THE FUTURE OF EVALUATION: PROSPECTS AND CHALLENGES

Reinhard Stockmann[1]

RESUMEN

No es posible predecir el futuro de la evaluación. Sin embargo, se pueden expresar algunos supuestos fundamentados con respecto a cómo se va a desarrollar, teniendo en cuenta la situación actual, y los procesos de cambio social en el futuro.

En primer lugar se muestra, mediante una serie de indicadores, que la evaluación se encuentra en un período de crecimiento histórico. Después de formular tres funciones sociales de la evaluación, se constata que el grado de institucionalización de cada una de ellas es muy distinto. En la mayoría de las organizaciones gubernamentales y no gubernamentales, la evaluación se ha establecido sólidamente como una función de gestión. En cambio, la evaluación está menos institucionalizada como herramienta de gobernanza democrática. Las estructuras más débilmente institucionalizadas se encuentran en la función "iluminadora" de la evaluación. Al respecto, para mejorarlas serían necesarias instituciones independientes, con presupuesto, que puedan actuar sin limitaciones administrativas o agendas externas.

La última parte del artículo se centra en los retos y los peligros que la evaluación debe enfrentar para que la importancia de las tres funciones de la evaluación aumente en vez de disminuir.

PALABRAS CLAVES: EVALUACIÓN; INVESTIGACIÓN SOCIAL; SOCIEDADES MODERNAS ; REFLEXIVAS; GOBERNANZA DEMOCRÁTICA ; SOSTENIBILIDAD ; LEGITIMACIÓN POLÍTICA; DESARROLLO HISTÓRICO DE LA EVALUACIÓN; FOMENTO DE LAS CAPACIDADES EN EVALUACIÓN

ABSTRACT

It is not possible to predict the future of evaluation. However, some well-founded assumptions can be made as to how evaluation is likely to develop on the basis of the current situation and the processes of social change in the future.

First, a series of indicators is used to show that evaluation is in a historic growth phase. After formulating three social functions of evaluations, it is ascertained that the extent to which these functions have been institutionalized varies considerably. Most governmental and non-governmental organizations have firmly established evaluation as a management function. As an instrument of management in democratic governance, evaluation is less strongly institutionalized. The weakest institutionalized structures are to be found when

1   Center for Evaluation (CEVAL); Saarland University, Postfach 15 11 50 D - 66041 Saarbrücken, Alemania. r.stockmann@ceval.de

it comes to the enlightenment function. What are needed here, above all, are independent institutions with their own budget, which could operate freely without administrative constraints and agendas.

The final part of the article focuses on the challenges and dangers that evaluation must respond to if the importance of the three functions of evaluation is to be increased rather than allowed to decline.

KEY WORDS: EVALUATION; SOCIAL RESEARCH; REFLEXIVE MODERN SOCIETIES ; DEMOCRATIC GOVERNANCE ;SUSTAINABILITY; POLITICAL LEGITIMACY ; HISTORICAL DEVELOPMENT OF EVALUATION; EVALUATION CAPACITY DEVELOPMENT; UTILITY OF EVALUATION; EVALUATION BOOM

INTRODUCTION

The prospects for the future of evaluation are very promising. Worldwide, the state and non-state demand is booming for evaluation. The quantity and range of evaluation services for consulting and scientific institutions that they offer as well as teaching and training opportunities, continue to grow. Worldwide! Based on the motherland of evaluation, the USA, in Europe, Latin America and Asia and more recently, also in Africa.

This also becomes evident when you take a look at the recent number of reports and studies, of which only a fraction ever are published and thus highlight only the tip of the 'Mountain of Books'. Already by the end of the last century, Carol H. Weiss (1998:10ff.) had spoken of a growing flood of studies. Ray Rist and Nicoletta Stame (2006) called their book, "From Studies to Streams", taking up this metaphor. The ever increasing demand creates an offer professionalizing itself more and more.

Evaluation has become a "booming business" (Leeuw 2009:3). Even if this situation offers great days for evaluators, it is nevertheless not a guarantee that this positive development keeps progressing this way. There are also dangers lurking for the evaluation boom. In addition, the evaluation must be flexibly adjusted to changing contexts in order to meet the needs of clients of evaluation.

This article is structured as follows: First, the attempt is made to outline the international prevalence of evaluation, before then the emphasis is put on the social functions that evaluation can fulfill in a society. Finally, the following questions are addressed: Which are the challenges and dangers evaluation is faced in the future and how will this affect the different evaluation functions.

THE CURRENT STATE OF EVALUATION

The international status of evaluation research is strongly influenced by the American motherland in theoretical and methodological terms but also with regard to topics and trends. For decades the USA is the country where the highest degree of professionalization worldwide has been achieved. There are a number of indicators of this:

In the USA evaluation is firmly anchored in institutional terms, in legislation, in the implementation of public programmes and in impact assessment. The American Evaluation Association (AEA) in the USA has the most members and certainly also the most influence. The 'Program Evaluation Standards'[2] issued by the AEA in 1989 and revised in 1994 and 2010, which were developed from the 'Standards for Evaluation of Educational Evaluations', were the force behind a large number of evaluation standards that have meanwhile been issued by other national associations worldwide. Other important efforts toward professionalization can be seen in the 'Guiding Principles for

---

2   http://www.eval.org/evaluationdocuments/progeval.html

Evaluators'[3], issued in 1995, and the lively debate on the possibilities for the certification of evaluators.[4]

The development of theoretical and methodological approaches and models in evaluation research is dominated by American authors. The training market for evaluators is also most well developed in the USA. Fitzpatrick, Sanders and Worthen (2004: 41) point out that "though there are fewer graduate programs training students in evaluation than there were in the heyday of the Great Society, the programs that continue in the United States (…) have matured into programs offering unique training opportunities – training tailored to fit the reconceptualized views of evaluation that had emerged (…)". Training programmes for evaluators have also expanded to cover the non-university sector, with many schools, state institutions, companies and different national professional associations offering such courses. There are also practical courses, pre-conference workshops, the Internet, journals and a deluge of practical guides and handbooks.[5]

As we will see later many other countries especially in Europe catch up. In general, but particularly in Europe, a *high degree of dynamism* with regard to the development of professionalization has made itself felt in the last two decades. The European Commission and its individual departments are the strongest forces working towards the

expansion and standardization of evaluation in the individual countries of Europe. Countries in which there has so far been no evaluation culture whatsoever must also gradually establish evaluation capacities in order to be able to meet the evaluation specifications tied to the implementation of EU programmes.

