

*Evert Vedung*  
Uppsala University Institute  
for Housing and Urban  
Research

*Morten Balle Hansen*  
Department of Political  
Science, Aalborg University

*Pekka T Kettunen*  
University of Jyväskylä

## Introduction to the special issue on evaluation

In recent times, there has been a remarkable expansion in the evaluation of public interventions, particularly concerning intervention processes but also intervention impacts. Along with New Public Management and its ideas of augmenting the use of Market Type Mechanisms (MTMs) and pushing for improved political and administrative leadership, evaluation – included in a broad governance doctrine alternatively called Performance Management and Results-oriented Management - has been one of the most important features of public sector reforms of liberal democracies since the 1980s (Hood 1991; Leeuw & Furubo 2008, Vedung 2010, Pollitt & Bouckaert 2011). Another more recent driver for evaluation is the evidence movement. At the same time, criticism of evaluation from various points of view has put new evaluative approaches on the agenda (Nordic examples include Hansen & Vedung, 2010, Karlsson 1995, 1996, 2001 and Krogstrup 1997).

Evaluation is practised by public authorities themselves, by watch-dog agencies, by private consultants and by academics. Evaluation findings and conclusions may be used for many purposes. The most obvious ones are accountability and improvement. One of the elementary rules of representative democracy is that implementers of public programmes, the agents, are accountable to those who make decisions concerning the adoption and amelioration of the programmes, the principals. The traditional forms of monitoring, financial accounting and

---

**Evert Vedung** is emeritus professor of political science, especially housing policy with office space at Uppsala University's Institute for Housing and Urban Research. He serves as a visiting professor (part-time) at Linnaeus University's Dept of Social Work, and as evaluation advisor (part-time) at Samordningsförbundet Västra Skaraborg. His works in English on evaluation include *Public Policy and Program Evaluation* (1997, 2010, author) and *Carrots, Sticks and Sermons* (1998, 2003, co-editor).

**Morten Balle Hansen**, PhD. Professor. Works at the department of Political Science, Aalborg University, Denmark and is a leader of the research group for Public Administration which is related to the Centre for Organization, Management and Administration (COMA). Has been doing research in the management of public sector organizations since the early 1990ies and published in journals such as *American Journal of Evaluation*, *Public Administration*, *Public Management Review* and *International Review of the Administrative Sciences*. Combines research in leadership behavior in public organizations with research in public policy processes in general and evaluation processes in particular. In recent years he has increasingly directed his research activities towards a comparative perspective on the dynamic relations between public management reforms, innovation processes and leadership.

**Pekka T. Kettunen** works as a professor of evaluation of social services at the University of Jyväskylä, Finland. His previous work has covered implementation and policy analysis, comparative local government research, and organisation and network analysis. He has published journal articles in e.g. *Local Government Studies*, *Regional and Federal Studies*, *European Journal of Political Research* and *Croatian Political Science Review*, and articles in edited books published by University College London, Lexington Books, Sage and Edwin Mellen Publishers. Professor Kettunen is currently working in the area of evaluation with such topics as evidence-based evaluation and the usage of evaluation findings in policy-making.

Offentlig förvaltning  
Scandinavian Journal of  
Public Administration  
16(3): 3-7

© Evert Vedung, Morten  
Balle Hansen, Pekka T  
Kettunen and Förvaltnings-

auditing, have increasingly been supplemented by policy and programme evaluation, performance review and client satisfaction surveys, thus putting pressure on the way accountability is framed and understood (Pollitt 2003). Formative in nature, the second form of evaluation use, improvement, is usually resorted to while a programme is being implemented, in concert with implementers and stakeholders. In practice, the two are interrelated. Evaluation findings can be used for the purpose of accountability, for example concerning the queuing time for social services, and at the same time for improving services, by suggesting how to make the services more user-oriented.

Given the role that evaluation findings and conclusions can play in public sector thinking and decision-making, we must note that producing these findings and drawing these conclusions is fraught with challenges. The time scale and resources of evaluation may not suffice for undertaking a comprehensive analysis and hence the question arises of whether the findings can be generalized. In addition, the question of how to acknowledge the variety of views concerning, say, the success of an intervention, may be a tricky one. Decision-makers, implementers, interest groups and clients may have differing views concerning what constitutes success in a particular intervention.

