# WHY ALL THIS EVALUATION? THEORETICAL NOTES AND EMPIRICAL OBSERVATIONS ON THE FUNCTIONS AND GROWTH OF EVALUATION, WITH DENMARK AS AN ILLUSTRATIVE CASE

Erik Albæk
Department of Political Science
Aarhus University

**Abstract:**  This article demonstrates that theoretical debates and developments in the area of organizational analysis can be a fruitful source of inspiration for evaluation utilization research. The article considers two sets of questions. One concerns the societal function of evaluation, the other its historical development. Three sets of well-established lenses, ground by competing views of organizations and organizational behavior, are used to address these concerns. The theories in question see organizations as rational systems, political systems, or cultural systems. The utilization of policy evaluation in Danish political-administrative practice is used as an illustrative empirical case. The article further demonstrates that inspiration from organizational analysis will bring us no closer to a unified understanding of what evaluation utilization is. On the contrary, it shows that debates on and controversies over evaluation utilization belong to the order of things, and that this order entails multiple realities depending on one's theoretical point of departure.

**Résumé:**  Cet article met en relief l'importance des débats théoriques et des développements dans le domaine des analyses organisationnelles en tant que source d'inspiration pour l'utilisation de l'évaluation. Cette importance a été démontrée à l'aide de deux types d'interrogation. La première concerne la fonction sociétale de l'évaluation; la seconde traite de son développement dans le temps. On s'est basé sur trois séries de propositions concurrentes—bien connues et largement discutées—par différentes écoles théoriques, considérant les organisations comme des systèmes rationnels, politiques ou culturels, respectivement. La pratique d'évaluation dans le système politico-administratif danois a servi comme exemple empirique illustratif. L'article démontre que l'inspiration émanant des analyses organisationnelles ne nous fournit pas, pour autant, une compréhension sans

équivoque et unanime de ce que représente l'utilisation de l'éva-
luation. Bien au contraire, les débats et les controverses exis-
tant sur ce sujet relèvent entièrement de l'ordre naturel des
choses, un ordre dont les multiples faces varient en fonction du
choix théorique entamé.

Although the roots of policy evaluation reach further
back, it was in the 1960s that it first took hold, spread, and became
common practice in Western political-administrative systems. Evalu-
ation activities proliferated, and soon evaluation developed into an
industry of its own. However, not all Western countries took up evalu-
ation at the same time. The U.S. came first. Here evaluation was
closely connected to the efforts of President Johnson's Great Society
to alleviate the hardships of the underprivileged and concomitantly
to rationalize public policy-making (Albæk, 1989–90). Only a few
years later, countries like Sweden, Canada, and West Germany fol-
lowed suit and to a large extent imitated the U.S. in their efforts to
introduce evaluation in politico-administrative practice; other coun-
tries, such as Norway, Denmark, France, Switzerland, and the Neth-
erlands, employed evaluation beyond ad hoc studies only well into
the 1980s (Albæk, 1993; Rist, 1990). When evaluation was first es-
tablished in all these countries, it grew at an explosive rate. Why?
What contributed to this fascination with evaluation? And why did
evaluation emerge and become institutionalized at such widely di-
vergent times in various national contexts? These are the questions
addressed in this article.

Much of the international literature on evaluation utilization is of
limited use in answering these kinds of questions. First, it is based
to a large extent on impressionistic observations, anecdotal evidence,
and empirical generalizations rather than a firm theoretical foun-
dation. Second, much utilization research is biased toward finding
"legitimate" evaluation functions, that is, types of evaluation utili-
zation that are congruent with evaluation researchers' own under-
standing of their role as an instrument for increasing "rationality"
in public policy making. Third, in this connection, much utilization
research is prescriptive in its aim.

As a result, evaluation utilization research has generally been theo-
retically underdeveloped, and models of evaluation utilization have
been underspecified. This article contends that one possible and fruit-
ful, but paradoxically often ignored, way to provide a firmer theo-
retical foundation for utilization research is simply to take
inspiration from organization theory in general.

This article considers two sets of questions: one concerns the societal functions of evaluation, the other its historical development. These are big and complicated sets of questions that we shall not pretend to answer satisfactorily here; nor is this the intention. The task instead is to demonstrate how the general debates in the area of organizational analysis are relevant to research on evaluation utilization. To this end we use the theoretical lenses of three major approaches to organizational theory and analysis, which view organizations as either rational systems, political systems, or cultural systems.

Theories help us to see—to bring out form and structure in an otherwise diffuse and blurred reality. But theories differ in their focus and explanatory logic. Therefore the same reality—in this case, the utilization of evaluation and evaluation research—takes on a different appearance, depending on the theoretical perspective of the observer. What is clear in one perspective will be blurred or even invisible from another perspective. No perspective will give an (evenly) clear view of everything. But by using different perspectives to observe the same reality, we can bring more facets of a complex situation to visibility than would be revealed by using just one perspective (Allison, 1971).

This article shows, on the one hand, that theoretical debates and developments in the area of organizational analysis can be a fruitful source of inspiration for the highly interdisciplinary field of evaluation utilization research. On the other hand, it demonstrates that inspiration from organizational analysis will bring us no closer to a unified understanding of what evaluation utilization is (Rist, 1995), let alone what evaluation is. On the contrary, it shows that debates on and controversies over evaluation utilization belong to the order of things, and that this order entails multiple realities, depending on one's theoretical point of departure.

What follows is a reinterpretation of social-scientific perspectives on organizational analysis as applied to the field of evaluation and evaluation research, here defined as the attempt to apply social-scientific theories, methods, and techniques in the systematic mapping and assessment of public policies and programs, their implementation, outputs, and outcomes, in order to affect future decisions (Vedung, 1992, p. 72). The utilization of policy evaluation in Danish political-administrative practice will function as an illustrative empirical case. These illustrations of evaluation utilization will arouse *déjà vu* even in readers with no prior knowledge of the

Danish system of government, who can doubtless cite similar examples of evaluation utilization from their own national contexts. Although the argumentation in this article is based on the characteristics of research-based evaluations, no sharp distinction is made between them and other forms of systematic evaluation practices. The terms *evaluation* and *evaluation research* will be used interchangeably except where it seems appropriate to draw a distinction.

## ORGANIZATIONS AS RATIONAL SYSTEMS

Historically, evaluation research has found its theoretical justification in the classic synoptic-rational organization model (Albæk, 1989–90). In this model of organizational behavior, assessing the consequences of alternatives in order to arrive at optimal decisions constitutes the central feature of all organizations. Preferences are predetermined authoritatively or on the basis of value consensus, allowing the rational decision to result from a process in which thoroughly analyzed means are related to well-understood problems. Optimal solutions are deemed achievable by presupposing that the process is structured by means of formalized division of labor, hierarchy, communication channels, and so on.

Evaluation has traditionally been seen as an instrument to be used in a rational, analytical decision chronology to secure high and efficient goal attainment (Weiss, 1972). In such a chronological sequence of stages, evaluation is introduced either as an *ex ante* assessment of whether given policy instruments can reasonably be expected to have the anticipated effect, or as an *ex post* analysis of the result. This is also the case in the few existing Nordic evaluation textbooks (Møller Pedersen, 1980; Premfors, 1989; Vedung, 1991). From its initial definition primarily as a means of assessing program outcomes or effects, evaluation has over time expanded its methodologies as well as the questions it addresses to encompass practically all decision stages of a policy or program—from conception through implementation through impact through policy reassessment (Albæk, 1993; Rist, 1995).

Because of its universalist understanding of organizations, descriptively as well as prescriptively, the classic synoptic-rational organizational theory provides only limited opportunities for analyzing variations in evaluation utilization as well as its development over time. Instead, contingency theory directs our attention to the diver-

sity of organizational forms. Researchers in this tradition have sought to identify relevant variables that influence organizational form and to specify the conditions under which these factors are likely to produce particular forms. Like the synoptic-rational model, contingency theory is rational in its orientation: organizations have goals that they pursue, and in attempting to attain these goals they seek efficient means. In successful organizations the structure, processes, culture, and leadership style, for example, ensure high goal achievement in the most efficient way (Galbraith, 1973; Lawrence & Lorsch, 1967; Thompson, 1967).