As a glance at the 'annual evaluation reviews' of the Directorate-General for the Budget shows, evaluations are meanwhile conducted in almost all areas of EU policy (http://ec.europa.eu/dgs/secretariat_general/evaluation/documents_en.htm). The number of evaluations has also increased considerably. In the three-year period from 1996 to 1998 the evaluation project count was 198, whilst there were almost three times as many (549) in the period from 2004 to 2006. During the following three years, the numbers remained almost constant, leading to more than 1.400 Evaluations conducted on behalf of the European Union between 2002 and 2009 (EU 2010). In 2009, about 237 evaluation projects are assembled by the European Commission, half of them with an retrospective view (interim and ex-post evaluations) and only 5 per cent prospective as ex-ante evaluations (EU 2010).

By summing up the last fifteen years, ex-ante evaluations are rare (approx. 20%) while mainly ex-post (approx. 40%) and interim evaluations (approx. 40%) have been conducted. However, during the last decade a discussion on the use of evaluation – comparable to the one in USA – approached and this also lead to considerations about an increasing usage of ex-ante evaluations.

The spread of projects financed through the Structural Funds[6] has considerably influenced the development of evaluation policy in the EU's member states in recent years (see e. g. Viñas 2009 for the case of Spain). As the statutory requirements stipulate the incorporation of evaluation in the management of projects within the Structural Funds, member states are forced to build evaluation

---

3   http://www.eval.org/publications/guidingprinciples.asp

4   Cf. in particular the contributions by Altschuld 1990; Jones & Worthen 1990; Smith 1990.

5   However, hardly anything is known about the quality of these basic and advanced training courses, since there are no studies on the results of the various forms of training. For this reason the call for such studies is loud: "(…) much more work – is needed in evaluating the outcomes of evaluator training" (Datta 2006: 420). This applies all the more to crash courses, which only last a few days and are supposed to empower programme managers to conduct evaluations themselves; a matter of some urgency, in view of the sharply rising importance of evaluations conducted 'in-house' by non-experts (cf. Datta 2006: 429).

6   Programmes to build regional infrastructure known under the term 'cohesion policy' [http://ec.europa.eu/regional_policy/thefunds/index_en.cfm]

capacities. Presumably this added to the increasing demand for professionalization of evaluation in Europe and especially in the new member states of the EU.

The number of newly founded professional evaluation associations is one indicator for the increasing importance of evaluations in the EU's new member states. There is already more than 20 national evaluation societies with memberships ranging from a dozen up to more than 700 in the case of the biggest association, the German/Austrian DeGEval Evaluation society (2012). The European Evaluation Society (EES), founded early at the beginning of the nineties, is an example for international cooperation in the field of evaluation. Evaluation societies promote the professionalization of evaluation and serve as a communication network for their members through various channels such as newsletters, press releases annual conferences and training sessions.

Summing up, it should be noted that both for most European countries and for the EU institutions,there has been a clear increase in policy and programme evaluations, that a large number of attempts to professionalise evaluation can be recognized, and that an 'evaluation culture' is spreading, a fact which led Christopher Pollitt (1998: 214) to say that "these are grand days for European evaluators".

This statement does not apply to Europe alone. With the help of a number of observations (indicators) it can be shown that the importance of evaluation is on the rise, *worldwide* (cf. Furubo, Rist & Sandahl 2002; Fitzpatrick, Sanders & Worthen 2004: 50f.; Dahler-Larsen 2006: 141ff.; Smith et al. 2011; Smith & Brandon 2011: 566; Speer, Jacob & Furubo 2013):

1) In many countries, evaluation is a fixed element in policy-making and a management control element in international organizations, national governments and their administrations and a wide range of non-profit organizations. Evaluation is also often a part of quality management or other procedures such as auditing, inspection, bench-

marking etc. Datta (2006: 420) points out that 'scientific-research-based programs and evaluations', 'evidence-based resource allocation', 'program logic models', and similar terms of our trade have become widely institutionalized for all manner of programs.

2) The number of national evaluation societies has grown considerably in recent years. According to a worldwide Internet search by Dahler-Larsen (2006: 142), the number of evaluation societies increased tenfold to 83 between 1984 and 2004. The strongest growth in recent years has been in Europe and Africa. On the official site of the global umbrella organization of the evaluation societies IOCE – 'International Organization for Cooperation in Evaluation' we find an interactive map of the world, where all, formal and informal evaluation societies and networks known by the IOCE at national and regional level in the world are listed[7]. Currently a total of 162 organizations are listed here, including 139 at national and 11 at international level. This means that also since the middle of the last century the number of associations has risen further. The 'International Organisation for Cooperation in Evaluation' (IOCE) was founded with funds from the Kellogg Foundation[8]. The IOCE sees itself as a loose worldwide amalgamation of regional and national evaluation organizations, "that collaborate to

- build evaluation leadership and capacity in developing countries
- foster the cross-fertilization of evaluation theory and practice around the world
- address international challenges in evaluation and
- assist the evaluation profession to take a more global approach to contributing to the identification and solution of world problems" (www.ioce.net).

3) The increasing demand has given rise to a *broad demand market* for evaluation, which

7   http://www.ioce.net

8   http://www.wkkf.org

is continuing to grow (cf. Leeuw, Toulemonde & Brouwers 1999: 487ff.).

The number of consulting firms concerned with evaluation has also risen sharply. Small and very small companies are in the majority here. Alongside higher education policy, development cooperation is probably – not only in Germany – the policy field most often evaluated, with the evaluation of policies and even more so that of programmes and projects coming from a long tradition and a comprehensive evaluation system having been set up (cf. Borrmann 1999; Borrmann et al. 2001; Borrmann & Stockmann 2009), and here in particular there are many individual experts. The major social research institutes, auditing companies and corporate consulting firms are only now beginning to discover the market for themselves.

Apart from consulting enterprises, there are a number of research institutions and universities active on the evaluation market and attempting to combine research and evaluation in the service of the client, basic and advanced training and communication in a fruitful way. Beyond this, institutions have been established by authorities and even ministries - such as the new evaluation institute of the German Federal Ministry for Economic Cooperation (www.evaluierungs-institut.de). They deal with the evaluations and audits in certain policy fields.

4) Above all thanks to the development of information and communication technologies and that of the World Wide Web, the *dissemination of evaluation findings* has been the subject of a tremendous surge. Even if many organizations still do not make their evaluation studies accessible to the general public, a host of findings from evaluations, audits and inspections are now already available on the Internet. Little is known about the extent to which this knowledge is used by others for shaping their own programmes or planning and conducting evaluations.

