Theory-Based Stakeholder Evaluation

Morten Balle Hansen[1] and Evert Vedung[2]

Abstract

This article introduces a new approach to program theory evaluation called theory-based stakeholder evaluation or the TSE model for short. Most theory-based approaches are program theory driven and some are stakeholder oriented as well. Practically, all of the latter fuse the program perceptions of the various stakeholder groups into one unitary program theory. The TSE model keeps the program theories of the diverse stakeholder groups apart from each other and from the program theory embedded in the institutionalized intervention itself. This represents, the authors argue, an important clarification and extension of the standard theory-based evaluation. The TSE model is elaborated to enhance theory-based evaluation of interventions characterized by conflicts and competing program theories. The authors argue that especially in evaluations of complex and complicated multilevel and multisite interventions, the presence of competing theories is likely and the TSE model may prove useful.

Keywords

multilevel program, stakeholder evaluation, theory-based evaluation, logic model, program theory

Introduction

In this article, we introduce a new approach to program theory evaluation called theory-based stakeholder evaluation or the TSE model for short. Although many theory-based evaluation (TBE) approaches entail a mixture of program theory and stakeholder model ideas, they fuse the program perceptions of the various stakeholder groups into one unitary program theory. However, in some evaluation contexts, the inherent complexity, conflicts, and dilemmas embedded in an intervention may be too strong for such a deliberative strategy to work. When confronted with such a context in a specific evaluation project, we felt a need to elaborate a novel approach to TBE. We call this approach the theory-based stakeholder evaluation (TSE) model; and in this article, we present the raison d'être of this model as well as some ideas concerning how to conduct it in evaluation practice. Our TSE model keeps the program theories of the diverse stakeholder groups apart from each other and from

[1]University of Southern Denmark, Campusvej, Odense, Denmark
[2]Uppsala University, Uppsala and Gävle, Sweden

Corresponding Author:
Morten Balle Hansen, Department of Political Science and Public Management, University of Southern Denmark, Odense Campus, Campusvej 55, DK-5230 Odense M, Denmark
Email: mbh@sam.sdu.dk

the program theory embedded in the institutionalized intervention itself. The utility of the TSE-based stakeholder model is demonstrated by showing how it has been applied in an evaluation of a country-wide, multilevel, and multisite administrative reform project in the Danish eldercare sector.

The article is structured as follows. First, literature emphasizing theory-based and stakeholder-based evaluation practices is reviewed. Second, the general line of reasoning and practicing of a theory-based stakeholder approach is elaborated and illustrated by means of an evaluation of a country-wide, multilevel, and multisite administrative reform project in the Danish eldercare sector. Third, the pros and cons of our theory-based stakeholder model are discussed and compared to other approaches. Finally, some conclusions and implications for future evaluation practices are advanced.

Literature Review

For more than 30 years, evaluators of social programs have developed a number of "theory-based" approaches to evaluation. TBE has been labeled in various ways (e.g., program theory, theory-driven, and logic models) but what unites the different approaches is the reconstruction of a causal model based on various sources to understand how a specific intervention (or program) is intended to bring about, or brings about, output and outcomes (Stame, 2004). A good overview of the relevant U.S. body of the literature on TBE is provided by Weiss (Weiss, 2007). In an earlier contribution (Worthen, 1996) traces the U.S. development back to the Fitz-Gibbon and Morris 1975 article "Theory-Based Evaluation" (FitzGibbon & Morris, 1975), but already Suchman (1967) referred to the notion of program theories (Weiss, 2007). Europeans have also contributed to this important part of the evaluation literature (Hoogerwerf, 1990; Leeuw, 1991, 2003; Stame, 2004; Vedung, 1997). TBE has made important progress over the years but also faces a number of problems (Weiss, 1997, 2007). One of the problems that we believe has acquired too little attention is how to conduct the approach in more complex and potentially conflict-ridden contexts. Most applications of TBE use the intervention theory approach while presuming, constructing, or imposing the hegemony of one dominant intervention theory.

Although evaluators typically work with several stakeholder groups, the purpose usually is to create one intervention theory on which all of them can agree. However, in some types of evaluations it may be impossible, extremely difficult or for some reason not desirable to reduce complex intervention perceptions into one overarching intervention theory. One reason may be that in their implementation, many interventions are meant to involve several groups of actors in very different working situations and with very different expectations to the intervention. This is particularly evident in nation-wide, multisite, and multilevel interventions. In these cases, aside from the theory embedded in the program itself, many more program theories often exist among various stakeholder groups in various localities and at various administrative and political levels concerning the implementation, output, and impact of the program as well as suggestions for its improvement. Even in the case of single-site, single-level evaluations, strong competing beliefs and perceptions concerning the means and ends of a program may be present (Dahler-Larsen, 2001). In such cases, we recommend applying the stakeholder approach in the sense of keeping the diverse stakeholder theories separate from each other.

The TSE model that we will explicate in this article will be grounded in literature on theory-based and stakeholder-based approaches to evaluation practice. In this literature review, we intend to answer two questions:

1. How do evaluators usually apply the theory-based evaluation approach?
2. How do evaluators usually combine the theory-based evaluation approach with the stakeholder approach?

We have consulted a broad selection of the quite substantial pertinent literature. We searched in the databases Social Science Citation Index, Sociological Abstracts, and International Bibliography of the Social Sciences using various search terms for theory-based and stakeholder approaches and soon discovered the enormous amount of literature within this field. We decided to delimit our further search to four prominent journals: *American Journal of Evaluation, Evaluation, Evaluation Review*, and *New Directions in Evaluation*. We have consulted all articles in these four journals, which use some version of the theory-based approach and a stakeholder approach. We have also consulted often cited important books and encyclopedia entries (Chen, 1990, 2005a; Donaldson, 2007; Rossi, Freeman, & Lipsey, 2004). Yet, for obvious reasons, we can include in our article only a few very important previous contributions to clarify how our TSE model differs from and resembles previous practices and thus represents a novel evaluation model.

TBE, Theory-Driven Evaluation, Theory-Led Evaluation, Program Theory, and Logic Models

Most evaluators and evaluation theorists advocating or pursuing theory-based (Friedman, 2001), theory-driven (Bledsoe & Graham, 2005; Sidani & Sechrest, 1999; Turnbull, 2002; Worthen, 1996), or theory-led (Molas-Gallart & Davies, 2006) evaluation discuss and apply the program theory idea (Rogers, 2007). For instance, Chen (Chen, 2005c), a major contributor to this literature and practice, at the beginning of his encyclopedia entry "Theory-driven evaluation" adds the words "or program theory-driven evaluation." In this article, we use "intervention" as the main term with "program" as a synonym. We suggest that "intervention theory" is the most appropriate general term, because interventions may take place at levels other than the program level (regime level, policy level, project level, treatment level, individual level, etc.).

What do evaluation theorists and evaluators mean when they promote or adopt the program theory approach? Surprisingly, in all the literature we have consulted, the program theory approach is grounded in stakeholder involvement. And in the main, the various stakeholder program theories are *not* kept apart. Instead, program theory evaluators attempt to fuse all stakeholder conceptions into one unitary program theory behind which all stakeholders may rally.

Stakeholder Involvement to Create One Unitary Program Theory

Let us consider what one of the most used textbooks on evaluation in the United States as well as the rest of the world (Rossi, Freeman, & Lipsey, 1999) has to say about program theory.