In impact evaluation the challenges increase. This is because real life is an open system in which causal relationships are difficult to pinpoint. Interregional initiatives, such as the EU Structural Fund programmes, may be argued to create jobs, but perhaps the emergence of jobs is determined by the markets, a fair assumption, and not by the programmes and projects as such. Similarly, in the area of social care we may like to claim that certain kinds of interventions are helpful in diminishing alcohol-related problems. The real reason for a potential diminishment could be, however, that individuals themselves understand the detrimental effect of over-consumption and draw their own conclusions. In many cases the public programmes have only a marginal effect, but sometimes this little input pushes an individual, be it a business executive or a social welfare recipient, in the desired direction.

In the North-Atlantic world since the 1960s, evaluation has come in waves that have left layer upon layer of sediments, which constitute latter-day evaluation realities. Starting in the mid-1990s, a wave of evidence-based policy-making and evaluation has grown in strength, at least at the level of rhetoric (Vedung 2010). The main tenets of the evidence wave are a focus on what works, i.e. causality, in the sense of which intervention produces the desired results, a preference for randomized controlled experimentation as the gold standard method to sustain such findings and systematic reviews of such findings for evaluative use in public sector activities and decisions. (Davies, Nutley & Smith 2000, Gray et al. 2009; Nordic contributions by Rieper & Hansen 2007, Hansen & Rieper 2009, 2010, and Krogstrup 2011). All of these aspects have created both enthusiasm and storms of criticism in various quarters.

In social work, for instance, efforts have been made to include non-experimental qualitative knowledge of professionals and clients as useful evidence of what works. Social work practice, it is argued, is based on a close rela-

tionship between the social worker and the client, each case representing a particular history. This alliance makes the difference, and talk, not pills, is the active agent. Some scholars even argue that clients should be put at the centre of evaluation, because they produce the best evidence of what works for their own recovery.

A related but somewhat different view of evidence can be deduced from champions of stakeholder evaluation. While far from new, stakeholder evaluation is held up as an alternative to randomized controlled trials (Greene 2000, 2005). If we think about the success of a particular programme, it is understandable that there are different views on this, and hence the evaluator ought to construct a balanced view concerning the success or to report the main grievances concerning the issue. Finally, evaluation should function not merely as a management partner but also as an independent and critical watch-dog.

This special issue presents a careful selection of papers originally presented at the Nordic Political Science Association workshop sessions on evaluation in Vaasa, Finland in August 9-12 2011. Revised, peer-reviewed and revised again, the articles focus on some of the above questions. Three of them are about theory and methodology. The other two are more empirical in nature.

How are various governance types and evaluation approaches interrelated? This is the fundamental focus of **Anders Hanberger's** article. After a thorough discussion of the issue, he proposes a framework for analysing the implications. He maintains that different governance arrangements affect the choice of evaluation models and the way they are emphasized. For example, the New Public Management model requires performance information while a network model emphasizes information concerning the functioning of the collaborative network relationships.

Evaluability assessment, argues **Peter Dahler-Larsen** in his contribution, is a tool that is supposed to help evaluators establish whether evaluation is appropriate in a given situation. Thus, evaluation is understood as a situational good. Today, however, evaluability assessment is neither much discussed nor particularly used, maintains the author. Mandatory, comprehensive and repetitive evaluation systems are gaining ground in public administration supported by general social, political and managerial norms and values, indicating that evaluation is believed to be a universal good. Can a form of evaluability assessment be revitalized in order to pave the way for a more modest, more reflexive, and more context-sensitive belief in evaluation? The author provides an updated version of evaluability assessment in the form of a list of recommendations. This is accompanied by a discussion of the limitations of such an approach.

Governance systems anchored in evaluation and focused on developing performance by comparing actors in organizational fields have been increasingly institutionalized at international as well as national levels in recent years. Systemic evaluation governance is an appropriate umbrella term for these, suggests **Hanne Foss Hansen**. Systemic (not systematic) evaluation governance takes several forms, for example, indicator systems, benchmarking, accreditation, certification and initiatives of evidence-based professional practice. On the basis

of a thorough discussion of the concept of systemic evaluation governance, she develops a typology of different forms of this specific governance type and analyses and discusses the logics and premises in which they are anchored. Her analysis is illustrated by empirical examples.

**Merete Watt Boolsen** explores how to understand an education reform and its evaluation. Her case analysis deals with an education reform in Greenland. She notes that quantitative analysis gives an insufficient explanation of the success of the reform. Qualitative analysis is also needed to provide a more nuanced and in-depth view on the way students and the Greenlandic people in general perceive education. Boolsen argues that the choice of perspective and methods of evaluation are part of a power struggle.