What is obvious – at least in Europe – is that the media take precious little notice of evaluation findings presented on the Internet. The fears of many government and non-government organizations that the media could latch on to negative evaluation findings and use them to their disadvantage – so that said organizations actually refrain from publishing their evaluation studies, let alone putting them on line – are therefore largely unfounded.

Having said that, if evaluation findings are picked up by the media on account of their politically explosive nature or because the topic happens to be *en vogue* (e.g. PISA)[9], "this may often be seen to have an impact which outweighs all the other efforts an evaluator has made to produce a good, respectable, and useful evaluation" (Dahler-Larsen 2006: 149; Dahler-Larsen 2011).

5) *Training activities offered worldwide* have increased sharply. While in the early years evaluations were carried out by people who were trained as educational- and social scientists or psychologists, „is not only the number of those growing who professionally and primarily deal with evaluation" (Preskill 2008:128f.; cf. Rugh 2011: 586ff.).

If in the USA – as has already been shown – there are numerous opportunities for training and advanced training, the range in Europe is also expanding. According to research carried out by Wolfgang Beywl and Katja Harich (2007: 121ff.), there are 14 university programmes in ten countries (Belgium, Denmark, Germany, England, France, Ireland, Italy, the Netherlands, Switzerland and Spain), though they do vary greatly with a view to objectives, target groups and duration.

6) The World Bank[10] began in 2001 with evaluation training courses for people working in international development cooperation and founded the 'International Development

---

9   PISA = Programme for International Student Assessment [http://www.oecd.org/pisa/]

10   Cf. Picciotto 2002 on the role of the World Bank as it relates to the funding of evaluation capacities in developing countries.

Evaluation Association' (IDEAS)[11] to support evaluation in developing countries. This aid organization, registered in England and with an office in South Africa, has declared its aim as being 'to advance and extend the practice of development evaluation by refining methods, strengthening capacity and expanding ownership'.

In January 2010 another important step regarding to strengthening the monitoring and evaluation (M&E) and performance management (PM) capacity of countries and their governments in order to achieve development outcomes was taken. The CLEAR Initiative was launched collaboratively by different donor organisations[12], with the vision to support "development anchored in evidence learning, and mutual accountability". CLEAR consists of two key components:

a) As the acronym CLEAR which stands for *Regional Centers for Learning on Evaluation and Results* indicates, such centers should be established in different regions. For this purpose five regional institutions were competitively selected to host the Centers. The Centers are established to provide in-region capacity development and technical assistance services on M&E and performance management. CLEAR supports these institutions for five years to help them achieve recognition as regional 'go to' centers for gaining technical and institutional knowledge of M&E. At last the centers should be able to to provide demand-driven and cost-effective services to each region, which includes trainings and advisory services on M&E. Additionally,

it should develop their capacities to conduct evaluations, using varied and appropriate methodologies. Five Centers were established in the different regions:

- Latin America Center: Centro de Investigación y Docencia Económicas (CIDE), Mexico D.F:, Mexico (22 applications from 7 countries, 3 institutions shortlisted)
- Francophone Africa Center: Centre Africain d'Etudes Supérieurs en Adminstration et Gestion (CESAG). Dakar, Senegal (24 applications from 11 countries, 6 institutions shortlisted)
- Anglophone Africa Center: University of Witwatersrand, Johannesburg, South Africa (56 applications from 17 countries, 5 institutions shortlisted)
- East Asia Center: Asia-Pacific Finance and Development Center (AFDC), Shanghai, China
- South Asia Center: J-PAL South Asia at the Institute for Financial and Management Research (IFMR). Chennai, India (24 applications from 6 countries, 3 institutions shortlisted)

b) CLEAR also aims at encouraging Global Learning. In doing so it offers capacity-building streams for all Centers, it supports peer-to-peer learning among the Centers particularly through a Global forum on key M&E issues.

These developments, observable worldwide, admit of the conclusion that following the years of expansion in the 1960s and 1970s in the USA and somewhat later in most countries in Europe a *second evaluation boom* which began in the 1990s can now be

---

11   http://www.ideas-int.org.

12   The African Development Bank, the Australian Agency for International Development, the Asian Development Bank, the Belgian Development Cooperation, the Inter-American Development Bank, the Rockefeller Foundation, the Swedish International Development Agency, the Swiss Agency for Development Cooperation, the UK Department for International Development, and the World Bank Group

seen, though it is far more global in nature. This second phase is no longer fuelled only by governmental (government and administration) and supra-governmental actors (in Europe the European Commission), but increasingly also by organizations of civil society.

Furubo, Rist and Sandahl (2002: 7ff.) made an attempt to characterise the evaluation culture of selected countries. They applied nine criteria, which do not, astonishing as it may be, include capacities for basic and advanced training in the respective countries! In their ranking-list, only the USA achieves in 2002 the maximum possible number of points. Canada, Australia, Sweden, the Netherlands and Great Britain follow. Even if their methodological procedure certainly does lay itself open to criticism[13], the list compiled by these authors not only underlines the position of the USA at the head of the field, but also makes it clear that many countries are trying to establish an evaluation culture and have meanwhile initiated appropriate steps toward professionalization.

In Europe, professionalization in the individual countries has progressed to very different degrees. Sweden, the Netherlands, Great Britain, Germany, Denmark, Norway, France and Finland are among those which Furubo, Rist and Sandahl (2002: 10) award a high or middling place on the ranking-list with regard to the degree of their evaluation culture.

These positions have changed dramatically during the last ten years: Speer, Jacob & Furubo (2013) published most recently an update of this table, using the same criteria for the same group of countries. Finland and Switzerland are now on the top of the ranking-list, while the USA, Canada, Australia and Sweden slightly lost ground. In general, a dramatic increase can be recognized for most countries, especially those who were behind the others ten years ago (Japan, Israel, New Zealand, Spain and Italy). Stagnations characterize the situation in the leading European countries (UK, Netherlands and Germany) instead and the gap between pioneers and the majority has been closed during the last Decade. Especially the two most important forerunners in Europe, Sweden and Germany, dropped to average (cf. Figure 1).

