



The Debates, Notes and Queries section provides an opportunity for evaluation-related issues to be debated as well as for more general interchange. Debates can take the form of sustained arguments by the advocates of different approaches or of briefer thoughts or notes. Contributors may also wish to comment and raise questions about material that has previously appeared in the journal or simply use the section to draw readers' attention to relevant issues, ongoing research, evaluation activities and other events.

Evaluation Activities in Europe: A Quick Scan of the Market in 1998

FRANS L. LEEUW

Netherlands Open University and Utrecht University

JACQUES TOULEMONDE

Centre for European Evaluation Expertise (C3E), France

ANDRÉ BROUWERS

Netherlands Open University

Introduction

In line with the title of the third EES Conference (Evaluation: profession, business or politics), the authors convened a workshop on 'Evaluation as a business'. About 50 participants gathered on Thursday 29 October 1998 in Rome. In this brief article we attempt to summarize the presentations and interventions in this workshop and we reflect on the present state of evaluation activities in Europe.

Monitoring Evaluation Activities: Why?

Increasing membership of the EES may be taken as evidence of the growth of evaluation activities in Europe. Many of those involved in this developing profession wonder whether we are at the beginning of a rising curve or at the brink of a mature market. Examples from other continents show interesting trends within the world of evaluation.

In the United States, General Accounting Office (GAO) reports (1987, 1998) show that the support for and use of programme evaluation in both the executive and legislative branches has 'unmistakenly declined' (Paliokas and Rist, 1998). Compared to the 1960s 'golden years' of evaluation that Rossi and Wright (1984: 332) referred to and to the 1980s, when GAO's PEMD (Programme Evaluation and Methodology Division) had its finest hours, things have changed. At the same time, evaluation seems to have survived as routinely accepted activity in many departments, states, local authorities and NGOs.

In Australia, the annual number of evaluation reports jumped from almost zero to hundreds between 1985 and 1995. The high political commitment in favour of evaluation has now fallen back, but the overall level of activity seems to be steady (MacKay, 1998).

In Europe, sustained attention has been paid to evaluation activities (Summa and Toulemonde, 1998). A recent survey in the field of socio-economic development policies shows that the evaluation market has developed remarkably during the past five years, as have the number of agents performing evaluation work (EC-DG XVI, forthcoming). Data on the Netherlands also show a strong increase in the number of (*ex post*) evaluations carried out by central government departments when the mid 1990s are compared to the late 1980s (Leeuw and Rozendal, 1994; Leeuw, 1997).

The supply of evaluation services often remains external to administrations. The number of consultancy firms active in the field of evaluation has risen sharply in all countries. Entrusting evaluation to outside evaluators, whether private consultants or universities, is seen as a guarantee of the independence of the team in relation to the authorities or principals concerned. This sector has developed in different ways: the emergence of teams linked to the academic world and the creation of departments or subsidiaries of international consultancy groups. The quantitative development of supply is also accompanied by qualitative developments generated by greater professionalism. Skills have also evolved in recent years, owing more to accumulated experience than to the specialization of evaluation service suppliers.

A glance at these contrasted trends in various parts of the globe suggests that the ups and downs of evaluation activities would deserve to be more carefully monitored, and hopefully explained.

Monitoring Evaluation Activities: How?

There are several ways to monitor growth (or decline) of evaluation activities.

A *first way* of monitoring trends is to survey the activities of commissioners, with questions like:

- To what extent do policy makers, politicians and practitioners make use of evaluations in their decisions? Research looking into the utilization of evaluations often follows this route;
- How many organizations (public, private or hybrid) are carrying out evaluations as a permanent activity?
- Why do administrations in some countries adopt evaluation earlier than

other countries? In the past, international observers have identified several waves of evaluation development, corresponding to different political demands (Derlien, 1990).

A *second way* consists in counting evaluation reports across client institutions. As far as we know, only a few databases of evaluation reports have been fully operational (e.g. in the Australian Department of Finance and the Treasury Board of Canada). This approach is probably only workable in a rather centralized evaluation system that depends heavily on a leading institution. Since evaluation activities increasingly permeate all layers of administrations and all levels of governments, evaluation databases may become unmanageable. Bibliometrics could be an interesting substitute, if corrected by surveys looking at differences in publishing behaviour across European countries.