In their 33-page chapter entitled "Expressing and Assessing Program Theory," the authors emphasize that stakeholders should be involved in the endeavor of articulating and representing the program theory. The authors argue "To describe the theory . . . , the evaluator must interact with the program stakeholders to draw out their implicit program theory, that is, the theory represented in their actions and assumptions" (p. 162). And later, "The most straightforward approach to developing a description of a program's theory is to obtain it from program personnel and other pertinent stakeholders." The most important sources of information describing a program and contributing to the articulation of the program theory are those persons with firsthand knowledge and experience of the program (p. 164). Yet, a review of program documents, site visits and observation is also helpful (pp. 162–164).

Because there is no hint of divergences among the stakeholders in all of this, the evaluator does not have to keep dissimilarities among stakeholders apart. In the end, the evaluator will have one unified program theory. The authors conclude, "This process continues until the stakeholders find little to criticize in the description." Whose theory is constructed in this approach? In our interpretation, it is the theory of the evaluator but negotiated with and agreed upon by the involved stakeholders. Our primary problem with this approach is that in a number of evaluation contexts,

including our example in the next section, such an agreement may be hard to accomplish and therefore perhaps not warranted.

The second example of how TBE combined with stakeholder evaluation gets *all* pertinent stakeholders to agree on and support *one* intervention theory is drawn from an article by Birckmayer and Weiss (Birckmayer & Weiss, 2000): "Possible *sources* of program theories are social science theories and research, prior evaluations, planner and practitioner expectations, the evaluator's knowledge and experience with programs of similar type, and her or his own logical thinking. Often, the evaluator will cycle through several of these sources—asking program planners and program managers, reviewing existing theories in the field, reviewing previous research and evaluations, hypothesizing a theory on the basis of this information, and then *negotiating* (our italics) her formulation with program managers and staff to come to an agreement that accords with their thinking" (426–427).

The program theory articulated is a theory of *how the program actually works* according to the program planners, program managers, and program staff once these categories of stakeholders *have participated in interactive sessions among themselves and with the evaluators as brokers*. Still, the end product is not one theory per stakeholder group but one program theory common to all groups involved. In our interpretation, this approach to theory construction is very similar to the approach of Rossi et al. in our first example. What is constructed is actually the theory of the evaluator but informed by relevant theory and negotiated with and agreed upon by the stakeholders.

The third example is taken from Chen who, since the publication of a famous book on the subject (Chen, 1990), has been considered one of the prominent proponents of TBE. In the encyclopedia articles "Theory-driven Evaluation" and "Program theory" (Chen, 2005b, 2005c), Chen offers a definition of program theory: "a set of implicit or explicit assumptions of how the program should be organized and why the program is expected to work" (p. 415). Stakeholders are discussed at length. It is clear that by program theory Chen means the program theory of the stakeholders on how the program actually works. The goal is to reach agreement among the stakeholders about one unitary program theory. "Program theory belongs to stakeholders" (p. 347). Any imposition of the evaluator's own values on stakeholders should studiously be avoided, according to Chen (p. 418).

As one among few, Chen notes that there might be several stakeholder program theories. Often, agreement among stakeholders about what the program theory should look like is not difficult to reach. Even if some components of the program do spark disagreement between key stakeholders, this is not an obstacle to evaluation. Rather, disagreement means that evaluators should test various hypotheses during their investigation (p. 419). Chen is open to the idea of keeping diverse stakeholder theories apart from each other but does not elaborate further on this idea or the reasons for it. We consider the approach we present here as an attempt to expand upon and deepen this particular idea in Chen's work. Especially, in a political context with multiple competing theories, it is important to keep the theories of different stakeholders apart.

To repeat, our literature review on stakeholder involvement and program theory approach has been far more extensive than our three above examples show. Yet, as far as we have been able to ascertain, attempts to combine the two approaches issue into *one* common, unitary theory (Leeuw, 2003; McLaughlin & Jordan, 1999). Our conclusion is in line with Donaldson's (Donaldson, 2007), who certifies that "evaluators typically work with stakeholders to develop a *common understanding* of how a program is presumed to solve a problem" (italics ours). In addition, Donaldson also accept this procedure for his own "participatory program theory evaluation" (pp. 22, 26, 34, 56, 62, 63, etc.). "A special emphasis was placed on showing how program . . . theory can help evaluators and stakeholders develop a common understanding of a program," he concludes his exposition (p. 39).

In many evaluative contexts, this may certainly be an appropriate theory-based approach. However, in complex multilevel and multisite programs or in conflict-ridden programs, it may not

be possible even for the best and most able process consultant to achieve such a common shared understanding. In such cases, the most important outcome of the evaluation may be to sort out the main features of the different program theories and the reasons for these disagreements.

Stakeholder Involvement Where Diverse Program Theories are Kept Apart

In our searches for contributions that are both stakeholder oriented and recommend keeping program theories of diverse stakeholders apart we have found only few cases. Under the heading "Competing Programme Theories," Dahler-Larsen (2001, p. 342ff) states that "more than one programme theory may come into play in a given context." His example is a school for social workers whose teachers hold two different views, based on dissimilar pedagogical philosophies, on what it means to be a good social worker. While keeping these two views clearly apart, he does not elaborate since his major concern in the article is something else. In addition, other authors touch upon the subject in passing (Chen, 2005c, p. 419; Weiss, 1997, p. 509).

Jos Vaessen (Vaessen, 2006) combines a program theory approach with application of various stakeholder values to program implementation and effects (2006, pp. 397, 403ff). First, he investigates how stakeholders think that a program works and is expected to achieve certain outcomes. Then he uses this (with other information like program documents) as "building blocks for *a* (descriptive) programme theory," in other words to build *one* unitary program theory. By eliciting the value criteria of the stakeholders, the evaluator can determine the major stakeholder value positions regarding a program. These criteria would then be the basis for the evaluation of *the* descriptive program theory. Although the value criteria and actually performed evaluations of the stakeholders are kept apart, their causal program theories are not; they are used as building blocks to erect one program theory. This makes Vaessen's contribution different from the TSE model we will elaborate on later in the article.

A more extended treatment is found in an article by Friedman (Friedman, 2001), which makes his contribution rather close to the TSE model that we will outline in this article. Friedman draws upon evaluation efforts by Bowen of a local adolescent pregnancy and parenting program to provide services that strengthen family life as well as self-support and personal independence.

In trying to understand the mixed outcomes that ensued, Bowen analyzed program theories at the following four levels: (1) Theories of the program planners, (2) Program theories held by the program staff members, that is, the case managers who worked with the teenagers, (3) A comparison of the case managers' espoused theory with their theory in use, and (4) Feedback to the case managers of the results of the evaluator's comparison between 2 and 3.

The approach of carefully distinguishing between the program theories of several stakeholder groups resembles our approach, which we will elaborate later in Section 3. However, our concern here is not to emphasize the particular distinction between "theory-in-use" and "espoused theory" but a more general interest in opening up the theory-based approach for tensions, disagreements, and conflicts that cannot be resolved by the evaluator in the process. What we find important in Friedman's contribution is the notion of "designed blindness" toward possible conflicts and tensions. We think that evaluation practice needs models to explicitly cope with strong dissension and political conflicts and we hope to make a useful contribution to achieving this end with our TSE model, which will be elaborated in the next section.

The TSE Model

As clarified in the last section, the intervention theories of stakeholders have been used so far by evaluators performing TBEs as inputs in their construction of one unitary program theory.

In the TSE model proposed here, the theories of the primary stakeholder groups are kept carefully apart and compared with each other to clarify their similarities and differences. In this section, we

will spell out the basic thrust of our TSE model and how it differs from previous ways of practicing TBE. We will do that in general terms and we will illustrate our model by means of an evaluation of a national, multilevel, and multisite program conducted in the Danish eldercare sector.