---

13   The most difficult problem with Furubo, Rist & Sandahl (2002) is that they form a national average although the individual sectors (disciplines) reveal major differences in the various countries. In many countries, for example, evaluations are conducted very often in university and development policy, so that an evaluation culture could be said toexist there, though this would hardly be appropriate for other policy fields. Thus individual sectors, viewed Europe-wide or even worldwide, sometimes have more similarities than different sectors in the same country (cf. Meyer & Stockmann 2006).

FIGURE 1
RANKING OF COUNTRIES ON EVALUATION CULTURE 2002 AND 2012

	I		II		III		IV		V		VI		VII		VIII		IX		2012	R	2002	R	Trend
Finland	2	2	1	2	1	2	1	2	1	2	1	1	1	2	1	2	1	2	17	1	10	12	+++
Switzerland	1	2	1	2	2	2	2	2	0	1	0	2	2	2	0	2	0	2	16	2	8	14	+++
Canada	2	2	2	2	2	2	2	2	2	2	1	1	2	2	2	2	2	2	16	3	17	2	-
United States	2	2	2	2	2	2	2	2	2	2	2	1	2	2	2	2	2	2	16	4	18	1	--
UnitedKingdom	2	2	2	2	2	2	2	2	1	2	1	1	2	2	1	2	2	1	15	5	15	6	=
Nether-lands	2	2	2	2	2	2	1	2	2	2	1	2	2	2	2	2	1	1	15	6	15	5	=
Korea	1	2	1	2	2	2	2	2	2	2	0	2	2	2	1	1	1	1	15	7	12	9	++
Sweden	2	2	2	2	2	2	1	2	2	2	1	1	2	2	2	2	2	2	15	8	16	4	-
Denmark	2	2	2	2	2	2	1	2	1	1	0	1	2	2	1	2	1	1	14	9	12	8	++
Australia	2	1	2	2	2	2	2	2	1	1	1	1	2	2	2	2	2	2	14	10	16	3	--
Norway	2	2	1	2	1	1	1	2	2	1	1	1	1	2	2	2	1	1	14	11	12	10	++
Germany	2	1	2	2	1	1	2	2	1	1	1	1	2	2	1	1	1	2	13	12	13	7	=
France	2	2	1	1	1	2	2	2	2	1	1	1	1	1	1	1	0	1	13	13	11	11	++
Japan	1	2	0	2	0	2	1	1	1	2	0	0	0	2	0	1	0	1	13	14	3	19	+++
Israel	1	1	1	2	1	1	2	2	1	1	0	1	1	2	1	1	1	1	12	15	9	13	+++
New Zealand	1	1	0	1	1	1	2	2	0	1	0	1	1	1	1	1	1	1	12	16	7	15	+++
Spain	1	1	0	2	1	1	2	2	1	1	0	1	0	1	0	0	0	1	11	17	5	18	+++
Italy	1	2	1	2	1	1	2	2	0	1	0	1	1	1	1	0	0	1	11	18	7	17	+++
Ireland	1	1	1	1	1	2	0	1	1	1	0	0	1	1	1	1	1	1	9	19	7	16	++
China		1		1		0		0		2		0		1		0		1	-		6		
Zimbabwe		1		1		0		0		1		0		0		1		0	-		4		
Total	30	32	24	35	27	32	30	36	23	27	11	20	27	33	22	27	19	26	261		223		
Mean	1,6	1,7	1,3	1,8	1,4	1,5	1,6	1,8	1,2	1,5	0,6	1,0	1,4	1,7	1,2	1,4	1,0	1,4	14		11		+++

I	Evaluation takes place in many policy domains
II	Supply of domestic evaluators in different disciplines
III	National discourse concerning evaluation
IV	Professional organizations
V	Degree of institutionalization – government
VI	Degree of institutionalization – parliament
VII	Pluralism of institutions or evaluators performing evaluations within each policy domain
VIII	Evaluation within the Supreme Audit Institution
IX	Proportion of outcome evaluations in relation to output and process evaluations
R	Ranking

Source: Furubo, Rist & Sandahl (2002: 10); Speer, Jacob & Furubo (2013)

Apparently, the field of evaluation plays an increasing role internationally. This can be proved by means of a set of indicators. Starting with the U.S., a culture of evaluation has been established in many countries. Here some small countries (Finland, Switzerland) have moved up into the top group (USA, Canada), that is, they have completed a strong catch-up development in the area of evaluation between the years 2002 and 2012. Measured by the average values of all indicators (2002: $\tilde{\chi}$=11, 2012: $\tilde{\chi}$=14), an obvious "lift effect" can be observed. Almost all countries show higher indicator values. Especially the countries in the bottom third of the ranking have caught up powerfully.

III. SOCIETAL FUNCTIONS OF EVALUATION

While initially the dissemination of evaluation in a society and its evaluation culture and its spread have been described, it is now to address the question for what *purposes evaluation* can be deployed in a society and then see the significance of these different evaluation functions in society.

First of all it has to be mentioned that *Evaluation is an invention of modernity.*[14] It is on the one hand linked to the vision of economic and social progress, the pursuit of growth and continuous improvement, and on the other hand to faith in the feasibility and controllability of social development. Evaluation offers itself both as an *instrument of enlightenment* which sheds light on development processes, and as an *instrument of control* which aims to influence those processes purposively. Above and beyond that, evaluation is also suitable for reflexive use, as an *instrument for criticism of modernity* itself. By helping with the aid of evaluation to record not only the intended impacts of interventions, but also their unintended ones, evaluation provides the empirical basis for social self-reflection.

It has meanwhile become clear that in modern societies, which are becoming more and more complex, development strategies and policies must be questioned more radically than before because of side effects which are undesired and in some cases decidedly harmful. This means that problems which up to now have only been treated as external (e.g. the environment), unintended consequences of purposively rational acts and the ability of those consequences to endure in the future (sustainability) must be integrated in the assessment to a greater extent. By these means, social action can be put on a more rational basis and the public capacity for management increased.

From these considerations, the *conclusion* may be drawn that *evaluation has never been as necessary as it is today.* Evaluation does not merely support faith in progress by simply *comparing the targets and achievements* of the desired objectives with actual statuses. By also focusing especially on *side effects and unintended impacts* in its analyses, it detaches itself from a purely technocratic view of things, questioning progress itself. Only with a *holistic perspective* and a *comprehensive impact approach* can it pay heed to the sustainability of the solutions implemented.