In a *third approach*, one could gather indirect evidence about evaluation trends through repeatedly taking stock of teaching and training programmes. The number of programmes and students could be counted, as well as the evaluation content of the curricula (Is evaluation a minor/major component in the programme? Does the programme focus on evaluation theory/practice?).

In preparing the EES Conference, we have followed a *fourth way* which is to survey the activities of evaluation service providers, both private sector consultants and academic research centres (see Box 1).

One or Several Markets in Europe?

Participants in the EES workshop in Rome acknowledged that many evaluators concentrate on a given policy domain and use evaluation methods and skills that are specific to their field. This was highlighted by one respondent in the following terms:

Evaluation criteria and methods used in country-wide or sectoral evaluations of aid programmes are still far from having led to commonly accepted definitions and rules, even within the EU. . . . Methods have still to be improved but above all an ongoing process of updating and harmonising definitions and methods [is important].

Evaluation activities in Europe seem even more fragmented when we look at national (sub)markets. One can easily understand that foreign evaluators face high barriers when trying to enter into a given national market, especially for reasons of language, knowledge of institutional settings and familiarity with administrative culture. A participant in the workshop stressed that it was even hard to hire an English evaluator in Ireland, although the pool of qualified national consultants is small in the latter country.

When it comes to evaluating European-wide policies, this fragmented supply creates a problem. The European Commission can barely find enough genuinely European suppliers. Many tenders and contracts involve *ad hoc* consortia merging with national evaluators. We may be seeing the development of a new (quasi)-profession, which consists of gathering a dozen national evaluators, of making them evaluate a given European policy from various national standpoints, and in working hard to draw synthetic conclusions for use at the

Box 1 – A small survey of European evaluators

EES headquarters in Stockholm mailed 518 questionnaires to people and organizations in 15 countries (all EU member states minus Luxembourg, plus Switzerland). Addressees belonged to one or several of the following lists:

- evaluation experts having answered the call for expression of interest recently issued by the European Commission (DG XIX);
- consultants within the EES membership;
- consultants within the C3E's contact list.

To an unknown extent, the mailing list was biased in favour of evaluators dealing with socio-economic development policies.

Question 1 concerned the type of organization the respondent was affiliated with (private, not for profit, quasi-governmental). Question 2 focused on the total budget of the organization, while Question 3 dealt with the part of this budget that was directly related to carrying out evaluations. Question 4 tried to gain an insight into a trend. We asked: 'Compared to the early nineties, has this proportion been stable, increased or decreased?' Question 5 focused on what the respondents considered their four biggest markets for evaluations. 'Markets' were defined by crossing policy domains and countries. Question 6 wanted to find out how the evaluation activities in the respondent's organization were divided given these four markets, while with Question 7 we wanted to gain some insight into what share the respondents' organization has in these four markets. Finally, we added a question allowing the respondents to add 'qualitative comments' on the situation of the 'evaluation market' in Europe.

The response was very limited. Only 32 of all the questionnaires were returned. This may be partly explained by outdated addresses in the mailing list. In addition, we had neither budget nor time for undertaking (telephone) follow-up. Half of the 32 organizations responding were private ($n = 16$), the other half non-profit ($n = 8$) or quasi-governmental ($n = 3$). Missing values account for the differences that appear in this paper. For the analysis we brought together the non-profit and quasi-governmental organizations.

European level. Surprisingly, most of the big audit companies also operate in a fragmented manner, as far as we speak of evaluation which is not part of their core business.

A Fast Growing Business?

As indicated in Box 1, only 32 organizations responded. Most of them ($n = 25$) report a *growth* in the proportion of their turn-over generated by programme and

policy evaluations.¹ Only two organizations report a decline, while the five remaining organizations report that the proportion [compared to the early 1990s] is stable.² This growth is stressed by one respondent in these terms: 'evaluation is [becoming a] standard procedure in the policy process'. With regard to private sector organizations, 5 out of 16 have to deal with a decline or a stable percentage of their evaluation activities. This is only true for 1 (out of 11) non-profit/quasi-governmental organizations (see Table 1).

Table 2 shows the main 'operating fields' ranked as follows: (1) economic development and enterprises;³ (2) social services and health; (3) development, aid and transition, (4) education, schools and training and (5) technology, innovation and research. Relatively less important are the fields of security, police, justice, agriculture and the environment.

Participants in the workshop agreed that evaluation demand is developing faster at European and national level than at regional and local level. While almost all European policies are now subject to periodic evaluation, a figure of 10 percent was mentioned for the proportion of regional policies that are subject to periodic evaluations in countries like the UK, Denmark, the Netherlands and France.