In our TSE model, we suggest three specific analytical tools intended to guide the systematic reconstruction and comparison of intervention theories, which will be elaborated and illustrated in this section (see Table 1). First, we suggest the application of a *principle of reason* to reconstruct partly tacit intervention theories. Second, we recommend that evaluators structure various parts of the intervention theories into a *tripartite scheme* of analysis, namely, situation theories, causal theories, and normative theories. Third, we also recommend that the raw theories should be organized according to the *extended system model*, that is, the input–conversion–output–outcome model. All relevant raw stakeholder theories can be usefully packed into this scheme. The concepts of the tripartite scheme and the extended system model are added by the evaluator; they are very rarely found in the raw intervention theories themselves.

We furthermore recommend that evaluators *insert pertinent actors* into the reconstructed program theories. By actors we mean, in this context, collective actors, not individual persons. The input–conversion–output–outcome model is actor free. It only exhibits actions (activities) and states of the world. However, it will become a more illuminating tool of analysis if actors are introduced. Examples of actors are agency top management, agency middle management, agency line-level workers who actually deliver the goods and services, and recipients. If a couple of these actors are added, the model looks as follows: input–administrative agencies–conversion–output–recipients–outcome.

There are alternative ways of proceeding once the embedded intervention theory and the intervention theories of the various primary stakeholders are elicited and arranged according to the above guidelines. An assortment of possibilities, including the ones suggested by us, will be extensively elaborated in this section.

First, however, we elucidate our principled mode of conceptualizing intervention theory, which is more inclusive than most other conceptions suggested in the literature.

Definition of Intervention Theory

Intervention theories[1] are defined as presuppositions (notions, conceptions, and assumptions) of how an intervention (a program, a policy, a treatment, and an organizational change) may have an impact on a given situation and change it or preserve it in ways that are preferable or not preferable to the situation without the intervention or with another intervention. In particular, we suggest that in its fully developed form, an intervention theory is composed of three elements:

Situation theory. Notions concerning relevant features of the context in which an intervention is supposed to take place. Some or all of the following notions might be relevant to include (a) size of the problem at present; (b) development of the problem up till the present time; (c) size of the problem after x years if nothing is done; (d) causes of the problem, and finally (e) impact of the problem at present and after x years if nothing is done.

Causal theory. Notions concerning how a given intervention/program directly or indirectly through its process of implementation and delivery of outputs will (a) have an impact on the causes of the underlying problem at issue so that the problem will disappear, be reduced, or prevented from becoming aggravated and (b) have possible effects in other areas during the entire process.

Normative theory. Notions concerning why the various aspects of the situation that are supposed to be affected by the intervention are preferable or not preferable to the situation without the intervention or with another intervention.

Some conceptions of program theory omit situation theory (e.g., Chen, 1990). Others omit both situation theory and normative theory from the concept. Bickman (1987, p. 5), for instance, defines program theory as "a plausible and sensible model of how a program is supposed to work" (Bickman, 1987). According to Bickman's conception, program theory is synonymous with our causal theory. In such conceptions, the normative theory is often taken for granted, while the situation theory is more or less included as part of the causal theory.

We suggest three arguments for the utility of the distinction between the three elements in a fully developed program theory. First, the tripartite distinction provides useful guidance in the process of constructing intervention theories. Second, in the case of competing intervention theories, which is the raison d'être of our TSE model, the three elements provide useful guidance for comparison. Third, in formative and deliberative evaluations, the clarification of differences and similarities on the three dimensions can be a powerful starting point for discussions and the possible emergence of an intersubjective consensus or enlightened disagreement.

We now proceed to our TSE model.

Five-Step Model of TSE

As can be seen in Figure 1, we suggest a sequential model of five steps when practicing TSE. We will now explicate the characteristics and practices of each of the five steps. We will illustrate our methodology by means of an evaluation of a national, multilevel, multisite information and communication technology (ICT) program adopted and implemented in the eldercare sector in Denmark (Hansen & Vedung, 2005).

I. Reconstruction of the Raw Intervention Theory Embedded in the Intervention Itself. The application of our TSE model starts with the evaluator's reconstruction of the raw theory inherent in the intervention being evaluated, including its situation, normative, and causal theory parts and its views of the intervention's implementation process, output, and outcome effects. We recommend that this reconstruction be based on the intervention decision per se, that is, how the intervention is described in the formal decision to embrace it as well as by examining written documents and oral viewpoints produced by the initiators at the time the program was adopted or maybe somewhat later. Contrary to several practitioners of TBE, we strongly emphasize the issue of whose theory it is that we are actually attempting to represent. In this first step, it is the theory inherently implanted in the program as adopted, as institutionalized, that we try to assemble, not the program theory of any stakeholder.

This type of reconstruction is not that unusual in the literature. A more uncommon trait is our use of a Principle of Reason, which resemble and is inspired by the famous "Principle of Charity," propagated by prominent philosophers (Davidson, 2006; Quine, 1960) and practiced by some social scientists (Galtung & Naess, 1955). It is an attempt to cope with the practical problem faced by evaluators when there is no extensive and clear formulation of the raw program theory in official documents or by formal initiators or pertinent stakeholders. It is an approach to interpretation that prescribes the interpreter to provide the best and strongest possible rendering of an argumentation. The aim of this methodological principle is to avoid attributing irrationality, logical fallacies, or falsehoods to the statements of others when another reasonable, coherent, sensible interpretation of their statements is possible. The Principle of Reason, or of Rationality as Popper called it (Popper, 1985), can be seen as the opposite strategy of setting up a straw man by choosing a very negative interpretation of the adversary's position as the object of injunctions. It implies that concerned agents are assumed to think and act adequately or appropriately to the situation in which they are embedded, that is, in accordance with the situation (Popper, 1985). The Reason principle should be regarded and used as a *methodological postulate*, that is, as a *tool* or an *instrument*. As a tool in the application of the intervention theory method, the principle should be used to reconstruct
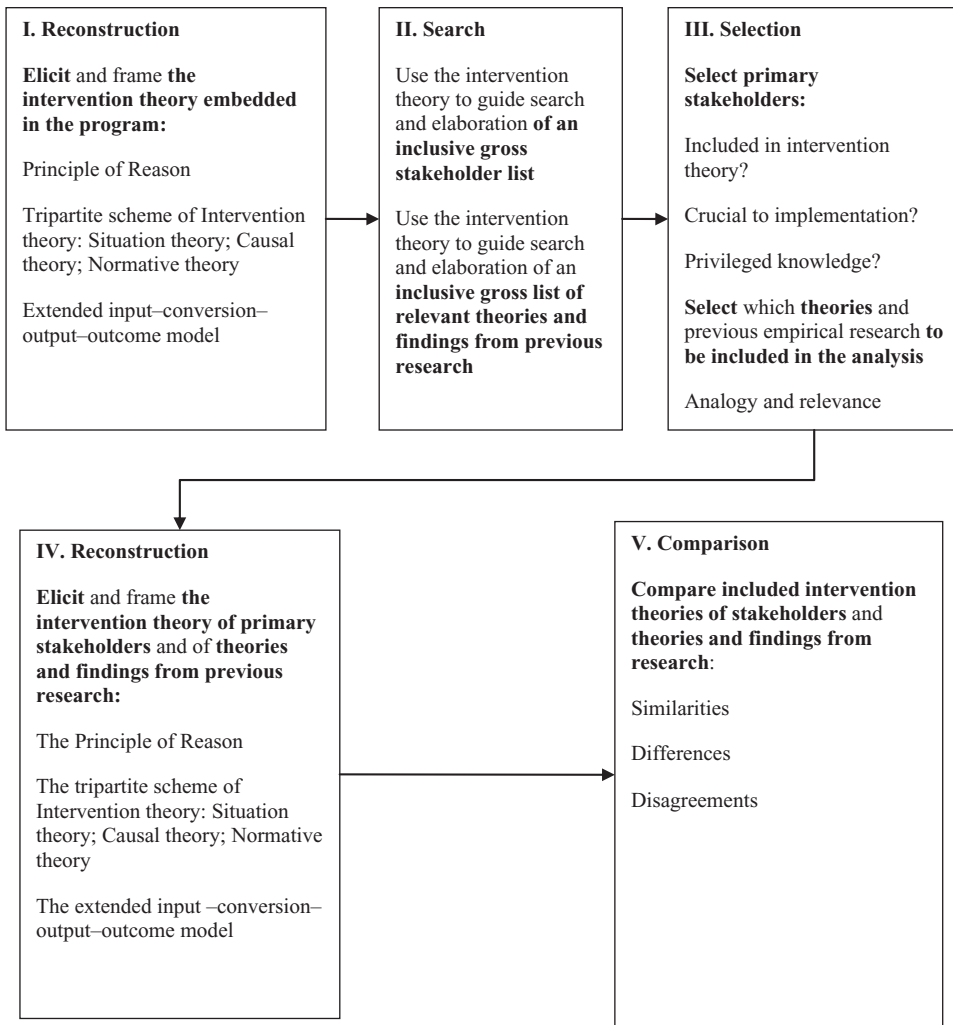
Figure 1. The five steps of theory-based stakeholder evaluation (TSE).