Threefold purpose of evaluation

From the previous remarks, it becomes clear that evaluation can be carried on (1) in the service of *social enlightenment* (cf. fig. 2). In this case it is primarily a question of assessing political strategies, programmes and measures with the instrument of evaluation to see whether or not they make a contribution to solving social problems. Creating transparency as regards the objectives and impacts of such strategies and measures enables assessments to be made on a rational basis. For example, by disclosing which political objectives are being achieved and which neglected, who benefits from such measures and who does not, which problems are solved and what risks are associated with solving them, evaluation can trigger public discussion. By doing so, it opens

---

14   For the relationship between evaluation and modernity see especially Dahler-Larsen 2011

up the possibility "to help society shape its own future in a qualified way through systematic, data-based feed-back. A society which seeks evaluation is one which prefers rational thought and critical inquiry to tradition, ideology, and prejudice" (Dahler-Larsen 2006: 143).

Be that as it may, evaluation must render its *assessment criteria transparent* in order not to be exposed to accusations along the lines of only having adopted the perspective of the political elites and decision-makers. Evaluation findings should be discussed in the *public sphere*, i.e. the central institution in which modern societies guarantee the exchange between the state and its citizens. Making evaluation findings accessible to the general public stimulates the debate about social problems and the proposed political solutions. Only if the assessment criteria are identified can evaluation promote an objective discourse, defuse ideologically motivated conflicts and contribute by means of solution-oriented recommendations to a consensus-oriented conclusion.

Evaluation findings are always *assessive judgements*. It is not until the criteria applied have been disclosed that the rationale can be seen in the judgements made in an evaluation; only then does the possibility manifest itself of arriving at other assessments by applying other criteria. It is not the evaluation findings based on systematically gathered data on specified aspects which represent a *subjective value judgement*, but the assessment criteria stipulated in advance, and at the end of the day that judgement cannot be objectified. As long as the assessment criteria in the public discourse about its findings are made transparent, evaluation contributes to separating these interest-guided value judgements from the objective realm of facts and thus making them accessible to social discussion.

By observing and assessing public action and rendering it transparent with the aid of its concepts and procedures, evaluation assumes a *social enlightenment* function which is similar to that of journalism. Eleanor Chelimsky (2008: 33), for many years director of the Institute for Program Evaluation at the

U.S. General Accounting Office (now known as the Government Accountability Office) and thus familiar with the system of politics and evaluation from the inside, formulates the special merit of evaluation thus: 'its spirit of scepticism and willingness to embrace dissent help keep the government honest'. By disseminating evaluation findings, it enhances the degree to which the public is kept informed about government action, but also about the activities of *civil society* with its many different non-government organizations (NGOs). It is only through the independent examination of the effectiveness and problem-solving competence of government programmes and measures that civil society is empowered to express competent criticism and elaborate alternative proposals for solutions.

Evaluation is not only part of society's control of the state, but also (2) an essential element of *democratic governance*. Evaluation is used on the one hand by the *legislatures,* being made compulsory in laws and ordinances for certain purposes and accordingly having to be implemented by the executive agencies. In other words the legislators use evaluation as a means of keeping an eye on the impacts of executive measures and thus enabling themselves to make objective judgements in further developing legal framework conditions in the parliaments and their subordinate (e.g. specialist) committees. Both the juridically fixed framework conditions, i.e. the extent of the obligation to evaluate, and the scope and type of the prescribed evaluations vary from country to country and are subject to change over time. In general it can be said that in the last twenty years in particular there has, in all modern societies, been a clear increase in the number of public evaluation assignments and the degree to which they are binding.

Not only the *legislatures* however, but also the *executive*, in other words the government and its ministries and public administration, are using evaluations more now. If these public institutions use evaluation to prove that they are achieving their set objectives (effectiveness), what impacts (including the unintended ones)

have been triggered, what the ratio of cost to benefit is (efficiency) etc., the *credibility and legitimacy of policy* can be improved. If clear and logical reasons can be provided as to why certain programmes are being discontinued, cut or expanded, the acceptance of decisions, or at least people's understanding of them, increases. At the same time, the disclosure of the difficulties associated with political measures and a knowledge of correlations and the impacts caused by political strategies also promote the readiness of civil society to take part actively in solving these problems and support the government with contributions of their own for the good of all.

However, a prerequisite for this is that the evaluation findings be used as a rational basis for political decisions. Donald Campbell (1969) picked up on this idea in his concept of the *"experimenting society",* in which a kind of "work sharing" between evaluation and political decision-making is propagated. According to that concept, the rational knowledge gained in evaluations should be translated directly and quasi automatically into political decisions. This form of link between evaluation and politics has been harshly criticized as a reduction of political issues to technical ones and referred to as "social engineering." Not only that, but studies showed early on "that the official political machinery did not actually behave according to the assumed rationalistic model" (Dahler-Larsen 2006: 143). The use of evaluation findings is a complex social and political process, which should be further rationalized in organizations for example by the introduction of knowledge management systems (cf. for example Becerra-Fernandez & Leidner 2008; Haun 2005; Amelingmeyer 2004; Götz & Schmid 2004; Winkler 2004; Willke 2004; Ipe 2003; Alvesson & Karreman 2001 for an overview). Conducting evaluations is without doubt not an adequate condition for *rational politics*, but it is at least a necessary one: unless the results achieved by governmental and administrative action are disclosed, it is very difficult indeed to form a democratic opinion on the basis of assessments which are really rational.

Evaluation can make a contribution not only to social enlightenment and to strengthening democratic participation and governance, but also (3) to improving the *manageability* of individual measures, programmes, organizations or even entire policy networks. The integration of evaluation in project management, for example by 'logic models' and the concept of 'project cycle management', already has a certain tradition in the modern industrial countries. In the past two decades, during the introduction of new management models in new public management and the establishment of far-reaching quality management models, evaluations have advanced increasingly to become an integral part of organizational structure and culture and of the work processes in organizations. Lastly, network management concepts have been under discussion for several years. Through the active involvement of civil society actors, these seek to establish a kind of 'institutional control' ('governance') as a complement to or perhaps even a replacement for government action.

Since evaluation can be organizationally integrated in *'feed-back' loops,* acquired knowledge, for example about the development and effectiveness of programmes, has again and again had an influence on its management. That knowledge can consistently support programme control in all phases of the political process and thus open up potential for learning. Thus the readiness and ability to integrate evaluation in the management structures of an organization have meanwhile become characteristics of modern organizations and a key to legitimacy, status, recognition, and sometimes funding (cf. Dahler-Larsen 2006: 147).