But—and this is the point—not all ways of organizing are equally effective for all organizations. The effective organizational form is contingent upon the situation of the organization. The basic assumptions of contingency theory can be summarized by the often-cited observations advanced by Galbraith (1973): (1) there is no best way to organize, and (2) not all ways of organizing are equally effective (under given circumstances). These should not be misunderstood to mean that all forms of organization are unique. On the contrary, the message in the explanatory parenthetical note is that under certain circumstances, certain forms of organization will be superior to others. This means that according to contingency theory, effective organizations with similar tasks will tend to have similar organizational forms if their environments are also comparable. Among the relevant factors identified by contingency theory, the influence of the environment on organizational structure has received special attention in theoretical debates as well as in empirical analyses.

Contingency theory can be used not only synchronically but also diachronically. The dynamic view of adaptation and change underlying contingency theory implies that comparisons can be made over time. According to the theory, organizations must effectively achieve their goals if they are to survive. Organizational effectiveness depends on the goodness of fit between the internal features of the organization and the series of conditions on which the goal achievement depends. If these conditions change, the organization must respond.

Rational Evaluation Utilization

Contingency theory is sufficiently rational that it is fully compatible with the generally accepted function of evaluation research as expressed in evaluation textbooks, designs, and applications:

evaluation research is a means to increase the effectiveness and efficiency of organizations. But contingency theory also reminds us that evaluation research is an effective and thus rational instrument *when* the situation calls for it.

It follows that the situation does not always call for evaluation, because not all structures, technologies, and processes are equally well suited to ensuring effectiveness and efficiency in the particular situations different organizations face. This is also true for evaluations. Therefore, according to contingency theory there is nothing especially surprising, and even less anything problematic, in the fact that not all organizations engage in wide-ranging, formalized evaluation activities, or that only a limited part of an organization's operations are involved when evaluation does occur.

Seen from the point of view of contingency theory, it would be very strange and in fact undesirable if evaluation were a universal practice even within the circumscribed terrain of the public sector, because public organizations vary considerably. Even apparently homogeneous authorities such as the Danish municipalities differ in size, demography, and locale (urban vs. rural). Some municipalities confront commercial turbulence, demographic pressures, or unemployment. Some municipalities must cope with a declining tax base and fiscal stress, whereas others thrive. The effective form of organization for each municipality cannot be determined in the abstract. Thus it would be mistaken to assume that municipalities that do not implement research-based evaluation of their activities are the least effective. Whether evaluation is appropriate depends on whether evaluation is an effective instrument for an authority given its circumstance.

Put simply, to evaluate means to stop in the middle of or at the end of an activity and assess whether the results correspond to the purpose and effort. In other words, evaluation is a form of feedback. Evaluations are executed both within and outside the organized behavior, and can stretch from a practically unconscious action to a routine check of standard operating procedures to a theoretically and methodologically complicated, research-based assessment of results.

When evaluation research is chosen as a feedback mechanism, there are three things to keep in mind. First, research-based evaluation demands considerable time, energy, and money. Second, evaluation research can generate only some information about a delimited, simplified segment of an entity. Exactly because evaluation research

can involve theoretically and methodologically sophisticated analysis, it can be a foreign and relatively "artificial" instrument compared to the multitude of alternative feedback mechanisms utilized in daily practice. Third, urgent problems can rarely wait for the production and accumulation of decisive scientific evidence or proof of effectiveness.

Decision makers, professionals, and street-level bureaucrats must therefore sometimes rely on more "natural" feedback mechanisms. These include the deeper and multifaceted experience and shrewdness of "ordinary knowledge," reflected in the embedded knowledge that constitutes the standard operating procedures of the organization, professional judgements, sensible rules of thumb, common sense (Lindblom & Cohen, 1979), and "tacit knowledge" (Polanyi, 1967; Schön, 1983). Another informal feedback mechanism is the information that runs through the fine-meshed network of continuous contacts between public authorities and special-interest organizations, which in a Danish context covers all stages of a decision chronology. A final example of an informal information loop is critical debates in scholarly circles, at professional and practitioner conferences, in professional journals, or in the media and via public participation.

Thus evaluation research should not be abstractly characterized as an unconditional benefit, because we must carefully consider in each case whether research-generated knowledge is worthwhile (literally and figuratively) compared to (1) the information value of alternative feedback mechanisms and (2) the costs in terms of time, energy, and money required or actually expended generating alternative forms of information and, in some cases, including new actors and procedures in the existing routine. The precise calculation of value and costs can have very different results even in apparently similar situations, as for example in the utilization of evaluation across the relatively homogeneous Danish municipalities.

Seen through the lens of contingency theory, we would expect the utilization of evaluation research across a variety of agencies to vary not only in degree but also in approach, depending on the function of the individual agency. In other words, different functions create different questions and different means to answer them. That evaluation research has itself acknowledged this is attested by the development of a considerable repertoire of evaluation designs to enable an optimal and very precise coupling of information needs and evaluation design. For example, legislators may have a general interest in knowing whether an approved policy actually has had the intended

effects and is remedying the problems the legislators wanted to deal with. They typically do not care about obtaining detailed information about the organization, management, or implementation, which they see as belonging to the administrative domain, not their own. However, such information is of interest to those responsible for implementing a provision, including institutional management and staff, but also users and user organizations, as it is of interest to sectoral officials in the municipal, county, or central administrations. These administrators might also want to know about the expenses and productivity associated with particular aspects of a given public policy or program, an interest they will often have in common with officials and ministers from agencies with cross-sectoral, coordinating, and economic functions. Typically, this information will be of little importance to professionals or street-level bureaucrats, who are face to face with the users of the public system. Practitioners will to a higher degree demand a continuous, formative evaluation of their daily activities that can help them assess whether they are on the right track or what sections of their operation should be adjusted. It is therefore obvious that the potential users of evaluation research will have some information needs in common, but often at differing levels of detail; at the same time, they will each have information needs specific to them (Albæk, 1993).

The contingency theoretical perspective can also contribute to an understanding of the considerable differences in the use and institutionalization of evaluation in the Western countries. First, the countries that pioneered the incorporation of evaluation research into their common administrative practice—USA, Canada, Germany, and Great Britain (Rist, 1990)—typically operated with public initiatives and budgets on a much larger scale than is the case in Denmark. Size and scale can create a need for a more systematic way of assessing a public policy's substantial, organizational, and economic efficiency. If the policy does not work or fails to address the problem, the waste of public resources and the impact on public support can very rapidly reach astronomical levels. Second, the administrative distance, geographical as well as organizational, is considerably larger in these countries than in Denmark. Administrative distance can also create a need for systematic evaluation. Estimating whether and to what extent an apparently reasonable approach designed in Washington will be effective when implemented in Oakland, California (Pressman & Wildavsky, 1973) is much more difficult than sitting in Copenhagen supervising what goes on in the northern Danish municipality of Hirtshals. Third, Denmark has a long corporatist tradition of alternative feedback mechanisms in

the form of strong, sectoral networks of continuous and institution-alized contacts between affected interest organizations, administrative agencies, and politicians.

Contingency theory can assist us in generating hypotheses not only about synchronic variations in public authorities' use of evaluation research, nationally as well as internationally. This perspective can also help to identify possible explanations of diachronic shifts, and thereby help us to understand why evaluation research is relatively recent yet spreading rapidly within the Danish administrative context.

From the middle of the 1960s, when evaluation research was insti-tutionalized in the U.S.A., until today, when evaluation has become a commonly used instrument in Danish administrative practice, the Danish public sector has developed dramatically. During this pe-riod the Danish welfare state expanded and consolidated. Through the mid-1960s, the Danish public sector's share of GDP was well below the OECD average. By the middle of the following decade, Denmark had rocketed to the top of the list. At the same time, the country was struck by severe economic problems as well as steer-ing, capacity, and legitimation problems. This combination of prob-lems meant that by the late 1970s, Denmark seemed to exhibit all the symptoms, to be in fact a model example, of the long-predicted "crisis of the welfare state."