**Matts-Åke Belin and Per Tillgren** evaluate the impact of an extremely interesting Swedish national programme adopted in 1997 called Vision Zero – the long-run implication of which was that no one should be killed or seriously injured as a result of traffic accidents in the road transport system. An important direct policy instrument in this programme was to assign system designers (not drivers) formal responsibility for safety in accordance with Vision Zero. The finding that the programme has a low level of achievement appears to be attributable to four principal explanatory factors well known from implementation theory while other interventions at, for example, the EU level tend to provide support for the intervention.

## Editors' Note

The intellectual and technical work of manufacturing this special issue has been fairly divided between the three editors. For that reason our names are provided in alphabetical order. In deciding to place Evert Vedung's name first, we have made one conscious exception to that principle. This is to honour his lead role in initiating, planning and implementing the evaluation workshop at the Nordic Political Science Association Conference in Vaasa, Finland, without which this special issue would not have been produced (for a photo, please consult Picture 23 at <http://www2.ibf.uu.se/PERSON/evert/pictures.html>).

The three editors take final responsibility for neither contents nor language and form of the articles of this special issue. That responsibility rests entirely with the authors. For a specific disclaimer by Evert Vedung, see the Belin-Tillgren contribution.

## References

- Davies, Huw T.O, Sandra M. Nutley & Peter T. Smith (eds) (2000) *What works? Evidence-based policy and practice in public services*, Bristol, Polity Press.
- Gray, Mel, Debbie Plath & Stephen A. Webb (2009) *Evidence-based Social Work - A Critical Stance*, Routledge, London.
- Greene, Jennifer (2000) 'Challenges in practicing deliberative democratic evaluation,' *New Directions for Evaluation*, 85: 13–26.

- Greene, Jennifer (2005) 'A Conversation With Jennifer Greene,' *The Evaluation Exchange*, 11 (3), Issue Topic: Democratic Evaluation.  
<http://www.hfrp.org/evaluation/the-evaluation-exchange/issue-archive/democratic-evaluation/a-conversation-with-jennifer-greene>
- Hansen, Hanne Foss & Olaf Rieper (2009) 'The Evidence Movement. The Development and Consequences of Methodologies in Review Practice.' *Evaluation*, 15 (2): 141-163.
- Hansen, Hanne Foss & Olaf Rieper (2010) 'Institutionalization of Second-Order Evidence-Producing Organizations' in Frans L. Leeuw, Tom Ling & Olaf Rieper (eds.) *The Evidence Book. Concepts, Generation, and Use of Evidence*. Transaction Publishers, New Brunswick, NJ.
- Hansen, Morten Balle & Evert Vedung (2010) 'Theory-based stakeholder evaluation.' *American Journal of Evaluation*, 31 (3): 295-313.  
<http://evi.sagepub.com/content/16/3/263.full.pdf+html>
- Hood, Christopher (1991) 'A public management for all seasons,' *Public Administration*, 69: 3-19.
- Karlsson, Ove (1995) *Att utvärdera mot vad? Om kriterieproblemet vid intres-sentutvärdering*, Stockholm, Lärarhögskolan: HLS Förlag.
- Karlsson, Ove (1996) 'A critical dialogue in evaluation: How can the interaction between evaluation and politics be tackled?' *Evaluation* 2 (4): 405-416.
- Karlsson, Ove (2001) 'Critical dialogue: Its value and meaning,' *Evaluation* 7 (2): 211-227.
- Krogstrup, Hanne Kathrine (1997) 'User participation in quality assessment: A dialogue and learning oriented evaluation method,' *Evaluation*, 3 (2): 205-224.
- Krogstrup, Hanne Katrine (2011) *Kampen om evidens. Resultatmåling, effekt-evaluering og evidens*. Hans Reitzels Forlag, København.
- Leeuw, Frans L & Jan-Eric Furubo (2008) 'Evaluation systems: What are they and why study them?' *Evaluation*, 14 (2): 157-169.
- Pollitt, Christopher & Geert Bouckaert (2011) *Public Management Reform. Comparative Analysis of New Public Management, Governance and the Neo-Weberian State*, Oxford University Press, Oxford.
- Rieper, Olaf & Hanne Foss Hansen (2007) *Metodedebatten om evidens*. AKF Rapport, København.
- Vedung, Evert (2010) 'Four Waves of Evaluation Diffusion,' *Evaluation*, 16 (3): 263-277. <http://evi.sagepub.com/content/16/3/263.full.pdf+html>