In Figure 2 the *three main purposes of evaluation* are presented again in an overview. It becomes clear that the three fields of deployment are closely connected. These three perspectives are not to be understood as exclusive, but they do reflect different opinions and philosophies, some of which lead into fundamental discussions. Whether the point of evaluation is to be seen in the generation

of knowledge, the further development of institutions or the maximization of the impacts of public programmes, remains a matter of debate between evaluators (cf. Chelimsky 2008: 33f.). Here, the view is taken that evaluation by no means serves only a single purpose, but rather a variety of objectives, subsumed here under the three aspects of *democratic enlightenment, procurement of legitimacy* for policy and *control* of politics (by means of programmes, projects and measures).

FIGURE 2
THREEFOLD PURPOSE OF EVALUATION
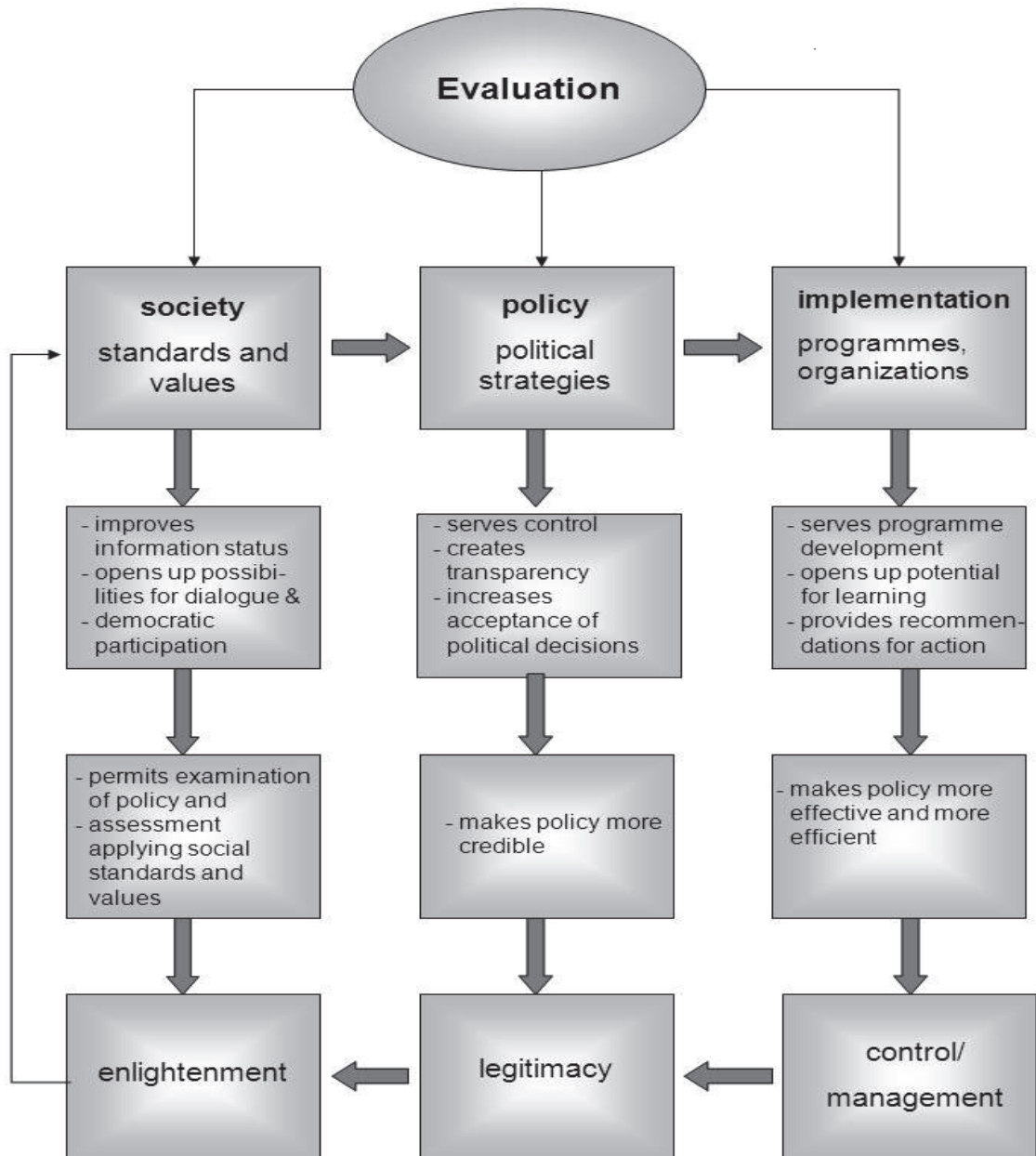
Source: © **Prof. Dr.** Reinhard Stockmann

INSTITUTIONALIZATION OF EVALUATION

In order for evaluation to be able to develop its functions as an instrument of enlightenment, legitimacy and control, evaluation capacities are necessary in various fields:

(1) to do justice to their *enlightenment function*, evaluation capacities need to be established *in society* such as can act as independently as possible of clients and entities which administer funding. Social enlightenment by means of evaluation can only be brought about purposively if the latter is not always straight-jacketed by clients' wishes. For this reason, on the one hand, *independent institutes* are *necessary*, that can decide freely where they consider social evaluation to be needed and what it is they wish to evaluate. Audit offices with a mandate of this kind can exercise such a function. However, institutions bound to certain policy fields, for example for ensuring quality in schools or quality of research services, or for examining the effectiveness and efficiency of job-market policy or that of development cooperation, can also bring about social enlightenment in the policy fields for which they were founded, but only if they are at the same time given a mandate which affords them access to the research object (for example to state, federal, communal or even EU programmes or statutory provisions). The idea of founding an international 'impact evaluation club' of the main donor countries in development cooperation in order to examine their effectiveness is to be placed in this category. Just as the German Evaluation Institute, which was founded in 2012 by the German Federal Ministry of Economic Cooperation and Development (BMZ), in order to assess the success of the German development cooperation measures *independently* (www. deval.org) (cf. Stockmann 2012b).

The provision of research funds can also help in order to strengthen the independence and credibility of evaluation, which is crucial for the enlightenment function of evaluation.

To ensure that this purpose can be fulfilled, funds would be necessary from which financial support could be applied not only for fundamental research projects but also for evaluation. Evaluation research, which has something of a strained relationship with pure or disciplinary research has a tough time in the research landscape. Its tasks and the topics it covers are often perceived in the world of science as clients' wishes, for the fulfilment of which said clients ought to pay. This impedes not only the theoretical and methodological progress of evaluation research – since the goals of a client with a very specific cognitive interest will hardly be concerned with that – but of course also its role as an instrument of enlightenment.