The questionnaire survey cannot inform us about the size of European market(s). It cannot even provide us with an order of magnitude. An interesting figure was given by one of the workshop participants about the overall cost (internal and external) of running a systematic evaluation function. In the case of European Structural Funds in Ireland, this cost amounted to about 1 percent of the total public expenses involved (CSF Evaluation Unit, 1998).

Table 3 shows the respondents' opinion about the share of their organization in their biggest evaluation markets. Interestingly, there are only five organizations indicating that their share in either one of the markets is more than 20 percent. This gives an image of an industry that is very far from the concentration that prevails in the 'audit industry'.

Figures and opinions show a sharp growth in evaluation activities, but they should be questioned in the light of the following comments received through the questionnaire survey. '[It is difficult] to decide how to define evaluations in respect of our activities, because in a sense all of our research work involves evaluation, since it assesses the current situation and tries to make recommendations for the

Table 1. The Nature of the Organization – Private or Not Private – By Changes in Proportions of their Budgets Directly Related to Carrying Out Evaluations (Increased and Stable or Decreased); *N* = 27, missing data are not included

	<i>Stable or decreased</i>	<i>Increased</i>	<i>Total</i>
Private	5	11	16
			59.3 %
Not private	1	10	11
			40.7 %
Total	6	21	27
	22.2 %	77.8 %	100.0 %

Evaluation 5(4)

Table 2. Means (M) and Standard Deviations (SD) of Approximate Percentages of the Organizations' Evaluation Activities by Policy Domains, divided for Private and Not Private Organizations (*N* = 27, missing data are not included)

	PRIVATE (<i>n</i> = 16)		NOT PRIVATE (<i>n</i> = 11)	
	M	SD	M	SD
ECON	39.4	41.4	29.5	29.5
SOCIAL	12.8	26.1	12.3	16.6
AID	10.3	26.4	0.9	3.0
EDUC	9.4	17.3	20.0	31.9
OTHER	4.4	17.5	5.0	7.1
TECH	4.4	17.5	9.0	22.2
AGRIC	2.5	7.7	0	0
INFRA	1.9	7.5	2.7	9.0
ENVIR	1.3	5.0	2.3	4.7
SECUR	1.3	5.0	0	0

ECON: Economic development, Enterprises; EDUC: Education, Schools, Training; SOCIAL: Social Services, Health; AID: Development, Aid and Transition; TECH: Technology, Innovation, Research; INFRA: Basic Infrastructures, Transport; AGRIC: Agriculture, Fisheries, Forestry; ENVIR: Environment; SECUR: Security, Justice, Police.

Table 3. Organizations' Market Shares and Policy Domains

	1–5%	6–20%	21–100%
EDUC	2	4	0
SECUR	0	0	0
SOCIAL	5	2	1
ECON	6	4	3
AGRIC	1	0	0
ENVIR	1	0	1
INFRA	2	0	0
TECH	3	1	1
AID	1	1	0
OTHER	1	0	0

ECON: Economic development, Enterprises; EDUC: Education, Schools, Training; SOCIAL: Social Services, Health; AID: Development, Aid and Transition; TECH: Technology, Innovation, Research; INFRA: Basic Infrastructures, Transport; AGRIC: Agriculture, Fisheries, Forestry; ENVIR: Environment; SECUR: Security, Justice, Police.

future'. Several participants in the workshop agreed that many studies tend to be labelled 'evaluations' nowadays, which would have been called action-research or organizational learning in the past. Reasons for these shifts in vocabulary can be found in either management fashions or the increasing number of rules that make evaluation compulsory. Can we rely upon our perception of an evaluation development?

Although some perceptions are more 'true' than others, perceptions can have a distinct impact on the real world. 'If men define situations as real, they are real in their consequences', is an old sociological adage. The reason is that politicians, public policy makers, managers and 'patrons of evaluation' take societal perceptions about evaluations into account. This happens when they have to decide on whether or not to establish an evaluation capacity structure in their organizations, whether or not to commission (more) evaluations or whether or not to use evaluation findings. When there is a belief that evaluations are rapidly becoming a 'growth industry' within the public sector, this may cause free rider behaviour by those organizational units that not yet have established such an evaluation capacity.