Around 1980, the question whether the Danish welfare state model had lost its dynamism and effectiveness emerged in the Danish po-litical debate. As in a number of other Western countries, the gov-ernment initiated comprehensive public-sector reforms aimed especially at increasing both efficiency and productivity. Its line of reasoning bore a striking resemblance to the main tenets of contin-gency theory. First, the reform programs were based on the idea that the conditions under which Western welfare states had been estab-lished had changed, and that the public sector's organizational forms had become outdated (March & Olsen, 1989, pp. 95–116; Olsen, 1988, 1991). Second, the reform initiatives were based on the assumption that it is possible to deliberately change public-sector organizational forms in order to reduce the mismatch between public operations and the constraints imposed by citizen demands and the economy.

Thus the increased use of evaluation and evaluation research may be interpreted in light of the reform initiatives of the 1980s and 1990s. The changes in the public sector mentioned above have

increased the apparent rationality of using an instrument such as evaluation research to enhance public-sector effectiveness and efficiency. Several factors pointed to by contingency theory had changed. First, the size and complexity of the public sector had increased unusually quickly. The increase in the number of tasks defined as public responsibility has been explosive, and the public sector continues to reach ever deeper into other spheres of society. The size of the public sector has in itself become a problem, in that it is becoming more and more difficult to manage, administratively as well as economically. When a new policy is accepted and implemented, it is typically through an existing and wide-ranging organizational network, whose exact structure and mode of operation are difficult to assess, design, and keep on track. For the same reasons, economic steering has been hampered. Furthermore, the nature of the tasks the public sector has taken on has caused more problems. Many public tasks, new and traditional, are much more complicated than before. Modern—perhaps post-modern?—society is becoming more complicated, fragmented, turbulent and unpredictable. This turns more and more tasks into so-called wicked problems (Rittel & Webber, 1973), where there is no clear consensus on or unequivocable knowledge of what the problem actually looks like or how to define it, its scope, or its possible cure. This has generally created greater skepticism over whether public actions work according to their intentions, along with a sense that the public sector's traditional feedback mechanisms no longer provide an adequate information base for policy making and administrative practice. These mechanisms need to be supplemented with more systematic, sometimes research-based evaluations of the effects, implementation, productivity, costs, and so on of public policies and programs.

Second, this direction is reinforced by the fact that the tight economic situation means it is no longer acceptable to solve public problems (solely) through increased appropriations. What is lost externally in terms of missed expansion opportunities must be gained internally by improving the public sector's efficiency and productivity. In this case, it may also be sensible to supplement existing, relatively unsystematic feedback with a more systematic and even research-based form of feedback.

Third, the public sector today faces a significantly more educated and critical population than previously. This has created the need for more systematic knowledge about the effects of public operations as well as about citizens' needs, wishes, and demands. Policy makers and implementers need to know how citizens (clients, consum-

ers) perceive the services they receive and how, for example, they prioritize different types of benefits.

Fourth, these societal changes have resulted in increased skepticism toward traditional bureaucratic forms of organization. Instead, politicians and voters seem to prefer decentralized administration practices, capable of finding flexible and user-adapted solutions (Albæk, 1994). Relaxing bureaucratic control simultaneously creates a need for a better and more systematic information base to monitor the productivity, effectiveness, and efficiency of the decentralized administrative units.

The above-mentioned changes in the public sector's contingency factors have developed over a long period. The fact that evaluation research was not introduced to any significant extent in Denmark until the 1980s and 1990s can, however, also be explained within the framework of a contingency theory perspective. No organization can withstand constant changes in its basic structure and mode of functioning. Instead, the existing form will be preserved as long as possible in order to maintain internal consistency and avoid disturbing the existing equilibrium. The organization is often flexible enough to withstand some discrepancy between environmental demands and organizational form, until an organizational "paradigm shift" occurs. This has, of course, also been the case for the introduction of evaluation research as supplement to the traditional Danish administrative feedback mechanisms. Even though for a number of years Danish policy makers and public administrators were acquainted with the use of evaluation research in other countries, and even though its introduction had for a long time been an obvious option within Danish administrative practice, old feedback mechanisms prevailed. The older practices were familiar and comfortable.

Furthermore, the use of evaluation research can be influenced by other changes or paradigm shifts in the public sector's mode of functioning. For example, the program approach was introduced relatively late in Danish political-administrative practice, and it appears to have influenced evaluation research. Whereas earlier, public reforms were usually carried out as basic, structural policy revisions, (e.g., the Danish school reforms), nowadays public policy commonly takes the shape of a program. In principle, this typically means a cohesive system of general and concrete initiatives that are more delimited in terms of time and target groups than traditional Danish policy revision. For example, the increased use of public-sector experimentation in Denmark typically has the characteristics of a

program. And precisely the more limited target of the program model improves chances of measuring the effects of the program as a whole as well as those of its individual subelements. Thus it is no coincidence that DANIDA, the Danish agency for international development assistance, having operated in an international environment where the program model has been known for years, not only was among the first agencies in Denmark to employ the program approach, but also was the first agency to systematically evaluate its program activities.

An important condition for introducing a new political-administrative technology is that the technology actually exist in a given organization's environment. In Denmark, evaluation technology became accessible relatively late. Two of the social sciences that have been central to the development of evaluation research, sociology and political/administrative science, did not become independent university studies in Denmark until the end of the 1950s. For many years, the supply of graduates from these programs was very modest. Therefore public authorities had no personnel with a relevant, professional background to demand evaluation research, and few adequately educated graduates to head evaluation research projects. The number of social science graduates—not to mention the number of people with PhDs—is still below, for example, U.S. levels, and Danish universities did not start teaching evaluation and evaluation methods on a significant scale until the end of the 1980s (Albæk & Winter, 1990).

ORGANIZATIONS AS POLITICAL SYSTEMS

When we choose a particular analytical lens or approach we bring some things into focus, but we also leave other aspects outside our field of vision (Allison, 1971). Rational perspectives capture the more stable and formal aspects of human behavior. By making rationality the overriding value and not a partial logic that functions only under certain conditions, the rational perspective does not address such phenomena as power and conflict, except as indicators of failure. A rational approach tends to ignore or discount the complex, nonrational side of human behavior, which constantly invades and disturbs even the most well-designed organizational structures and processes. Therefore, an additional perspective that focuses on those aspects of organizations the rational perspective ignores will add to our understanding of both organizations and the utilization of evaluation research. One such additional approach views organizations

as political systems and therefore focuses on such phenomena as conflict, power, strategy, tactics, bargaining, and coalition-formation (e.g., Cyert & March, 1963).

Both the rational and the political perspectives on organizational behavior build upon the fundamental assumption that organizational behavior is rational, that is, based on goal-oriented choices that couple problems with well-calculated solutions. In the rational perspective, organizations are steered by clear and consistent objectives and policies that are set by the organization's top personnel, who have the formal authority to do so. The organizational form ensures the optimal coupling of goals and thoroughly analyzed and assessed alternatives. In the political perspective, on the other hand, the rational calculation is localized to the individual level. Organizations are seen as political arenas where a complex multiplicity of individual and group interests come into play. Organizational goals, objectives, and policies do not derive from the top of the system, but rather emerge from the continuous processes, negotiations, and interactions among the members of the organization. Policy determination therefore is a product of a process in which a series of actors with different interests, priorities, and demands bargain, form coalitions, and make use of their (often unevenly distributed) resources to influence the result of the process—the policy decision—so that it conforms as closely as possible to their individual interests, views, and needs. In other words, the connection between problems and solutions in this perspective emerges from a process of negotiations, where the organizational form ensures that bargaining produces a choice among problems and possible solutions. Consequently, the organizational form is not necessarily a neutral factor or level playing field for conflicting interests. Rather, the organizational form constitutes an institutional arrangement that in itself reinforces the political nature of decision making and influences the bargaining possibilities of various actors. Furthermore, the choice of organizational form can become an instrument for shifting the balance of power among the interests engaged in bargaining (Moe, 1989, 1990; North, 1990; Pierson, 1993; Shepsle, 1989).