eventual tacit or vaguely intimated ingredients of the raw intervention theory, given its premises and situational understanding. If we find some gaps or omissions in the raw intervention theory inherent in program decisions and preparatory documents, we do not end our interpretation by concluding that the theory is basically flawed in these respects. Instead, we assume that it is meant to be adequate or appropriate to the situation and try to fill out the gaps and omissions from this point of departure. Of course, the same procedure is used as far as stakeholder theories are concerned.

Besides the *Principle of Reason* and *the tripartite scheme* of situation, causal and normative theory described in Section 3.1, we suggest that the *extended input–conversion–output–outcome system model* provides useful categories for the evaluator's reconstruction and analysis of the situation, causal, and normative notions in the material.

Illustration: An evaluation of a nation-wide, multilevel, multisite administrative ICT reform in the Danish eldercare sector. How does one apply the suggested methodology to actual evaluation

practice? Our illustration, which has been an important vehicle for the elaboration of our TSE model, is taken from an evaluation of an ICT reform. Since the 1980s, numerous attempts to take advantage of the possibilities provided by the new ICT have been made in the public sector. The illustration is derived from a research project conducted from 2002 to 2005 of a controversial ICT intervention called Common language (CL, Fælles Sprog) in the organization of the Danish eldercare sector that took place from 1992 to 2004.

In our research project, the intervention theory of the CL program as such was initially developed through deskwork in an attempt to reconstruct how the program was supposed to function according to the assumptions laid down in 1998, in the first final version of the program, authorized and recommended by "Local Government Denmark (KL)," the national Danish association of municipalities. We also consulted published speeches and conference reports from the 1990s and interviewed a few key actors. The basic structure of the program-inherent intervention theory can be usefully presented by integrating the distinction between situation theory, causal theory, and normative theory into the input–administrative agencies–conversion–output–recipients–outcome model discussed previously. This way of representing this particular intervention theory is shown in Figure 2.

It is important to emphasize that Figure 2 does not represent our or any specific stakeholder's intervention theory. It represents our interpretation of the theory inherently embedded in the CL program structured according to the input–administrative agencies–conversion–output–recipients–outcome framework as well as the tripartite distinction between situation, causal, and normative theory in which the principle of reason outlined in Section 3.2 has been used. Thus, neither in the authoritative 1998 decision by Local Government Denmark nor in texts related to it or among single decision makers has Figure 2 been presented in its current form. It is an ideal type put together by us, evaluators, to clarify the logical structure of a raw intervention theory. At the same time, Figure 2 does not represent the intervention theory of us, evaluators, either. Using our conceptual tools, we have attempted to build a model loyal to the substantive and logical structure embedded in the decision and the written and oral arguments in favor of that decision presented to us when we were doing our research. It should also be emphasized that the model presents the theory of how the intervention/program was supposed to work—not how it actually worked once it had been adopted and implemented in Danish municipalities.

Following the structure of Figure 2, we now proceed with a short clarification of the composition of the intervention theory of this complex and complicated multilevel, multisite ICT reform of the Danish municipal eldercare sector.

Figure 2 starts with *the situation theory* inherent in the program (box A). The situation theory describes the problems that the intervention is supposed to face and counteract. In this case, the increasing costs to eldercare, the perceived lack of transparency, accountability, and control, and the unclear criteria for entitlements to eldercare service. Sometimes, the situation theory also includes some notions of causes of the problems to be tackled by the intervention.