(2) Evaluation capacities are also necessary to support *democratic governance* and (3) *program management*, so that governments and their subordinate authorities can examine the implementation of their own strategies and policies and accompany them evaluatively. Evaluations can – as has been explained – serve not only to *increase political legitimacy,* but also to *improve control potential,* in order to make the work of implementing organizations and programme managers more efficient and more effective.

This is a challenge not only for the government sector, but also, in view of its increasing involvement in policy networks, for the non-profit sector as a whole. Non-profit organizations in particular, which work with private donations, ought to have a particular interest in making full use of their control potential and proving, by means of evaluations, that they are doing effective work and that the donations they have received have been well invested.

Recently, not only government institutions (e.g. ministries, authorities and administrative institutions of all kinds) but also private non-profit organizations (e.g. foundations, associations, clubs and relief organizations) have been seen using control and quality assurance instruments which were originally developed in the corporate

sector of profit organizations. Whilst some instruments such as financial controlling can be transposed relatively easily, others run into great difficulties when used in non-profit organizations (cf. Stockmann 2006 and Kuhlmann et al. 2004 for a summary). This is mainly because of the situative and organizational differences between profit and non-profit organizations (including government institutions). As these new management concepts become established, the non-profit sector is the very place where evaluation is becoming an indispensable instrument, providing and assessing the data necessary for management decisions.

In order to be able to fulfil these management tasks, both *internal evaluation capacities*, also in the funding and implementing organizations, and *external evaluation capacities*, in the form of scientific institutes, private companies and individual experts, are necessary (cf. Figure 3).

External evaluation capacities tend to be used more for independent analysis in the service of social enlightenment and democratic legitimation, whilst both internal and external evaluation capacities can be deployed for programme control.

The *multi-functional use* of evaluation capacities makes it clear that there is no perfect allocation to the evaluation purposes defined here. External evaluation institutions can be called upon for all three purposes. The general view is that the more independent they are, the more credible their contribution to social enlightenment, democratic procurement of legitimacy and programme control will be. Internal evaluation institutions seldom contribute to social enlightenment – though courts of auditors can do so – and they are, on account of their restricted credibility, but little used in the portrayal of the legitimacy of government or non-government organizations which implement programmes; instead, they mainly serve the internal control of projects and programmes and sometimes also of policies, and the shaping of organizational quality and knowledge management.

FIGURE 3
RELATIONSHIP BETWEEN EVALUATION PURPOSES AND CAPACITIES
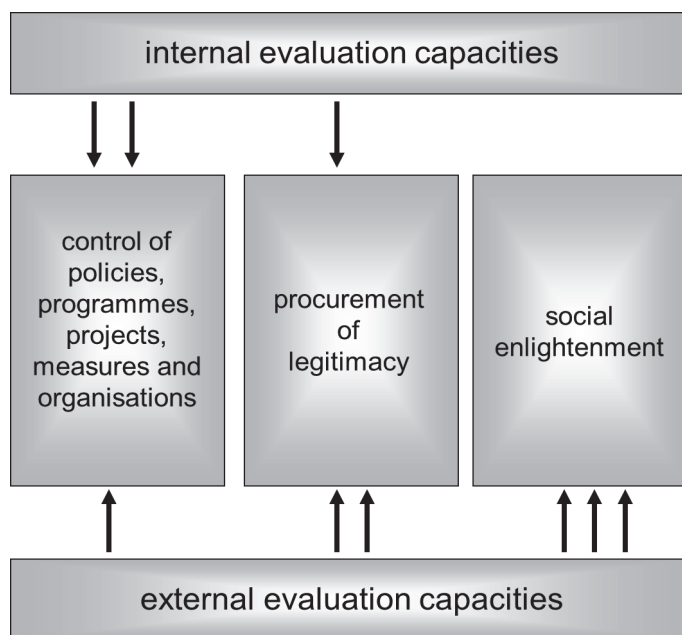
© *Source:* Prof. Dr. Reinhard Stockmann

If we visualize once again the development and current situation of evaluation and the challenges it faces against the backdrop of its role in society, it becomes clear that evaluation is widespread and that it is being used routinely more and more as an instrument of control for the *assessment of programmes, projects and measures* and as part of the *control and management system in organizations*. Evaluations are found much less often in the context of democratic governance for increasing the *legitimacy and credibility* of policy. Evaluations which cover such broad, whole policy fields (higher education policy, school quality, job-market policy) or sections of such fields are unusual. Evaluation is currently least able to do justice to the aspiration of making a contribution to *social enlightenment*. If evaluation restricts itself primarily to contract research – as is currently the case worldwide – this means, in principle, that anything which government or non-government institutions see fit to evaluate can be evaluated. There is thus no assurance that that which is necessary from a social point of view will also be evaluated.

IV. NEW CHALLENGES AND DANGERS

If we are looking next to the challenges evaluation is faced in the future we have to take in mind that evaluation is clearly policy driven. Both the first boom in the 1960s and 70s in the USA and somewhat later in most countries in Europe, and the second boom which began in the 1990s were triggered by increased state demand. Not only the boom years, but also the evaluation doldrums of the 1980s which were to be seen, at least in Europe, being attributable to policy, i.e. to considerably reduced demand.

It is not a big surprise that the development of evaluation depends on governmental – and nowadays also non-governmental – clients, because evaluation is an applied social science which is supposed to make a contribution to solving social problems. Accordingly, the demand for it increases when there is a greater need for problem solving, for example when a large number of reform programs are initiated, when selection decisions have to be made, due to scarce budget funds or when New Public Management approaches require output and impact orientation (cf. Schedler u. Proeller 2003: 62f.).

In other words the questions of whether evaluation is to take place or not, whether or not the market for it will grow, stagnate or shrink and even of what subjects are covered by it, are, to a great extent, politically influenced, that is to say influenced by the willingness of clients to deploy funds for evaluation.

However, if evaluation is also to fulfil the purpose of making a contribution to *social enlightenment*, this situation certainly does represent a problem, for "evaluation will tend to take place where money flows rather than where there is a societal need for evaluation" (Dahler-Larsen 2006: 148). This means that "there is no guarantee, however, that important areas of life in society which need evaluation are actually evaluated" (ibid.) Of course evaluation carried out on behalf of government or non-government actors can also contribute to social enlightenment, but there is simply no guarantee for this, for there is of course – as has already been mentioned several times – no compulsion for clients to render their findings public or investigate problem areas which have a high degree of social relevance.