The dominant understanding of science has been that research, properly conducted, can produce objective knowledge. In other words, scientific research addresses that which is, not that which should be. Both the classic ideal of the scientific process and the synoptic-rational decision model prescribe procedures—and only procedures. This means particular steps that have to be followed regardless of the subject under investigation or the objective of the inquiry if the

findings are to be received and accepted as "scientific" or "rational." This symmetry between the logical structures of both administrative decision-making and research processes forms the basis for the general view that social research can contribute to the social engineering work implicit in the public decision-making process, because social scientists as well as political-administrative decision makers have clearly defined and separate roles that do not compromise their respective professional norms and self-conceptions (Albæk, 1989–90).

It is this ideal coupling between social research and the political-administrative decision-making process that the political approach to organizational analysis challenges. First, we must question the assumption that decisions conform to the synoptic, rational ideal. This does not necessarily mean abandoning the premise that social research should be value free and objective. The political perspective points out that the very choice to utilize social research is political. Second, the political perspective may also call into question the impartiality of social research. This observation draws on a long and growing tradition within the fields of philosophy of science and sociology of knowledge. These critiques have in part been directed at the selection of subjects, which tends to favor the more conservative (conserving) side of a conflict. This was the message delivered during the first phase of the student revolt in the 1960s, and mainstream social research later responded by focusing on previously overlooked or understudied social groups, their living conditions and history (e.g., workers and women). But this critique has also been directed at even the formulation and choice of concepts and terminology, theories, and methods, which tend to reinforce existing power relations. This perspective is generally shared by research based on Marxist and critical theory (Bernstein, 1976; Bobrow & Dryzek, 1987; Fay, 1975; Fischer, 1980; Torgerson, 1986). In this connection, researchers have developed research strategies that allow them to take a political stand in their work, especially in favor of society's underprivileged groups, as in action research (Baklien, 1993; McTaggart, 1991; Whyte, 1991).

Political evaluation utilization

Viewed through the lens of the political perspective, social actors in a given decision arena will use research and research data to maximize their individual goal realization. The use of evaluation research increases in part because of the widespread conception of science as a neutral and objective enterprise. In other words, the research com-

munity delivers facts, not politically influenced opinions. But precisely the widely held belief in the impartiality of science yields its political utility. In the political game it is an asset to possess "knowledge" produced by a neutral institution outside of the political arena to support one's arguments. Thus the function of evaluation research is not to provide the actors with new insights, but to support existing views. In short, evaluation provides political ammunition (Albæk, 1989–90; for Nordic analyses, see Eliason, 1988; Nilsson, 1992; Premfors, 1982).

There are two principal ways in which evaluation research is utilized in this context. In the first case, a consumer of evaluation research merely interprets existing data so as to confirm his or her own biases. Of course, this is often done for a good reason, because solutions to societal problems can never be based solely on pure, cognitive knowledge. Societal problems are too complex to be judged on the basis of precise assessments or predictions of effects. In addition, evaluation research by itself cannot determine whether to view the proverbial glass as half full or half empty. In other words, research that conforms to the dictates of scientific objectivity generally remains neutral on the normative interpretation of its findings. Thus, solutions to social problems cannot be "discovered," they must be "willed," that is, they are chosen (Lindblom, 1990).

As an example, a Danish anti-unemployment program intended to promote the channeling of unemployed college and university graduates into nontraditional forms of employment was evaluated in the early 1980s. The evaluation showed that among program participants who completed a training period with close organizational links to the company providing the training, roughly three-quarters would find subsequent employment, most of them in nontraditional employment situations (Albæk, Krogh, Madsen, & Thomsen, 1984). The publication of the evaluation findings received intense media attention; both supporters and adversaries of the reform agreed that the program was a success. But consensus about the results of the evaluation did not produce political agreement about whether the program should be continued. The Social Democratic Party had launched the program while still in power and favored an expansion of the program in order to get more people into nontraditional forms of employment in an otherwise rigid labor market for college and university graduates. The Conservative-led government coalition had little use for an anti-unemployment program invented by its Social Democratic predecessor, and favored terminating the program because it had already demonstrated the intended wrecking ball effect:

that it was in fact possible for unemployed graduates to find employment outside their traditional fields, especially in private businesses. Accordingly, the program was terminated. Presumably the private sector, market forces, and individual initiative would generate the necessary conditions for well-educated individuals to switch career paths.

The other main form of the political use of evaluation research consists of directly purchasing evaluation results that support one's views. There are several variations of this form of sponsored evaluation, some of which do not necessarily challenge the integrity of evaluation research as (in principle) a neutral institution. For instance, as part of its "Modernization Program for the Public Sector," the Conservative-led government launched an experimental program in 1985 that permitted "deregulation" of selected municipalities under specified conditions (i.e., some local governments were granted dispensation from central government control, for example in the financing and organization of the delivery of social services). This "Free Municipality Experiment Program" was not subject to a general evaluation, but the government selected some of its favorite successful projects for evaluation. Documentation of positive results in an evaluation report would bolster government arguments in favor of legislative amendments to convert the project's provisional status into a permanent feature of the Danish administrative system. In one instance, a ministry collected data from free municipalities *after* the ministry had proposed a legislative amendment and at a time when the Parliament was so far along in its legislative work that the evaluation could not be used as a formal rational input into the decision-making process. The ministry sought to publicize the achievements of the free municipalities that had experimented with regulations like those in the new law. Perhaps the ministry wanted to use the evaluation to provide information about the proposed reforms; a less charitable observer might call the ministry's evaluation initiative propaganda (Albæk, 1994).

Perhaps the best-known Danish examples of legitimizing evaluation in recent years come from the increasingly commonplace practice among public agencies of contracting with external management consultants to assess cooperation problems in the organization. For example, an agency's upper management might commission an external evaluation as ammunition in negotiations over removing a staff member they have already identified as deficient. The top brass probably does not need to attempt to influence the evaluator's work because the problem as well as its solution are fairly obvious. The

purpose of evaluation is not to gain new insights, but rather to minimize the political costs—to management as well as trade unions—of an otherwise unpopular decision by doing what everyone, or almost everyone, agrees is the only reasonable solution (Mitnick, 1993).

However, sometimes the neutrality of agency-sponsored evaluation research is compromised. First, the selection of evaluators is to some extent political and sometimes even the object of great political controversy among actors with a vested interest in the policy or program to be evaluated. Second, determination of the scope, design, and methodology of the evaluation is also sometimes the result of blatantly political negotiations. Third, it does happen that evaluators modify their evaluation results to conform to the perceptions, positions, and demands of sponsors or other vested interests. In private conversations, evaluators often offer examples of how they have revised their report after the sponsor has read it and commented on it. This is not to say that the evaluation results are directly manipulated; but conclusions can be reformulated vaguely and less critically.

The structure of the market for evaluation research in itself makes such behavior understandable, even politically and economically rational. Suppliers of evaluation research—i.e., evaluation researchers—enter into an exchange relationship with the sponsors: the latter pays for an "objective" and "neutral" legitimation of its views, the former delivers the goods to earn a living. Obviously, we are dealing with a phenomenon that is hard to prove and that few involved in such transactions would substantiate, but a strong feeling that the problem is widespread has led to a debate about the appropriate organization of the evaluation market in Denmark, and the risks the status quo entails for the quality and integrity of evaluation research (Hansen, 1992; Notkin, 1992).

Finally, it should be mentioned that to a certain degree evaluation research as a field of professional endeavor has ensured that this form of applied research is part of the political game. For example, innovations in evaluation designs have brought greater political engagement for those who use them. Examples are "utilization-focused evaluation" (Patton, 1978), "participatory evaluation" (Cousins & Earl, 1995), and "empowerment evaluation" (Fetterman, 1996), which all resemble action (evaluation) research in its goal of actively assisting in the creation of action possibilities for various policy actors (Baklien, 1993; King, 1995; McTaggart, 1991; Whyte, 1991). Similarly, "stakeholder evaluation" (Bryk, 1983) as well as "participatory, post-positivist policy analysis" (Torgerson, 1986) take into

account in their research designs that a given public policy or program will have a number of interests with different and partly conflicting goals attached to them.