After the situation theory follows *the causal theory* (boxes B–J), which is spelled out in greater detail to clarify the program's purported cause–effect structure (see below). Finally, Figure 2 ends with *the normative theory*, which expresses the desirability of the envisioned outcome of the intervention. In this case, the rational dream is improved efficiency, accountability, transparency, and justice in the decision-making process concerning eldercare at multiple levels, in the delivery of eldercare services at the local municipal level, and in the legal rights of the elderly users of eldercare services. How are these improvements to be accomplished according to the *causal theory* (boxes B–J)? The basic idea in this ICT intervention is that the entire process in particular municipalities, from the adoption of the CL and through the various stages of its implementation and service delivery, should be organized in a digitalized ICT compatible way. The intervention is really an example of digital era governance (DEG) in practice (Dunleavy, Margetts, Bastow, & Tinkler, 2006). To integrate ICT, the invention of a digital CL suitable for computers and adapted to the
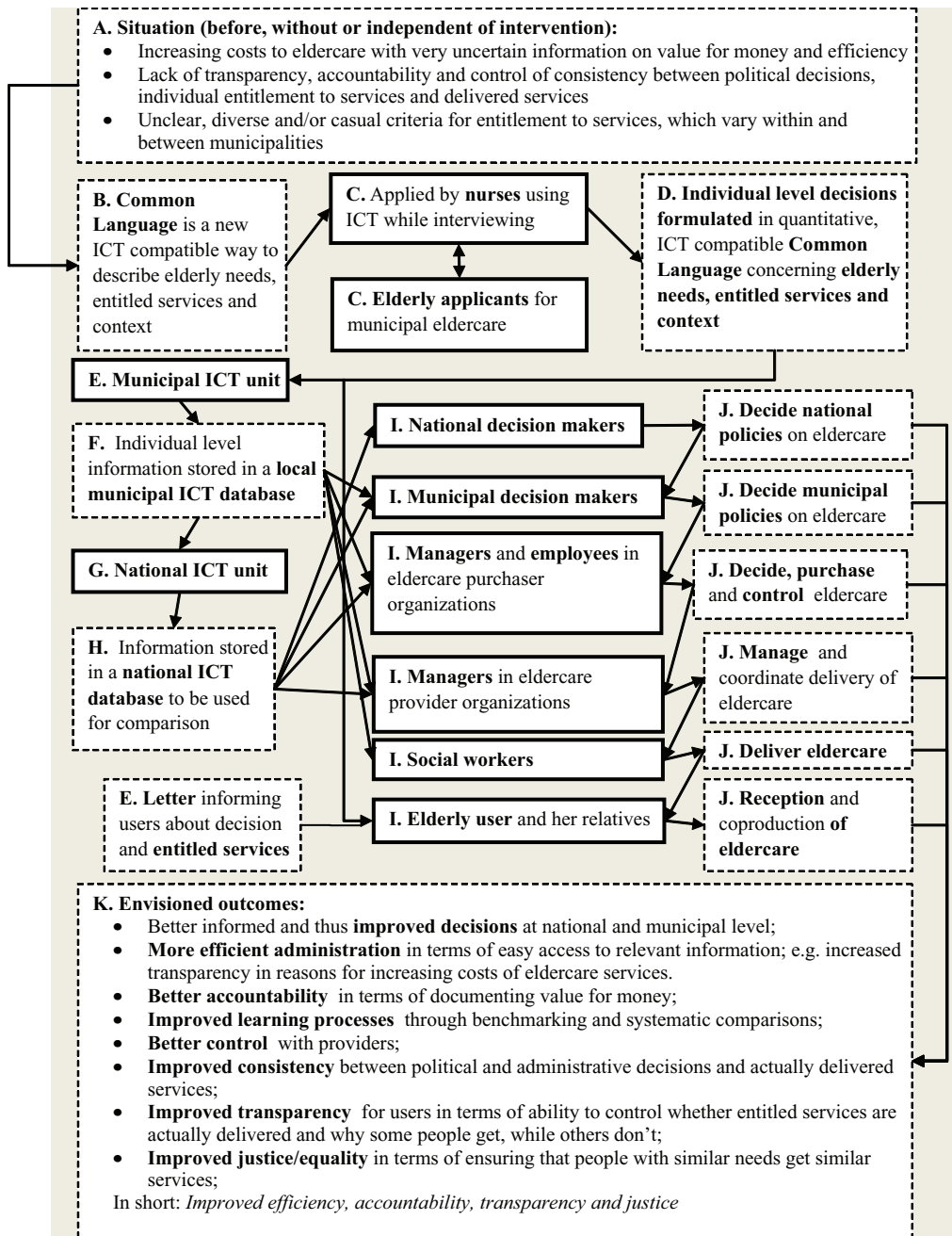
A. Situation (before, without or independent of intervention):
- Increasing costs to eldercare with very uncertain information on value for money and efficiency
- Lack of transparency, accountability and control of consistency between political decisions, individual entitlement to services and delivered services
- Unclear, diverse and/or casual criteria for entitlement to services, which vary within and between municipalities

B. Common Language is a new ICT compatible way to describe elderly needs, entitled services and context

C. Applied by **nurses** using ICT while interviewing

C. Elderly applicants for municipal eldercare

D. Individual level decisions formulated in quantitative, ICT compatible **Common Language** concerning **elderly needs, entitled services and context**

E. Municipal ICT unit

F. Individual level information stored in a **local municipal ICT database**

G. National ICT unit

H. Information stored in a **national ICT database** to be used for comparison

I. National decision makers

I. Municipal decision makers

I. Managers and **employees** in eldercare purchaser organizations

I. Managers in eldercare provider organizations

I. Social workers

E. Letter informing users about decision and **entitled services**

I. Elderly user and her relatives

J. Decide national policies on eldercare

J. Decide municipal policies on eldercare

J. Decide, purchase and **control** eldercare

J. Manage and coordinate delivery of eldercare

J. Deliver eldercare

J. Reception and coproduction **of eldercare**

K. Envisioned outcomes:
- Better informed and thus **improved decisions** at national and municipal level;
- **More efficient administration** in terms of easy access to relevant information; e.g. increased transparency in reasons for increasing costs of eldercare services.
- **Better accountability** in terms of documenting value for money;
- **Improved learning processes** through benchmarking and systematic comparisons;
- **Better control** with providers;
- **Improved consistency** between political and administrative decisions and actually delivered services;
- **Improved transparency** for users in terms of ability to control whether entitled services are actually delivered and why some people get, while others don't;
- **Improved justice/equality** in terms of ensuring that people with similar needs get similar services;

In short: *Improved efficiency, accountability, transparency and justice*

Figure 2. Embedded intervention theory of the full-scale intervention (actors in bold boxes).

context of eldercare was necessary (see box B), hence the name of the intervention. In the first version of the CL, the one analyzed here, this language included a controversial category scheme enabling its operator to enter data on individual elderly needs (physical and psychical functional ability), of municipal services that might be provided to meet these needs (e.g., cleaning, homecare,

and training) and of contextual factors. This language was envisioned to be integrated in the entire process as is illustrated in boxes B–J.

According to the program theory embedded in the program, elderly applicants for municipal eldercare are supposed to be interviewed by nurses using the CL to categorize and make decisions concerning the needs, entitled services, and context of each elderly person (boxes C–D). These individual-level decisions are assumed to be communicated to the municipal ICT unit in digital form and to each elderly applicant/user in a letter (box E). The municipal ICT unit is supposed to store the information in a local database (box F) and transmit it to a national ICT unit (box G), which is expected to store the standardized information in a national database (box H). According to the theory, the information from the local and national databases are assumed to be used by national and municipal political decision makers and by managers and employees in the municipal purchaser and provider organizations (box I) to improve decisions, management, coordination, and delivery of eldercare (box J). Elderly clients and their relatives are presumed to use the entitlement letter (box E) to control whether the delivery of services is in accordance with those entitled (boxes I and J). Social workers delivering eldercare (boxes I and J) are expected to carry personal digital assistants (PDAs) with the entitlement decisions stored in CL and use these PDAs to electronically report actually delivered eldercare in the same categories immediately after the work is done. Because they are expected to use this information to control and coordinate eldercare provision, managers of eldercare provision are held accountable by the managers and employees of the municipal purchaser organization (boxes I–J), who get the information immediately from the municipal ICT database (box F). The municipal purchaser organization can use the information from the municipal and national database (boxes H and F) to make internal and external comparisons to learn from the more efficient internal units and external municipalities. When elderly users or their relatives complain about eldercare provision to the purchaser organization, employees can easily check whether the services delivered are in accordance with the services to which they are entitled. They can also easily check the reasons for the decisions concerning entitlements. At higher levels, municipal and national decision makers can use the information to get value for money accounts, to check whether individual-level decisions are in accordance with political decisions on principled service levels to be provided, and to benchmark the performance of some municipal units against the performance of others.

The reconstruction above resembles quite closely the recommendations found in other explications of TBE. However, there is one major difference. Because we emphasize the distinction between multiple program theories, we have been very careful to spell out whose program theory it is that we are reconstructing. As we have just argued in the review section, other explications of TBE emphasize this point too little. Furthermore, as a general rule, we have recommended starting with the institutionalized theory embedded in the program. Finally, we have deliberately described the institutionalized program theory in some detail to illustrate how the tripartite scheme, the extended system model, and the principle of reason can be usefully applied to reconstruct the intervention theory embedded in the program.

Having elicited and reconstructed the theory embedded in the program, we now proceed to the next step (box II in Figure 1) in our TSE model.

II. Search. The second step involves using the refined intervention theory now derived to guide the search for all the possible relevant stakeholders.

A *stakeholder* is an actor (a collectivity in our case, not an individual person) expressing or actually entertaining a concern for the intervention, its activities, implementation, outputs, and outcomes or otherwise having an interest vested in or being affected by the intervention. Concerns may be counted in terms of money, status, current employment, power, face, opportunity, lifestyle, or some other coin (Guba & Lincoln, 1989, p. 51). Stakeholder groups can also be pigeonholed according to diverse criteria (ethnic group, social class, gender, age, lifestyle, level of education, current job,

etc.). We recommend classifying stakeholder groups from the point of view of their position in the social structure related to the intervention.