Still an Infant Industry

Evaluation can be an 'infant industry', a 'growth industry' or an 'industry in decline', while perceptions about the 'industry' that prevail can differ substantially. During a recent World Bank Symposium on Evaluation Capacity Building, Picciotto (1998) referred to programme evaluation as an 'infant industry'. An infant industry in our opinion is made up of different (small) firms simultaneously entering a promising market. This picture is in line with the results of our small survey. A mature market would probably be dominated by a few big international companies, as the audit industry is at present.

An infant industry also faces problems with unstable standards, like those expressed by several respondents to our questionnaire: 'Evaluations are not yet fully accepted as feedback mechanisms'; 'The evaluation process is becoming routinized and simplified and de-skilled. This is a danger'. Top-down pressures in favour of evaluation tend to weaken the commitment of civil servants that commission evaluations. One participant of the workshop even spoke of cynicism and asked the question whether demand would keep rising or would fall if we were no longer obliged to evaluate.

Some present features of the evaluation business were pointed out as typical of the shaky rules that govern an infant industry: over-ambitious terms of reference and evaluators promising much more than they are capable of delivering, and ambiguous requirements and behaviour with respect to independence.

Moreover, much evaluation activity keeps focusing on resources instead of outcomes, on administrative processes instead of social and economic stakes. There may even be a little drift into what auditors typically do. In the UK, Mike Power (1994, 1995) referred to the 'Audit Explosion' and the 'Audit Society'. He has put forward the insight that 'auditors become second order certifiers of the *form* of systems rather than their *substance*'. He links this development to the knowledge base that underlines the work of auditors and puts forward the hypothesis that 'system auditing is attractive because it is easier to audit the system which controls a process than the outputs of that process. A drift from inspection to certification occurs when standards of control replace standards of output'.⁴

In the longer run, there may be a danger that if we keep calling almost every

Evaluation 5(4)

‘policy analysis’, ‘process description’ and ‘organizational diagnosis’ an ‘evaluation’, a *procedure-driven approach* that resembles the current ‘auditing approach’ (looking into processes instead of outcomes for example) may be on the horizon.

Helping the Evaluation Business to Become Mature

Several workshop participants took the view that evaluation activity will mature by itself through *learning by doing*. Nevertheless, we think that evaluation societies at national and European level can and should help the evaluation business to become mature. As a means of speeding up this maturation process, we can obviously think of discovering and circulating examples of good practice. This happened at least twice during the EES Conference, as shown in Box 2.

We suggest that the EES, together with national evaluation societies, should continue to investigate evaluation activities, possibly on a more permanent basis. Topical workshops could be organized in various European countries under the same heading, like ‘recording and circulating lessons from managing evaluations’. Though small, it may help.

Box 2 – Two examples of interesting business practices

These ideas have emerged in the debates during the EES Conference.

A Dutch evaluation commissioner records those consultants that fail to meet their commitments, which is taken into account in awarding future tenders.

A Belgian evaluator strengthens his own independence by having several interim reports successively validated by the client. Report 1 displays the information basis that will support the whole evaluation. Report 2 shows the analyses and conclusions. Report 3 (final) includes recommendations. Once one report is validated, it cannot be questioned again, even if the following reports include statements that are not welcome.

Notes

1. Needless to say that there is always the possibility that respondents that do not encounter ‘growth’ belong to the non-respondents.
2. For further analysis we brought together organizations with declining market shares and organizations with stable market shares for performing evaluations (labelled ‘stable or decreased’).
3. This first ranking may result from a bias in the mailing list.
4. This *love for systems* has also been found in the comparative study of the OECD’s Public Management Group. OECD-PUMA has inventorized developments and consequences of auditing in some of their member states.