The political perspective on organizations also helps us to account for variation in the demand for evaluation designs, methodologies, and techniques. To paraphrase an American saying, Where you stand with regard to what constitutes useful, legitimizing knowledge depends on where you sit in an organization. Such standpoints will be influenced by substantive as well as by institutional goals (Heffron, 1989). Sectoral ministries as well as other agencies responsible for the implementation of a given policy or program may not only—if at all—have an interest in documenting the effectiveness and implementation of the program in order to make program adjustments; they may also have an interest in documentation that can be used to minimize the chance of a downsizing of the authority's activities, and instead maximize its chance of an upgrading. Therefore their interest will concentrate on, for example, output and outcome evaluations, which can demonstrate the utility of a program for the intended target group, preferably supplemented by consumer surveys and thick descriptions documenting citizen satisfaction with the program or dissatisfaction with its limited scope. Thus evaluation helps document that what the organization does, it does well, and if more resources were made available the organization could do even more even better.

In contrast, such authorities' interest in productivity, efficiency, and cost-benefit analyses may be limited, whereas agencies with cross-sectoral, coordinating, and especially monitoring and economic functions typically demand the information provided by these types of analysis. Because special-interest organizations are rarely affiliated with these authorities, and the authorities therefore are precluded from utilizing special-interest organizations as a resource in negotiations with other actors in the decision-making arena, public officials might want to have alternative sources to bolster their arguments with hard facts in the form of quantified evaluation results.

The political perspective can be used not only to analyze some of the political functions of evaluation research in the policy-making process. It can also be used to explain variations across countries as well as temporal shifts in the use of evaluation research. If evaluation research is more common in some countries than in others, it may be because of differences in the level of conflict in the policy-

making arena: the greater the scope and intensity of conflict, the greater will be the use of evaluation research.

This could be one explanation of the obvious difference in the use of evaluation research in the U.S. and Denmark. Economically, politically, ethnically, and culturally, American society is extremely heterogeneous; Danish society is far more homogeneous. The development of the American welfare state was marked by a period of intense activity during the War on Poverty and the Great Society programs, roughly the period in which evaluation research was incorporated into American politics and administration. Welfare programs have historically encountered considerable political opposition, whereas the universalist Danish welfare state enjoyed wide political support in all social groups; the American political-administrative system with its separation of powers, checks and balances, federalism, and weak party discipline enables and provokes conflicts more than the Danish political-administrative system. Thus, the number of actors demanding evaluation to control or criticize government performance has been much smaller in Denmark than in the U.S., and Danish governments have only to a limited extent needed evaluations to legitimize policy programs (Albæk & Winter, 1990).

The significant increase in the demand for evaluations in the Danish public sector during the last decade can be attributed to the increased level of conflict in the Danish political-administrative system, for resources declined at the same time as citizen demand for expanded welfare benefits and services increased. During periods of prosperity, abundant economic means, and popular support for public spending, organizations have sufficient resources to meet diverging goals simultaneously, sequentially (Cyert & March, 1963), or incrementally (Lindblom, 1959). But when the economy tightens, this approach is barred (Mouritzen, 1991; Schick, 1983); the level of conflict grows, and difficult choices must be made.

Under conditions of economic scarcity, evaluations of public-sector activities can be important weapons in their own right in political arguments both for and against boosting productivity, cutting costs, and resetting priorities. Second, the notion that it is not only necessary but also possible to deliberately redesign public organizations in order to reduce sub-optimizing incentives for political-administrative actors has become widespread. Thus measures to introduce or expand decentralization, management-by-objectives, user influence, or quasi-markets have become popular. Loosening bureaucratic control creates the need for monitoring functions and information systems if such public-sector restructurings are to achieve their intended

goals. Third, established institutions harbor conservative tendencies. Any kind of change disturbs the status quo, disrupting work habits, routines, and the existing power balance and provoking resistance. An evaluation can sometimes help overcome conservatism and resistance just as experimental programs can, and in fact evaluation and experimentation have accompanied each other in Danish administrative practices throughout the 1980s and 1990s. Public decision-makers are conscious of the political functions of experimental programs and evaluation research and willingly report on this subject when asked directly (Adamsen, Fisker, & Jorgensen, 1990; Albæk, forthcoming).

## ORGANIZATIONS AS CULTURAL SYSTEMS

A cultural perspective operates with an image of organizations significantly different from that of both the rational and the political views, in which utility-optimizing actors are located in a relatively predictable, stable, and controllable universe. To cultural theorists, social actors live in a decidedly uncertain, chaotic, and anarchistic world. Social actors are less concerned with producing substantial results than with discovering a sense of direction and discerning meaningful patterns and purpose to their activities. Therefore the cultural perspective does not focus on making (rational) choices, but on the production of norms and meaning. When people encounter uncertainty and ambiguity, they generate collectively recognized symbols and symbolic actions—including myths, rituals, metaphors, and stories—that help them reduce ambiguity and confusion, increase predictability, and develop guidelines that simplify and explain their perceptions of reality (Edelman, 1964, 1971; March & Olsen, 1989, pp. 39–52; Ott, 1989).

Myths are like the proverbial forest obscured by the trees: we are surrounded by their details but do not see them in their entirety because this would require a perspective external to them. This is especially true for the dominant and highly institutionalized belief that we neither live in nor are controlled by myths (Meyer & Rowan, 1992a; Meyer, Scott, & Deal, 1994). We see ourselves as secularized individuals who make independent choices and adjust our behavior on the basis of a rational calculation of consequences. We seek greater control over our lives by interpreting our world—especially our organizational world—and our actions in rational terms.

From a rational perspective, an organization consists of a number of subunits comprised of individuals and the connections between and

among these subunits and individuals. The elements comprising the organization can be instrumentally manipulated so that an organizational design emerges, determined by the organization's purpose, technology, and surroundings. The productivity and efficiency of an organization can be intentionally improved by changing its forms, that is, its structure, processes, and ideology; and that is also the explicit purpose behind the initiation of organizational changes (Brunsson & Olsen, 1993, pp. 1–3). But perhaps the notion that organizations can be designed deliberately is in itself a myth. In other words, from the cultural perspective, contingency theory is a mythology that has been incorporated into the reorganizing ritual of organizations. Reorganizations may serve noninstrumental functions, for example, to signal adherence to prevailing values and myths in society through the particular organizational form. In particular, organizations characterized by multiple and unstable goals, barely specifiable technology, inadequate knowledge of ends-means relations, and inadequate means to measure vaguely defined criteria for effectiveness and efficiency will need to legitimize themselves through isomorphic adaptation to society's norms for what such organizations are supposed to look like. According to the cultural perspective, organizations are judged not only—and not even primarily—by what they do, but by their appearance. Their environments are not always interested in the products and services created by an organization. Organizations are also judged by their structures, processes, and ideologies and whether these are perceived by important societal groups as rational, efficient, sensible, fair, natural, and modern (Brunsson & Olsen, 1993; Meyer & Rowan, 1992a, 1992b; Weiler, 1983).

If an organization or its environment changes, the organizational form is expected to change. As myths about the correct, rational organization tend to follow fashionable trends (Meyer & Rowan, 1992a; Røvik, 1992), organizations will constantly—almost routinely (Brunsson & Olsen, 1993)—need to change their appearance to ensure they meet expectations.

However, the coupling of the adapted organizational form to the normal activities of the organization is often so loose as to be almost nonexistent. Consequently, reorganizations can be regarded as the result of widespread "organizational hypocrisy" (Brunsson, 1989), that is, adaption to prevailing images of the right organizational form is a ceremonial facade whose real function is to protect the organization from the demands and expectations of its surroundings while in fact the organization goes about its business as usual (Meyer & Scott, 1992).

The noninstrumental, symbolic functions of organizational forms are directed not only at the outside world, but also at its own members. Structure and structural changes can give the organization's members meaning and faith in the rational legitimacy of their workplace. Likewise, organization processes are not necessarily result oriented. Many meetings end without any decisions made or problems solved; instead, meetings can be seen as occasions for the organization's members to express themselves, vent frustrations, or negotiate new, mutual understandings and beliefs (March & Olsen, 1976).