Illustration of Step II from the Danish ICT intervention system. As emerges from the bold boxes in Figure 2, the relevant stakeholders are a rather diversified batch:

- Municipal nurses categorizing elderly needs and deciding individual eldercare entitlements
- Elderly applicants for municipal eldercare
- Municipal ICT unit
- National decision makers
- Municipal decision makers
- Municipal managers in eldercare purchaser organizations
- Municipal managers in eldercare provider organizations
- Municipal social workers providing eldercare to individual users
- Elderly users and their relatives

It is easy to imagine that there might be conflicts between the various groups. The elderly applicants, users, and their relatives might be dissatisfied with the entitlement decisions of the municipal nurses and the eldercare actually provided to them. The municipal nurses might be dissatisfied with the necessity to use information technology (IT) equipment. The municipal social workers might be dissatisfied with the loss of degrees of freedoms imposed on them by the IT system and of the decisions of the municipal managers in provider organizations and with the attitudes of the elderly users. National decision makers might be dissatisfied with the low quality of data provided by the system.

III. Selection. The third step in the model is to decide which stakeholders should be included in the continued evaluation process. This decision involves important value judgments. The basic assumption in the theory-based stakeholder model is that values, beliefs, and capabilities vary substantially between actors in a social system and that important parts of this variation arise from differences in position within the social structure. The notion is that ''where you stand depends on where you sit.'' In this case, due to their different positions, we expect that municipal nurses responsible for categorizing elderly needs and deciding individual eldercare entitlements to exhibit values, beliefs, and capabilities different from those of municipal managers in eldercare provider organizations responsible for coordinating and controlling the delivery of eldercare services and thus are likely to evolve different theories of the functioning of the program.

Because most evaluations have scarce resources, a deliberate choice between which actors to include in the analysis has to be made. This choice depends on a number of premises, among which the purpose of the evaluation is perhaps the most important.

Various criteria for inclusion and exclusion have been suggested in the stakeholder evaluation literature. We suggest four criteria for selecting primary stakeholders:

1. Stakeholders pointed out in the program's own intervention theory should be considered.
2. Stakeholders decisive for the vagaries of implementation and practical functioning of the program should be considered.
3. Stakeholders presumed to hold privileged context-specific knowledge of the implementation, functioning, and possible improvement of the program should be considered. Often, this group is similar to the previous group.
4. Actors actively engaged in the policy area in which the evaluated program is implemented may be taken into consideration, particularly if their activity seems to trigger consequences for the execution and performance of the program.

Illustration of Step III from the Danish ICT intervention. In the evaluation of our Danish case, we decided to carry out a two-stage inclusion of actors. First, following the logic of the reconstructed embedded program theory in Figure 2, rough estimates of the intervention theories of seven groups of actors were made: (1) National decision makers, (2) municipal decision makers, (3) municipal managers in eldercare purchaser organizations, (4) municipal nurses categorizing elderly needs and deciding individual entitlements, (5) municipal managers in eldercare provider organizations, (6) municipal social workers providing eldercare, and finally, (7) elderly users of eldercare.

Because the purpose of our evaluation was to suggest how to improve the CL and its utilization in the organization of the eldercare policy process, it soon became apparent that most of the seven actor groups had a very superficial and vague understanding of the CL and its application. Following the criteria of importance to implementation and in-depth knowledge of the functioning of the program, we decided to select three of the stakeholder groups for closer inspection. These groups were, following the above numbers: (4) municipal nurses categorizing and deciding individual entitlements, (5) municipal managers organizing the daily provision of eldercare, and (3) municipal managers eldercare purchaser organizations. These three stakeholder groups had an in-depth understanding of, but very different and conflicting perspectives on the CL program.

IV. Reconstruction. Once the primary stakeholders have been chosen for the evaluative analysis, their intervention theories must be uncovered and structured by the evaluator according to the same methodology as in Step I.

Individual and focus group interviews are often the most important means in the process of uncovering, interpreting, and framing stakeholder program theories. Articles in magazines representing the various professions involved in the program (e.g., teachers, nurses, doctors, and social workers) may also be useful sources of information. As in the first stage, when the intervention theory inherent in the intervention itself was reconstructed, the principle of reason should be practiced. In this stage, the principle of reason is often more important because intervention theories other than the one embedded in the program itself are often not supported by written statements since many stakeholders represent groups where oral and partly tacit professional knowledge carries weight. "Stakeholder theory is implicit theory," argues Chen (2005a, p. 41). "Stakeholders' implicit theories are not likely to be systematically and explicitly calculated, and so it is up to evaluators to help stakeholders elaborate their ideas."

Illustration of Step IV from the Danish ICT intervention. In the case presented here of the three primary stakeholder groups selected for additional inspection, we organized focus group interviews with each of the three groups attending to the tripartite scheme situation theory, causal theory, and normative theory. Based on these interviews, we reconstructed the intervention theory nurtured by each of the groups. These written reconstructions were later approved by each of the participants in the focus groups. To illustrate the potential advantages of the TSE model, we have selected a few findings from this analysis to be shown in the next section.

V. Comparison. The next and crucial step in the TSE model is to scrutinize all the intervention theories included in the evaluation by carefully elucidating their similarities and differences on relevant dimensions. The set of intervention theories at stake are the one constructed from the raw theory inherent in the intervention and the theories of the various included stakeholders. Particularly, in this stage, our distinction between three crucial elements of an intervention theory (situation theory, causal theory, and normative theory) is useful.

A number of questions should be raised in this stage of the evaluation process. Do all the intervention theories comprise different situation theories, causal theories, and normative theories or are they similar in important respects? Does our comparative analysis indicate glaring, apparently

irreconcilable conflicts about the workings of the intervention? Can some of these differences be resolved by simply testing competing hypotheses as suggested by some proponents of the theory-based approach? Do important divergent interests or ideological differences indicate that strong obstacles to the implementation of the intervention are at hand or can be foreseen? Or is it possible to amalgamate and integrate these differences into a new, more subtle, and nuanced intervention theory? Does the analysis indicate important trade-offs between different desirable goals? Are these goals given different weights by different groups of stakeholders?

In particular, we consider three questions to be crucial: (a) On what aspects of the intervention theory (situation theory, causal theory, and normative theory) do the primary stakeholder theories differ, but not necessarily disagree (differences)? (b) On what aspects do the primary stakeholder theories clearly disagree (disagreements)? (c) On what aspects do all the primary stakeholder theories agree (similarities)?

Illustration of Step V from the Danish ICT intervention. Turning to the first question concerning differences, striking dissimilarities between the three groups were detected. For instance, as one might expect, only the managers of the purchaser organizations were interested in the external benchmarking part of the program's embedded intervention theory. These managers agreed that the current quality of the generated data was not adequate for comparison, but they disagreed within the group whether it was worth the effort to improve on this aspect. Furthermore, only the nurses making individual-level decisions on elderly needs and entitlements had elaborated theories concerning the structure of categories in the CL program. They advanced proposals for improvement of this structure not suggested by any of the other stakeholder groups, while the managers of eldercare provision brought forward suggestions concerning how to improve the use of the program in the coordination of eldercare provision.

