, concrete and abstract, substantive and symbolic, and findings and process are the eight models of use and indicated by boldface. As discussed above, symbolic use, by definition, does not give evaluators independent credibility. Thus all cells corresponding to “symbolic use - independent” are left unused.

As studies on the process use are relatively new, there has been little evidence on whether a distinction between instrumental (concrete) and conceptual (abstract) in the process use is possible, as a recent study (Alkin and Taut 2003) claims. There is concern about the conceptual process use is non-intentional so as to be called ‘evaluation influence’ rather than ‘evaluation use’. It is also suspected that the tactical model is always concrete and never abstract. For these reasons, the present study tentatively assigns concrete use to interactive and tactical models (see Table 4). Future empirical studies will develop the discussion.

Features of this taxonomy distinct from existing studies are as follows: (1) it places symbolic use contrasting to instrumental/conceptual use in theoretical terms; (2) it defines persuasive/legitimizing model to contrast with instrumental/conceptual use; and (3) it introduces the substantive-symbolic distinction in the process use and
conceptualizes interactive model and tactical model. Furthermore, the present study illuminates individual two models each in instrumental use and conceptual use by introducing a distinction between independent credibility and user-oriented evaluation process and use.

Note that there are often cases where the evaluation use cannot be clearly classified into one model. In terms of the process of the use, the independent/user-oriented distinction should be much clearer, assuming that the procedures and institutions for the evaluation and use are different. In terms of instrumental use and conceptual use, these are not exclusive. Both can be applied as complementary at the same time (Marra 2000). The purpose of use can have substantive or symbolic intention, but the real use often includes both intentions (Feldman and March 1981; Pelz 1978). The use of evaluation findings and the use of evaluation process are clearly distinguishable as the objects of use are different. In this fashion, each model of evaluation use does not always take place independently from other models, but often times several models appear at once.

We have discussed some cases not clearly distinguishable between concrete and abstract, or substantive and symbolic use, but this does not mean that individual definitions and distinctions are ambiguous. Aside from remaining issues on the strictness in this conceptual taxonomy, inter-model or concurrent use can be observed in the reality. Nonetheless, there are possibilities that discourses about whether use is substantive or symbolic significantly vary by observer. For example, whilst a researcher judges an evaluation exercise as being politically used, decision-makers dismiss it or intentionally avoid from making such judgment. In brief, not only the use of evaluation can be political, but also discourses on the use can be political (argumentative). Future discussions on the politicization of evaluation should take into account such a reflexive perspective.

Dichotomies like independent/user-oriented, concrete use/abstract use, substantive use/symbolic use, and findings use/process use are not new in each but the taxonomy integrating these dichotomies is novel and original in the present study. As some literature (Weiss 1978, 1979; Vedung 1997; Owen and Rogers 1999, pp.110-113) only enumerates types of use without indicating a clear classification and distinction, this paper aims to arrange them. It is also necessary here to show a clear classification where some earlier taxonomic studies are confusing. For instance, Smith (1988, pp.10-11) proposes the directly-observable/perceptual distinction, which seems similar to the concrete/abstract distinction, but this taxonomy is mixing methodological and epistemological categories and less obvious and persuasive. Similarly, Johnson (1998, p.105) describes that behavioral use “mainly includes instrumental use, but may also include symbolic use, legitimative use, and action oriented process use” and “cognitive use includes cognitive oriented process use, enlightenment and conceptual use, and individual learning”. Each of them includes various complex concepts and is very confusing.

Dichotomies not taken in the present study include immediate/long-term (Smith 1988; Kikhart 2000), intended/unintended (Kirkhart 2000), partial or incremental/holistic (Smith 1988), and active/passive (Westerheijden 1997). The first two are omitted in order to distinguish ‘use’ from ‘influence’ and limit ‘use’ to less long-term and intentional consequence. In addition, there are relationships in which immediate and direct use loosely corresponds to concrete use and long-term and indirect use to abstract use, so it is better not to overlap classifications. The third distinction in the above, referring to the extent of use rather than the way of use, is outside of the scope of this paper. The last distinction is less appropriate in both pragmatics and typology because it is unlikely for external actors (and users themselves) to discern in a rational way whether user’s consciousness is active or passive. Another reason is that the independent/user-oriented distinction performs a similar role to this - users can be more passive in the former and more active in the latter.
3. Conclusion

This paper picks up a relatively minor discipline called evaluation utilization studies and develops the discussion about what is the use of evaluation. Focusing on how evaluation is used, it analyzed the existing literature and made theoretical dichotomies from a comprehensive perspective concerning the process of use (independent/user-oriented, concrete/abstract), the purpose of use (substantive/symbolic) and the object of use (findings/process). These evaluation uses are then classified into eight models - technocratic, engineering, enlightenment, classical bureaucratic, persuasive, legitimizing, interactive and tactical. Practical and academic implications of this study are first that we come to understand the meaning of evaluation in more pluralistic and comprehensive terms. The deeper the understanding of conceptual use and process use scrutinized in the present study, the less myopic the practice of evaluation and decision-making. In addition, the academic articulation of symbolic use of evaluation in the government policy process makes evaluation in public policy more practical. Earlier evaluation studies tend to downplay political aspects by sticking to ‘objective’ or ‘scientific’ theories and methodologies of evaluation. Beyond the three categories - instrumental, conceptual and symbolic - as a still dominant typology in evaluation utilization studies, this paper acquires a balanced perspective on substantive and political implications in the use of evaluation, through Weiss’s discussion on the politicization of evaluation and the revaluation of Whiteman’s taxonomy. By accumulating findings of the use of evaluation in political and social contexts, the constructive relationship between evaluation and use will develop and evaluation studies toward better use of evaluation will flourish.

For a further study, it may be necessary to discuss where (for what), when evaluation is used in a decision-making and policymaking process and examine the relationship between the findings and the above models of use. The present case assumes that the use of public policy evaluation by bureaucrats, but it is required to theoretically and empirically illustrate to what extent this scope can be generalized. In addition to this, it is an important task to analyze why evaluation is used - factors affecting the use (i.e. inputs), contrasting to the present study on the use of evaluation (i.e. outputs). These factors include the nature of the issue, evaluator and user characteristics, process and product properties, evaluator-user interaction, information needs, legal and political context, and information processing style. From such input-output case analyses, we can develop studies on the characteristics of the use of evaluation in Japan and on the evaluation and decision-making process and institution.

Note


References


(Accepted 20th August, 2009)