Cultural Perspectives on Evaluation Utilization

Evaluation is generally viewed as a rational activity intended to facilitate the conscious improvement of effectiveness and efficiency. Evaluation activities consume time, energy, and economic resources; they often result in huge reports full of information and sophisticated number-crunching; and they are often presented at well-staged meetings with all the theatrical effects worthy of a religious ceremony, occasionally with the press and other interested parties in attendance.

Still, we continue to hear an almost monotonous lamentation—especially from evaluators—that such reports are immediately archived only to collect dust in the sponsors' files. No one is influenced by the research results, nobody remembers them. Anecdotal and impressionistic evidence based on the reported experiences of evaluation researchers is confirmed by the results of empirical research on research utilization, which has generally had a hard time documenting the kind of instrumental research utilization that the rational perspective—and thus the generally accepted objective of evaluation research—aims for (Albæk, 1989–90).

We are apparently facing a paradox: more and more resources are invested in the production of evaluation research intended to immediately and directly improve the information basis of decision makers; yet this effort conflicts with our empirical knowledge of how evaluation research is actually used in the policy-making process. Why, then, do policy makers keep wasting more and more resources by sponsoring such a futile activity?

The answer to this question is that perhaps we are looking for the wrong kind of evaluation utilization. Seen from a cultural perspective, the function of evaluation research is not primarily—and per-

haps not at all—to assist in decision making. Instead, its function is to create the image of a serious, responsible, and sensibly managed organization. This function can be internally directed, at the organization's members, or externally, at its environment.

An organization's members have a need for myths, rituals, ceremonies, and symbols that allow them to interpret their own behavior as rational. An evaluation can serve this purpose, exactly because it appears to be a rational activity. In other words, an evaluation possesses the essential qualities to function as a rational ritual. If we know that our workplace carries out evaluations, if we participate in them, if we are present when evaluation results are presented, if we participate in discussions on the reports, then we may be able to convince ourselves and our colleagues that what we and our workplace do is rational, responsible, serious, and the like. Evaluations can also—in the same way as, for example, planning (March, 1972, p. 427)—give us occasion to interpret previous incidents as rational when they notoriously were not (March & Olsen, 1976, pp. 338–350). In other words, they enable us to rationalize our actions after the fact. Furthermore, evaluations can help to reduce complex, societal problems to a choice between relatively well-defined alternatives. They can play a role in making a chaotic world seem more orderly and easy to handle and giving an organization's members the impression that they have a firm grasp on things.

These functions turn the focus of attention not only inward toward the organization's own members, but also outward toward its environment. By carrying out evaluations, public agencies can signal that they seriously intend to pursue stated goals and that they have things under control. First, affected social groups may be reassured by the launching of an evaluation, as it may indicate to them that their problems and concerns are taken seriously. When the Danish State Railways has several dramatic accidents within a short period, the implementation of a safety evaluation, whether it comes up with useful results or not, can help reduce the public's fears and channel media attention to other areas.

Second, evaluations can help improve the image of decision makers so that they appear as strong, efficient, and serious leaders who are in control of a rationally organized and well-run organization. We would be naive to conclude, when the Danish Ministry of Education called upon "leading" foreign experts to evaluate the quality of the Danish school system based on only three days of visits (Egelund & Larsen, 1989), that the ministry sought to acquire new, ground-

breaking, and substantial knowledge about the Danish educational system or even some knowledge they didn't already have. Instead, by publishing the evaluation under the glare of media attention, they showed the public that they take their task seriously and that they use the most modern organization technology and the most advanced experts to supply their decision bases.

In general, it becomes necessary for organizations to perform evaluation rituals when evaluation is an element in societal myths about how rational, efficient, and responsible organizations look. Thus, today most applications for public-sector experimental projects specify that the project subsequently be evaluated, not because the evaluation will have consequences for the project itself or other future activities, but simply because it is now fashionable to evaluate all projects. Similarly, Danish universities recently introduced instruction evaluation as a standard procedure; however, there is no indication that evaluations have any influence on future decisions concerning the assignment of teachers at different instruction levels, on curriculum quantity and quality, or on teaching methods. Perhaps that is not the purpose of instruction evaluation after all. The purpose might alternatively be to signal responsiveness to students' persistent criticism concerning the low quality of instruction, and to an education minister and his ministry who see it as their job to "modernize" the Danish educational system by means of a series of rational and productivity-boosting initiatives, one of which is evaluation. Other cases show that organizations perform evaluations without prior thoughts on the purpose of the evaluations. They evaluate because everybody else evaluates.

As was the case with the rational and the political organization perspective, it is possible not only to analyze the societal functions of evaluation research through a cultural lens, but also to explain changes over time. Ironically, Danish governments throughout the 1980s pursued initiatives to reform the public sector under the auspices of a "modernization program," implicitly indicating the developmental dynamic revealed by the cultural perspective. "Modernization" is closely associated with the rational myth; organizational change is legitimized by presenting it as "modern," which also implies that the reforms reflect current trends. Organizations want to keep up with the current trends, just as everyone to some extent follows clothing fashions. That is why organizational consultants have been characterized as "traveling salesmen in organization fashion" (Røvik, 1992).

In a Nordic context, the concepts of "workplace democracy," "codetermination," and "planning" were ubiquitous in the 1970s and influenced myths about what the appropriate and rational organization looked like. Any self-respecting organization, private or public, introduced worker participation initiatives, and one public agency after another produced detailed planning documents and established special planning departments. Similarly, the word "planning" turned up, almost without exception, in the names of social science institutes at the newly founded universities in Denmark in the 1970s. Twenty years later, the institutes still have the same names, but today there is little if any separate instruction in planning at the institutions of higher education. Danish authorities do not make a show of planning and workplace democracy initiatives when they want to signal that they are "modern." Instead, the private sector's organization principles are in vogue (Olsen, 1988, p. 7). According to the cultural perspective, it is no surprise that so many authorities have discovered evaluation research, because it corresponds to the organizational thinking that was employed during the modernization waves in the 1980s and 1990s. Or, as a Swede characterized the fashion shift between the 1970s and the 1980s, "from planning to performance" (Nilsson, 1993).

Furthermore, the cultural perspective enables us to explain variations in the use of evaluation research. Organization fashions are like all other fashions: they are not necessarily universal. In spite of American cultural imperialism, the Scandinavian countries do not copy everything from "over there." When, for example, DANIDA incorporated policy programs and evaluation into the departments's administrative practice, it may have been because the organization operated in an international environment that was dominated by American administrative culture, and where consequently such phenomena as programs and evaluation were common.

However, the cultural perspective can also provide other and supplemental reasons for the increased use of evaluation research. We have previously discussed how the explosive growth of the public sector during the welfare state's expansion period led to a significant increase in the number of so-called wicked problems with rather muddled, indefinable, and uncertain solution alternatives. According to the cultural perspective, perhaps precisely in such highly ambiguous situations organizations and their environments require symbolic production to reduce uncertainty and construct an interpretation of the situation in order to act.

The fact that a phenomenon is defined as a public matter signals an expectation that the public authorities are capable of finding a solution to that problem. This holds true even when possible solutions may not be accessible to Danish decision makers, as could be argued is the case in connection with unemployment and drug abuse. But because popular expectations are directed at the Danish public authorities, they must prove to the population, at the very least through symbolic action, that they take the matter seriously, that they are working on it, and that relief is on the way. Thus evaluation and evaluation research may increase the credibility and acceptability of public policies among the population.

All in all, it may be asked whether this flickering, turbulent, and fragmented postmodern time we live in makes it more necessary than earlier for organizations to justify their existence. To be an authority does not have the ring of authority any longer. Public actors—whether a teacher in front of the blackboard, a politician making a statement to the press, or an environmental inspector visiting a company—no longer enjoy the confidence afforded the big, once awe-inspiring institutions behind them. Instead they must constantly justify why they are there, which "introduces an element of show instead of substance" (Syberg, 1994). Thus public policies suffer from a legitimation deficit which may be compensated for by adding legitimacy to the policy process, for instance through the utilization of research-based knowledge or evaluation (Weiler, 1983).