With regard to the second question concerning disagreements between stakeholders, the conflicts inherent in the program were not so much between the three stakeholder groups selected for deeper scrutiny. They disagreed on the relative importance of certain issues but not on their relevance. In contrast, for instance, one of the conflicts inherent in the program was clearly articulated by some of the social workers delivering eldercare. The program implies a certain loss of autonomy for social workers delivering eldercare. Decision-making competence is to some extent removed from the social worker delivering eldercare to the nurse deciding on individual entitlements. Social workers aired a certain frustration about this, while some employees cited it as one factor behind the increasing difficulties in hiring social workers for eldercare. Thus, a causal theory suggesting that the program had created less attractive working conditions, thereby making it difficult to attract competent employees for eldercare provision, was suggested by this stakeholder group.

As to the third question concerning similarities, the three primary groups, who were selected due to their privileged knowledge and crucial role in the functioning of the program as implemented, agreed that the CL program had caused a significant standardization of the criteria on which it was decided who was supposed to get which services (causal theory), and that this represented an improvement (normative theory) compared to previous practices with a more casual and less transparent decision-making process (theory of previous situation). Thus, these stakeholders agreed that an advantageous discipline concerning individual decisions on entitlements to services had been achieved by the CL program. They also agreed that a certain danger of mechanical inflexible eldercare was associated with the program. The dynamic changing needs of elderly users were not adequately captured in the current version of the CL according to the three groups.

In their analysis of dissimilarities, disagreements, and similarities, the evaluators may draw upon *social science theories*. Which theories from scientific discourse would be relevant to apply? We used the embedded intervention theory as a vehicle for analogical reasoning and came up with various ideas to be explored. Looking at Figure 2, it is immediately obvious that it is an information system that can

be used for documentation and management control (Anthony & Govindarajan, 2001; Swiss, 1991). Many social workers believed that the system represented a control regime that symbolized upper level decision makers' lack of trust in their work and enforced a rigid standardized work ethic unsuitable to the unique and changing needs of elderly users. These concerns of the employees were partly shared by interest organizations representing the elderly. Another basic question concerning such systems is whether they produce reliable and relevant information or whether the information is distorted in the process (Swiss, 1991) and thus represent a GIGO (Garbage In Garbage Out) problem. There was a strong conflict dimension among the stakeholder theories related to the GIGO problem. In addition, the CL intervention is a clear-cut standardization effort. A classic problem in public administration, standardization offers advantages and disadvantages to different stakeholders. Street-level operators lose some of their maneuvering space to adapt services to individual cases, while upper-level decision makers gain in transparency and oversight. In addition, the *standardization–individualization dimension* represents an important conflict dimension in this program.

Summing Up: What is Special About TSE?

The starting point of our proposed model is a situation in which strong dissension and political conflicts prevail around the intervention to be evaluated. In such situations, we argue that the construction of one unitary intervention theory is not the best way to use the advantages of the TBE. Instead, it becomes important to distinguish between different theories. We have suggested the TSE model and five steps for its application as a feasible alternative to the unitary approach in such situations. Instead of trying to reconcile disagreements and conflicts, the evaluator must compare, clarify, and separate the different viewpoints of primary stakeholder groups based on the theory-based approach outlined above.

Apart from the recommendation to reconstruct a whole set of stakeholder intervention theories, our theory-based stakeholder model exhibits several other features. We advise that intervention theories be structured according to the *tripartite scheme* of situation theory, causal theory, and normative theory. We also suggest that the *extended system model* could be a useful classificatory device, that is, the input–conversion–output–outcome model. Furthermore, we propose that evaluators *insert pertinent actors* into the extended system model to give it the following appearance: input–administrative agencies–conversion–output–recipients–outcome. The tripartite scheme and the extended system model are supposed to be used by the evaluator as analytical tools to add structure, order, and clarity to the original raw intervention theories. By themselves, they are not intervention theories; they are instruments to be used in the building of intervention theories.

Finally, as a methodological tool in the eliciting of raw intervention theories, we have suggested the *principle of reason*. This principle is used to elicit and reconstruct tacit or vaguely intimated ingredients (gaps and omissions) of the raw intervention theories given their premises and situational understandings.

Discussion

In her 1997 article, Carol Weiss discusses 12 problems for TBE to make more progress (Weiss, 1997). The arguments expounded in our article may be related to the first four problems Weiss identified: That a program theory sometimes is unclear, that there may be confusion about the components of the theory, that the articulated theory is not the only possible one, and that it may be difficult to identify or construct the theory. We believe that a great deal of the confusion related to these four problems is caused by the fact that interventions are often political in nature and that different groups of stakeholders often disagree about both the process of implementation and the mechanisms by which certain outcomes will be achieved (what we have called the causal part of the

intervention theory), value criteria proposed (what we have called the normative theory), and the nature of the problems to be attacked (what we have called the situation theory). So far, the TBEs have coped with these problems using different types of deliberation and process consultant techniques to arrive at some negotiated agreement about one unitary intervention theory.

However, in some types of evaluations, this strategy is not possible because of substantive and multilevel complexities and political conflicts inherent in the intervention system. Even if possible, such a strategy may be problematic. If practiced in the context of an intervention with strong tensions, it may enforce a superficial unity around the viewpoints of those in top of the hierarchy (Hanberger & Schild, 2004).

The basic purpose of the TSE approach is to reduce the "designed blindness" to political conflict that we believe has characterized some work within this tradition. However, there is no such thing as a free lunch. The TSE model adds complexity to the task of the evaluator and is more resource demanding. It is demanding because it is imperative that the evaluator answers the question of whose theory she is interpreting and because there is more than one theory to elicit. Thus, there is a trade-off to be considered.

One of the difficulties in practicing TSE is to determine the most significant lines of conflict associated with a specific intervention. These lines of conflict are related to differences in interests, values, perceptions, and experiences often caused by a "where you stand depends on where you sit" mechanism. Often this mechanism is triggered by different positions in the formal social structure related to an intervention, as in the example provided in this article. In our Danish CL example, differences in professional training and organizational tasks triggered different perspectives and conflicts. However, these lines of conflict may not always be the important ones in social interventions. Depending on type of intervention, conflict lines may follow other social cleavages, such as different ethnic or religious affiliations. In multinational interventions, we may envision cleavages caused by different national cultures. Thus, the fundamental challenge to the evaluator in the TSE model is to analyze and understand the most important conflicts embedded in an intervention.

What advantages does the TSE model offer in return for the added complexity of the tasks of the evaluator? In our opinion, there are especially three potential advantages that may be realized by adopting the TSE model:

1. Benefits to decision makers: When the various stakeholder intervention theories are finely grained, the evaluation can track each link in these chains of assumptions. The findings of such an evaluation will show which links are well supported by stakeholder views, which links are weakly supported, and where in the chain the intervention realization stream seems shaky. This information from stakeholders with privileged knowledge about implementation processes, output, and mechanisms shaping outcome effects is important because, if deemed reliable, it enables responsible decision makers to make more relevant and better decisions.

2. Benefits to other stakeholders: Stakeholders other than leaders may gain knowledge, inspiration, and understanding of each others' positions from TSE. Some focal stakeholders may know little about other stakeholders' roles in and views on the intervention system. Their perceptions of the system may be narrow and very selective in some negative sense. They may also nurture beliefs that serve defensive functions, a situation that Friedman (2001) calls designed blindness. These stakeholders may learn and widen their horizons from being informed about intervention presuppositions of others. They may acquire some understanding and even appreciation of each other's perspectives. They may learn new skills, new interpretations of goals, new and different value criteria, and new interpretations of chains of causality.