References

- CSF Evaluation Unit (1998) *Review of Ongoing Evaluation Function in the Community Support Framework (CSF) for Ireland*. Dublin: Department of Finance.
- Derlien, Hans-Ulrich (1990) 'Genesis and Structure of Evaluation Efforts in Comparative Perspective', in R.C. Rist (ed.) *Program Evaluation and the Management of Government, Patterns and Prospects across Eight Nations*. New Brunswick, NJ: Transaction Publishers.
- European Commission DG XVI (forthcoming) 'Evaluation Design and Management', in *MEANS Collection: Evaluating Socio-Economic Programmes 1*. Luxembourg: OOEPEC.
- GAO (1998) *Program Evaluation: Agencies Challenged by New Demand for Information Program Results*. Washington, DC.
- Leeuw, Frans L. (1996) 'Performance Auditing, New Public Management and Performance Improvement: Questions and Answers', in *Accounting, Auditing and Accountability Journal* 9(2): 92–102.
- Leeuw, Frans L. (1997) 'Evaluatie-Onderzoek en de Nederlandse Rijksdienst Anno 1997: Een Nabeschouwing', in A.Hendriks (ed.) *Evaluatie-Onderzoek in de Rijksdienst*, pp. 161–70. Den Haag: Ministerie van Financiën.
- Leeuw, Frans L. and Piet J. Rozendal (1994) 'Policy Evaluation and the Netherlands' Government: Scope, Utilization and Organizational Learning', in Frans L. Leeuw, Ray C. Rist and Richard C. Sonnichsen (eds) *Can Governments Learn: Comparative Perspectives on Evaluation and Organizational Learning*, pp. 67–89. London: Transaction Books.
- MacKay, Keith (1998) 'Evaluation in the Australian Government', Paper, INTEVAL Working Group on Program and Policy Evaluation, Dublin, Annual Meeting, May 21–4.
- Paliokas, K. and Ray C. Rist (1998) 'The Rise and Fall (and Rise Again?) of the Evaluation Function in the US Government', Paper, INTEVAL Working Group on Program and Policy Evaluation, Dublin, Annual Meeting, May 21–4.
- Picciotto, Robert (1998) 'Why Bother About ECD?', in R. MacKay (ed.) *Public Sector Performance: the Critical Role of Evaluation*, pp. 17–21. Selected proceedings from a World Bank Seminar, Washington, DC.
- Power, Mike (1994) *The Audit Explosion*. London: Demos Publishers.
- Power, Mike (1995) 'Audit and the Decline of Inspection', Paper, Public Finance Foundation, London.
- Rossi, P. and J. D.Wright (1984) 'Evaluation Research: an Assessment', *Annual Review of Sociology* 10: 331–52.
- Summa, Hillka and Jacques Toulemonde (1998) 'Evaluation in the European Union', Paper, INTEVAL Working Group on Program and Policy Evaluation, Dublin, Annual Meeting, May 21–4.
- Taylor, M. (1996) 'Between Public and Private: Accountability in Voluntary Organisations', *Policy and Politics* 24(1): 57–72.

FRANS L. LEEUW will be Chief Review Officer of Higher Education in the Netherlands as of January 2000 and is professor of Sociology, Utrecht University. He received his PhD from Leyden University. Previously he was director of the Performance Audit and Evaluation Department of the Netherlands Court of Audit (1987–96) and associate professor Social Policy Research at Leyden University (1978–87). In 1980 he was Fulbright Scholar at the University of North

Evaluation 5(4)

Carolina at Chapel Hill. His research interests are policy theories, auditing and evaluation, higher education management and digitization of higher education. He is President-elect of the European Evaluation Society (EES). Recent articles have appeared in *Rationality and Society*, *Evaluation* and in *Accounting, Auditing & Accountability*. Please address correspondence to Open Universiteit, Postbus 2960, 6401 DL Heerlen, The Netherlands.

JACQUES TOULEMONDE is director and co-founder of the Centre for European Evaluation Expertise where he co-ordinated an extensive methodological programme on the evaluation of European Structural Funds. His specializations are: partnership evaluation, economic evaluation techniques, evaluation capacity building. He is a member of the International Evaluation Research Group Evaluation (INTEVAL) and of the Board of the European Evaluation Society. He teaches evaluation in several French universities and is the author and editor of articles and books dealing with the institutionalization and professionalization of evaluation, causality analysis and evaluation techniques. He has co-directed a collective book entitled *Politics and Practice of Intergovernmental Evaluation*.

ANDRÉ BROUWERS MSc is working at the Netherlands Open University, Department of Humanities and Social Sciences. His main interests concern the statistical area of Structural Equation Modelling, in research on idea-generating and teacher burnout. Recent publications with Welko Tomic are 'Idea Generating Among Secondary School Teachers' (in *Catalyst for Change*, 1999, 28(2), 5–7), 'A Longitudinal Study of Teacher Burnout and Perceived Self-Efficacy in Classroom Management' (in *Teaching and Teacher Education*, in press) and 'A Longitudinal Study of the Relationships between Three Burnout Dimensions among Secondary School Teachers' (in *Educational Practice and Theory*, 1999, 21).