CONCLUSION

Three different perspectives from organizational theory have been presented as the basis for an interpretation of the societal functions of evaluation as well as of its historical development in Denmark: the rational, political, and cultural perspectives. We observed that the functions and historical development of evaluation varied depending on the theoretical perspective.

The function of using a theoretical perspective in science is almost identical to that of the production of symbols and meaning in the cultural perspective: theories impose structure in an otherwise chaotic, random, and shapeless world. The purpose of one particular perspective is not to explain everything from that particular view, but to suggest one possible interpretation. In science, it is usually an advantage to isolate a given perspective and take the interpretation of reality as far as that perspective will allow. This allows us to assess the

strengths and limitations of a particular perspective. But it also means that there are no clear answers to the overall question of this article: Why all this evaluation business? The answers will depend on the theoretical perspective adopted in analyzing the question.

The interpretations presented suggest two additional comments. First, the three interpretations do not seem to be mutually exclusive, even though we occasionally have tried to push their internal logics as far from each other as possible. When, for example, organizations adapt their organizational form to meet the expectations of their environments, this can be interpreted from both a cultural and a political perspective. Adapting their organizational form is how organizations adhere to prevailing norms and values, but it is also how they survive and protect their core activities (Meyer & Rowan, 1992a; Scott, 1977). Similarly, substitution of one organizational form for another can be interpreted not only as the institutionalization of a new definition of reality, but as a new winning definition of reality, where actors with different interests in and resources for strategic influence over the process of opinion shaping have competed against each other for organizational control (Brint & Karabel, 1991; DiMaggio, 1988). When politically formulated arguments are used in a political conflict, that does not mean that all arguments are (equally) marketable; Western democratic culture supports the norm that not only does every citizen have the democratic-political right to voice his or her opinions, but also that these opinions must be justified, that is, they must be subjected to rational scrutiny (Albæk, 1995; Lindblom, 1990; Majone, 1989; Stone, 1988; Torgerson, 1986). Today it seems more necessary than ever to justify one's existence through symbolic reproduction. That probably has less to do with general fashion trends than with the fact that the public sector's contingency factors have changed so that the legitimizing production of meaning becomes more and more necessary just to survive (March & Olsen, 1989; Weiler, 1983). These examples indicate that our knowledge of the societal functions of evaluation is to be found inside existing organizational theory perspectives, as well as at their intersection (Nørgaard, 1996; Winter, 1991).

Second, the three perspectives discussed in this article do not capture all societal functions of evaluation and evaluation research, and perhaps only a few of them. For example, empirical research on research utilization has documented that the so-called enlightenment (Weiss, 1977) and conceptual (Peltz, 1978) functions are among the most widespread forms of research utilization—also in Scandinavia (Lampinen, 1992; Naustdalslid & Reitan, 1992; Nilsson, 1992;

Premfors, 1982). Ideas, concepts, and generalizations born of research indirectly and diffusely enter and become part of that huge reservoir of knowledge and experience that helps practical actors orient themselves in their daily activities and incrementally adjust and change these. Conceptual utilization does not easily fit into any of the three above-mentioned perspectives. This suggests that we must supplement our theoretical and empirical analyses of evaluation utilization with other perspectives from organizational theory.

## ACKNOWLEDGEMENTS

## REFERENCES

Adamsen, L., Fisker, J., & Jorgensen, K. (1990). *Forsøgsstrategi: Samfundsmæssige konsekvenser og fremtids perspektiver*. København: AKF Forlaget.

Albæk, E. (1989–90). Policy evaluation: Design and utilization. *Knowledge in Society, 2*(2), 6–19.

Albæk, E. (1993). Evalueringsforskning i Norden. *Politica, 25*(1), 6–26.

Albæk, E. (1994). The Danish case: Rational or political change? In H. Baldersheim & K. Ståhlberg (Eds.), *Towards the self-regulating municipality: Free communes and administrative modernization in Scandinavia* (pp. 41–68). Aldershot: Dartmouth.

Albæk, E. (1995). Between knowledge and power: Utilization of social science in public policy making, *Policy Sciences, 28*, 79–100.

Albæk, E. (forthcoming). Holdning til omstilling og forsøg blandt de nordiske kommunaldirektører. In L. Rose (ed.), *Kommuner og kommunale ledere i Norden*.

Albæk, E., Krogh, I., Mygind Madsen, A., & Risbjerg Thomsen, S. (1984). *Utraditionel beskæftigelse til langvarigt uddannede*. Aarhus: Aarhus Universitet, Institut for Statskundskab.

Albæk, E., & Winter, S. (1990). Evaluation in Denmark: The state of the art. In R. C. Rist (ed.), *Program evaluation and the management of*

*government: Patterns and prospects across eight nations* (pp. 95–104). New Brunswick, NJ: Transaction Publishers.

Allison, G.T. (1971). *The essence of decision: Explaining the Cuban missile crisis*. Boston: Little, Brown.

Baklien, B. (1993). Evalueringsforskning i Norge. *Politica, 25*(1), 56–63.

Bernstein, R.J. (1976). *The restructuring of social and political theory*. Philadelphia: University of Pennsylvania Press.

Bobrow, D.B., & Dryzek, J.S. (1987). *Policy analysis by design*. Pittsburgh: University of Pittsburgh Press.

Brint, S., & Karabel, J. (1991). Institutional origins and transformations. In W. Powell & P. DiMaggio (Eds.), *The new institutionalism in organizational analysis* (pp. 311–37). Chicago: University of Chicago Press.

Brunsson, N. (1989). *The organization of hypocrisy: Talk, decisions and actions in organizations*. Chichester: Wiley & Sons.

Brunsson, N., & Olsen, J.P. (1993). *The reforming organization*. London/ New York: Routledge.

Bryk, A.S. (Ed.). (1983). *Stakeholder-based evaluation* (New Directions for Program Evaluation, no. 17). San Francisco: Jossey-Bass.

Cousins, J.B., & Earl, L.M. (Eds.). (1995). *Participatory evaluation in education: Studies in evaluation use and organizational learning*. London: Falmer Press.

Cyert, R.M., & March, J.G. (1963). *A behavioral theory of the firm*. Englewood Cliffs, NJ: Prentice Hall.

DiMaggio, P. (1988). Interest and agency in institutional theory. In L.G. Zucker (Ed.), *Institutional patterns and organizations* (pp. 3-21). Cambridge, MA: Ballinger.

Edelman, M. (1964). *The symbolic uses of politics*. Urbana: University of Illinois Press.

Edelman, M. (1971). *Politics as symbolic action: Mass arousal and quiescence*. New York: Academic Press.

Egelund, N., & Larsen, B. (1989). *En udenlandsk vurdering af folkeskolen*. København: Under visningsministeriet.

Eliason, L.C. (1988). *The politics of expertise: Legitimating educational reform in Sweden and the Federal Republic of Germany*. Unpublished doctoral diss., Stanford University, Palo Alto, CA.

Fay, B. (1975). *Social theory and political practice*. London: George Allen and Unwin.

Fetterman, D. (Ed.). (1996). *Empowerment evaluation*. Newbury Park, CA: Sage.

Fischer, F. (1980). *Politics, values and public policy: The problem of methodology*. Boulder, CO: Westview Press.

Galbraith, J. (1973). *Designing complex organizations*. Reading, MA: Addison-Wesley.

Hansen, E.J. (1992). Kampen om forskningen—Elementer til forstąelse af sektorforskningens elendighed. *Dansk Sociologi, 3*(3), 4–17.

Heffron, F. (1989). *Organization theory and public organizations: The political connection*. Englewood Cliffs, NJ: Prentice Hall.

King, J.A. (1995, October 23–25). Evolving forms of participatory approaches to evaluation: Fostering organizational learning as a consequence of evaluation. Minneapolis: University of Minnesota. Paper prepared for discussion at *Systemic and individual consequences of evaluation as a tool for management and governing of public organizations*. Symposium conducted at Stockholm, Sweden.