3. Benefits to democracy: By keeping the various stakeholder intervention theories separate not only their similarities but also their differences are brought to light. In this way, many opinions are brought to bear on the intervention system in various stages and at various levels, and

pluralism is enhanced. From a democratic point of view, this may be an advantage. To look for open and potential clashes of opinion is reasonable from a democratic point of view. The rationality behind this is sometimes called communicative rationality or deliberative democracy. This means that findings from an evaluation are used in talks, conversations, dialogues, argumentation, and deliberations. The various intervention theories are assessed and counterassessed in conversations and discussions. The practical uses of information on diverse stakeholder intervention theories are constructed and reconstructed in interactive processes of arguing, presenting opinions, supporting some viewpoints, and opposing others (Albaek, 1995; Majone, 1989; Valovirta, 2002).

Conclusion

This article has introduced a new approach to TBE. Entailing a fusion of TBE and stakeholder model ideas, we call it theory-based stakeholder evaluation, or the TSE model for short. Theory-based approaches to evaluation have made significant contributions to current evaluation practice but have been based on a problem-solving model, which is often inadequate because it pays lip service to the politics of evaluation. We have argued that it is important that the theory-based approach is extended to contexts characterized by dissension and conflict. The TSE model, we have argued, represents such an extension and adds important clarifications of former approaches.

It is designed to be useful in contexts ridden with strong dissension and political conflicts around the intervention to be evaluated. Such situations are more likely in situations with large multilevel and multisite interventions such as state-level, federal-, or EU-level or even global programs to combat for instance climate change. However, even in local interventions, tensions and disagreements may be impossible to solve for even the best process evaluator.

Of course, the TSE model is no panacea and there are unresolved problems associated with it. However, hopefully, it will be of use for future attempts to practice TBE in political contexts.

Note

1. The definition is elaborated from Vedung (1997, pp. 138ff, 224–227), who uses the term "intervention theory." Its definition and application has been further elaborated in Hansen and Vedung (Hansen & Vedung, 2005, pp. 147–185).

References

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

Anthony, R. N., & Govindarajan, V. (2001). *Management control systems.* New York, NY: McGraw-Hill International.

Bickman, L. (1987). *Using program theory in evaluation. New directions for program evaluation (Series no.33).* San Francisco, CA: Jossey Bass.

Birckmayer, J. D., & Weiss, C. H. (2000). Theory-based evaluation in practice—What do we learn? *Evaluation Review*, *24*, 407-431.

Bledsoe, K. L., & Graham, J. A. (2005). The use of multiple evaluation approaches in program evaluation. *American Journal of Evaluation*, *26*, 302-319.

Chen, H. T. (1990). *Theory-driven evaluations.* Newbury Park, CA: SAGE.

Chen, H. T. (2005a). *Practical program evaluation. Assessing and improving planning, implementation, and effectiveness.* Thousand Oaks, CA: SAGE.

Chen, H. T. (2005b). Program theory. In S. Mathison (Ed.), *Encyclopedia of evaluation* (pp. 340-342). Thousand Oaks, CA: SAGE.

Chen, H. T. (2005c). Theory-driven evaluation. In S. Mathison (Ed.), *Encyclopedia of evaluation* (pp. 415-419). Thousand Oaks, CA: SAGE.

Dahler-Larsen, P. (2001). From programme theory to constructivism: On tragic, magic and competing programmes. *Evaluation*, *7*, 331-349.

Davidson, D. (2006). Radical interpretation. In D. Davidson (Ed.), *The essential Davidson* (pp. 184-195). Oxford, England: Oxford University Press.

Donaldson, S. I. (2007). *Program theory-driven evaluation science: Strategies and applications.* Mahwah, NJ: Lawrence Erlbaum.

Dunleavy, P., Margetts, H., Bastow, S., & Tinkler, J. (2006). New public management is dead—long live digital-era governance. *Journal of Public Administration Research and Theory*, *16*, 467-494.

FitzGibbon, C. T., & Morris, L. L. (1975). Theory-based evaluation. *Evaluation Comment*, *5*, 1-4.

Friedman, V. J. (2001). Designed blindness: An action science perspective on program theory evaluation. *American Journal of Evaluation*, *22*, 161-181.

Galtung, J., & Naess, A. (1955). *Gandhis politiske etikk.* Oslo, Norway: Universitetsforlaget.

Hanberger, A., & Schild, I. (2004). Strategies to evaluate a university-industry knowledge-exchange programme. *Evaluation*, *10*, 475-492.

Hansen, M. B., & Vedung, E. (2005). *Fælles sprog i ældreplejens organisering. Evaluering af et standardiseret kategorisystem* [Common Language in the organization of elder care. Evaluation of a standardized classification system]. Odense, Denmark: Syddansk Universitetsforlag.

Hoogerwerf, A. (1990). Reconstructing policy theory. *Evaluation and Program Planning*, *13*, 285-291.

Leeuw, F. L. (1991). Policy theories, knowledge utilization, and evaluation. *Knowledge and Policy*, *4*, 73-92.

Leeuw, F. L. (2003). Reconstructing program theories: Methods available and problems to be solved. *American Journal of Evaluation*, *24*, 5-20.

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

McLaughlin, J. A., & Jordan, G. B. (1999). Logic models: A tool for telling your program's performance story. *Evaluation and Program Planning*, *22*, 65-72.

Molas-Gallart, J., & Davies, A. (2006). Toward theory-led evaluation—The experience of European science, technology, and innovation policies. *American Journal of Evaluation*, *27*, 64-82.

Popper, K. R. (1985). The rationality principle. In D. Miller (Ed.), *Popper selections* (pp. 357-365). Princeton, NJ: Princeton University Press.

Quine, W. V. (1960). *Word and object.* Cambridge, MA: MIT Press.

Rogers, P. J. (2007). Theory-based evaluation: Reflections ten years on. *New Directions for Evaluation*, *114*, 63-67.

Rossi, P. H., Freeman, H. E., & Lipsey, M. W. (1999). *Evaluation: A systematic approach* (6 ed.). Thousand Oaks, CA: SAGE.

Rossi, P. H., Freeman, H. E., & Lipsey, M. W. (2004). *Evaluation: A systematic approach* (7 ed.). Thousand Oaks, CA: SAGE.

Sidani, S., & Sechrest, L. (1999). Putting program theory into operation. *American Journal of Evaluation*, *20*, 227-238.

Stame, N. (2004). Theory-based evaluation and varieties of complexity. *Evaluation*, *10*, 58-76.

Suchman, E. A. (1967). *Evaluative research: Principles and practice in public service and social action programs.* New York, NY: Russell Sage.

Swiss, J. E. (1991). *Public management systems: Monitoring and managing government performance.* Englewood Cliffs, NJ: Prentice Hall.

Turnbull, B. (2002). Program theory building: A strategy for deriving cumulative evaluation knowledge. *American Journal of Evaluation*, *23*, 275-290.

Vaessen, J. (2006). Programme theory evaluation, multicriteria decision aid and stakeholder values: a methodological framework. *Evaluation*, *12*, 397-417.

Valovirta, V. (2002). Evaluation utilization as argumentation. *Evaluation*, *8*, 60-80.

Vedung, E. (1997). *Public policy and program evaluation.* New Brunswick, NJ: Transaction Publishing.

Weiss, C. H. (1997). How can theory-based evaluation make greater headway? *Evaluation Review*, *21*, 501-524.

Weiss, C. H. (2007). Theory-based evaluation: Past, present, and future. *New Directions for Evaluation*, *114*, 68-81.

Worthen, B. R. (1996). The origins of theory-based evaluation. *Evaluation Practice*, *17*, 169-172.