Lampinen, O. (1992). *The utilization of social science research in public policy*. Helsinki: VAPK-Publishing.

Lawrence, P.R., & Lorsch, J.W. (1967). *Organization and environment: Managing differentiation and integration*. Cambridge, MA: Harvard Graduate School of Business Administration.

Lindblom, C.E. (1959). The "science" of muddling through. *Public Administration Review, 19*, 79–88.

Lindblom, C.E. (1990). *Inquiry and change: The troubled attempt to understand and shape society*. New Haven: Yale University Press.

Lindblom, C.E., & Cohen, D.K. (1979). *Usable knowledge: Social science and social problem solving*. New Haven: Yale University Press.

Majone, G. (1989). *Evidence, argument, and persuasion in the policy process*. New Haven: Yale University Press.

March, J.G. (1972). Model bias in social action. *Review of Educational Research, 42*, 413–29.

March, J.G., & Olsen, J.P. (Eds.). (1976). *Ambiguity and choice in organizations*. Bergen: Universitetsforlaget.

March, J.G., & Olsen, J.P. (1989). *Rediscovering institutions: The organizational basis of politics*. New York: Free Press.

McTaggart, R. (1991). *Action research: A short modern history*. Geelong, Australia: Deaking University Press.

Meyer, J.W., & Rowan, B. (1992a). Institutionalized organizations: Formal structure as myth and ceremony. In J.W. Meyer & W.R. Scott (Eds.), *Organizational environments: Ritual and rationality* (pp. 21-44). London: Sage.

Meyer, J.W., & Rowan, B. (1992b). The structure of educational organizations. In J.W. Meyer & W.R. Scott (Eds.), *Organizational environments: Ritual and rationality* (pp. 71-97). London: Sage.

Meyer, J.W., & Scott, W.R. (Eds.). (1992). *Organizational environments: Ritual and rationality*. London: Sage.

Meyer J.W., Scott, W.R., & Deal. (1994). Ontology and rationalization in the Western cultural account. In W.R. Scott, J.W. Meyer, & Associates, *Institutional environments and organizations: Structural complexity and individualism* (pp. 9-27). Thousand Oaks, CA: Sage.

Mitnick, B.M. (1993). Strategic behavior and the creation of agents. In W.R. Scott, J.W. Meyer, & Associates (Ed.), *Corporate political agency: The construction of competition in public affairs* (pp. 90-124). Newbury Park, CA: Sage.

Moe, T.M. (1989). The politics of bureaucratic structure. In J.D. Chubb & P.E. Peterson (Eds.), *Can the government govern?* (pp. 267-329). Washington, DC: Brookings Institution.

Moe, T.M. (1990). The politics of structural choice: Towards a theory of public bureaucracy. In O.E. Williamson (Ed.), *Organization theory: From Chester Bernard to the present and beyond* (pp. 116-153). New York: Oxford University Press.

Møller Pedersen, K. (1980). *Hvad er effektmåling?* København: AKF Forlaget.

Mouritzen, P.E. (1991). *Den politiske cyklus: En undersøgelse af vælgere, politikere og bureaukrater i kommunalpolitik under stigende ressourceknaphed*. Aarhus: Politica.

Naustdalslid, J., & Reitan, M. (1992). *Kunnskap og styring: Om forskningens rolle i politikk og forvaltning* (NIBR-rapport 1992, no. 15). Oslo: NIBR.

Nilsson, K. (1992). *Policy, interest and power: Studies in strategies of research utilization* (Meddelanden frạn Socialhögskolan 1992, no. 1). Lund: Socialhögskolan.

Nilsson, K. (1993). Frạn planering till resultat—Om utvärderingsforskning i Sverige. *Politica, 25*(1), 47–55.

Nørgaard, A.S. (1996). Rediscovering reasonable rationality in institutional analysis. *European Journal of Political Research, 29*(1), 31–57.

North, D. (1990). *Institutions, institutional change, and economic performance*. Cambridge: Cambridge University Press.

Notkin, A. (1992). Forskere i snor. *Weekendavisen*, 11.–17./9.

Olsen, J.P. (1988). The modernization of public administration in the Nordic countries. *Administrative Studies, 7*(1), 2–17.

Olsen, J.P. (1991). Modernization programs in perspective: Institutional analysis of organizational change. *Governance, 4*(2), 125–49.

Ott, J.S. (1989). *The organizational culture perspective*. Pacific Grove, CA: Brooks/Cole Publishing.

Patton, M.Q. (1978). *Utilization-focused evaluation*. Beverly Hills: Sage.

Peltz, D.C. (1978). Some expanded perspectives on use of social science in public policy. In J.M. Yinger & S.J. Cutler (Eds.), *Major social issues: A multidisciplinary view* (pp. 346-357). New York: Free Press.

Pierson, P. (1993). When effect becomes cause: Policy feedback and political change. *World Politics, 45*(4), 595–628.

Polanyi, M. (1967). *The tacit dimension*. New York: Doubleday.

Premfors, R. (1982). *Research and policy-making in Swedish higher education*. Stockholm: Group for the Study of Higher Education and Research Policy, University of Stockholm.

Premfors, R. (1989). *Policyanalys: Kunskap, praktik och etik i offentlig verksamhet*. Lund: Studentlitteratur.

Pressman, J.F., & Wildavsky, A. (1973). *Implementation: How great expectations in Washington are dashed in Oakland; or, Why it's amazing that federal programs work at all, this being a saga of the economic development administration as told by two sympathetic observers who seek to build morals on a foundation of ruined hopes*. Berkeley: University of California Press.

Rist, R. (Ed.). (1990). *Program evaluation and the management of government: Patterns and prospects across eight nations*. New Brunswick, NJ: Transaction Publishers.

Rist, R. (1995). Introduction. In R. Rist (Ed.), *Policy evaluation: Linking theory to practice* (pp. xiii-xxvi). London: Elgar.

Rittel, H.W., & Webber, M. (1973). Dilemmas in general theory of planning. *Policy Sciences, 4*(2), 155–169.

Røvik, K.A. (1992). *Den "syke" stat: Myter og moter i omstillingsarbeidet*. Oslo: Universitets forlaget.

Schick, A. (1983). Incremental budgeting in a decremental age. *Policy Sciences, 16*, 1–25.

Schön, D.A. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.

Scott, W.R. (1977). The adolescence of institutional theory. *Administrative Science Quarterly, 32*, 493–511.

Shepsle, K.A. (1989). Studying institutions: Some lessons from the rational choice perspective. *Journal of Theoretical Politics, 1*(2), 131–147.

Stone, D.A. (1988). *Policy paradox and political reason*. Glenview, IL: Scott, Foresman.

Syberg, K. (1994, February). Vi bor alle på Versailles (Interview with Ulrik Horst Petersen). *Information, 19–20*, 2.

Thompson, J.D. (1967). *Organizations in action*. New York: McGraw-Hill.

Torgerson, D. (1986). Between knowledge and politics: Three faces of policy analysis. *Policy Sciences, 19*, 33–60.

Vedung, E. (1991). *Utvärdering i politik och förvaltning*. Lund: Studentlitteratur.

Vedung, E. (1992). Five observations on evaluation in Sweden. In J. Mayne, J. Hudson, M.L. Bemelmans-Videc, & R. Conner (Eds.), *Advancing public policy evaluation: Learning from international experiences* (pp. 71-84). Amsterdam: Elsevier.

Weiler, H.N. (1983). Legalization, expertise, and participation: Strategies of compensatory legitimation in educational policy. *Comparative Education Review, 27*(2), 259–277.

Weiss, C.H. (1972). *Evaluation research: Methods for assessing program effectiveness*. Englewood Cliffs, NJ: Prentice Hall.

Weiss, C.H. (1977). Research for policy's sake: The enlightenment function of social research. *Policy Analysis, 3*, 521–45.

Whyte, W.F. (Ed.). (1991). *Participatory action research*. Newbury Park, CA: Sage.

Winter, S. (1991). Udviklingen i beslutningsprocesteori: En introduktion. *Politica, 23*(4), 357–374.