If you want to take a look *which functions* of evaluation *will be important in the future*, and in what the clients will be interested, it is important to know the change of the societal framework conditions, because the development of evaluation is dependent on changes at the global, national and local levels.

I would like to mention just a few of these *societal context factors*:

1) Firstly, as mentioned at the beginning, it is not only the growth and shrinkage phases of evaluation that are politically determined, but also the focus on themes. This is to say, to a great extent the clients of evaluations influence which topics are worked on by evaluations and hence, which evaluation functions are booming. While during recent decades, it has mainly been the management

function of evaluation that has been in the foreground and still is of a great importance for triggering internal learning processes and enabling rationally established management processes, in recent years there has been a powerful trend towards accountability (cf. Vedung 2010: 263ff.). Perhaps it is caused by a severe shortage of financial means for social and political programs. It is above all governmental organisations, but also non-profit organisations, that increasingly want methodically convincing proof of the effectiveness and impact of their measures and political strategies (cf. Rugh 2011: 591). The huge demand for randomized control trials, rigorous evaluations and robust methods of evaluation are a clear indicator of this development, which is likely to strengthen the legitimacy function of evaluation, in the terminology used here.

2) Secondly, the increasing importance of global issues (for example climate change, financial crisis, migration) go beyond national frontiers and demand joint actions for solving these problems and call for a transnational network of control mechanisms. Countries, political solutions and power interests (of both governments and civil society) need to be integrated and geared toward sustainability. Because of these developments, the boundaries imposed by national evaluation cultures need to be overcome, the way for transnational joint evaluations be cleared and the functions of evaluation need to be redefined. All three functions of evaluation are affected by this development.

3) Thirdly, the economic rise and increasing political significance of emerging countries (particularly China, India, Brazil and South Africa) are contributing to the complexity of international politics and extending and complicating transnational relationships. A new world order with a multipolar power structure is coming about, with some emerging countries introducing different concepts of the importance of human rights, democratic participation, policy-making and development aid etc. into the

international discussion. These processes will affect evaluation and the different functions of evaluation too. The demands made on evaluation could change, and the western, primarily Anglo-Saxon dominance in evaluation research could begin to falter.

4) Fourthly, should the demand for rationalization of politics as it comes along with concepts like evidence based policy and New Public Management be kept up or even be strengthened the requirements for its rational control will also increase. The result could be an increased demand for evaluation, for example in order to make rational selection decisions in times of limited financial resources, thus demonstrating the impact of policy programs and legitimating their implementation. This means that especially the legitimacy function of evaluations should be strengthened.

5) Fifthly, if the demand for evaluation increases further in modern societies which already use this instrument (possibly with changed requirements) and in countries which follow modern development patterns, the need for qualified evaluators will grow too. Thus the demands for quality and quality assurance in the evaluation process will increase. Thus, it is not surprising that the certification of evaluators is steadily becoming the focus for international discussion. (cf. Brandt 2009; Jacob and Boisvert 2010: 349ff.; Hwalek 2011: 582ff.; Picciotto 2011: 165ff.). Only the Canadian Evaluation Society (www.evaluationcanada.ca) has so far offered such certification. In every other country, 'evaluator' is not a protected title for which a particular course of training or academic qualification is necessary. If a prediction were to be made then it would be that a continuing process of professionalization is likely to lead to the certification of evaluators on the basis of prescribed catalogues of competences. The wide range of Master of Evaluation courses around the globe – and other, non-academic training courses – will further increase the already observable pressure for professionalization.

6) Sixthly, the increase in demand for evaluation competences at *national level* – which may well have changed as compared with today – will be fostered by the general public in (post-)modern societies, that public calling more and more for sustainability-oriented policy-making that integrates economic, social and ecological issues. In this context especially, the role of evaluation as an instrument for criticizing modernity (the social enlightenment function of evaluation) should be emphasized, an instrument which focuses not only on the intended effects of political strategies but also on their unintended effects. Thus the tasks of neutral and objective policy analysis and assessment fall to evaluation.

7) Seventhly, the keyword 'dissolution of political boundaries' describes a development which is characterized by an intensified interdependence of governmental and civil society actors and leads to expanding opportunities and forms of participation in political decisions and their implementation.

For evaluation, this development is likely to mean an increased demand for network evaluations and participatory approaches and a strengthening of the democratic and reflexive function of evaluation.

8) Eightly, the World Wide Web has led to a revolution in communication structures. Little work has been done so far to investigate what this means for evaluation and its social functions. The web society not only raises the question of how to evaluate digital policies and interventions (cf. Leeuw 2012; Leeuw and Leeuw 2012), but also how evaluators can use the Internet for new methods of evaluation (cf. Kistler 2011: 567ff.). The Internet offers opportunities for 'networked evaluation teams', for a more intensive inclusion of stakeholders in the planning and implementation of evaluations and therefore also for greater transparency. These development opportunities could impact all three functions of evaluation. It is the enlightenment function which could profit most, because the Internet increases the chances of the various stakeholders to participate and can also speedily provide global transparency. With its innumerable forums, ongoing online groups, online conferences, workshops and interest groups, the Internet offers an immense variety of opportunities for dialogue and exchange, so that problems in the planning and implementation of evaluations, evaluation methods and results can be communicated and discussed at lightning speed. All of this also increases the chances of a global, integrative evaluation community developing (cf. Labin 2011: 576).

These changes will affect the interests of the sponsors of evaluation. And therefore they affect the functions of evaluation and the extent to which they find application in a society.

The development of evaluation is not only influenced by such challenges and the ability of evaluation to adapt to these new challenges. But there are also *dangers* lurking in the observed evaluation boom:

Firstly, the expansion of evaluation in more and more policy fields and areas of activities also carries with it the dangers that can compromise the positive functions of evaluation. This is the case when evaluation is more and more solidified into routine.

The more evaluation is used as an instrument of control to support democratic governance or to support organizational management processes, the more the *routinization* of procedures can lead to fatigue, to a rule which is supposed to be complied with but no longer has any meaningful substance. Already now the inflationary use of evaluations gives reason for defense and reactance for example in the field of school and university evaluations. This danger also looms if the stakeholders are regularly involved and the findings made public but no consequences – or insufficient consequences – drawn from them. If evaluations make it clear that certain policies are not achieving the desired results and impacts, but are nevertheless maintained out of consideration to a given clientele or on account of lobbyist pressure, evaluation proves to be nothing more than a time-consuming

and costly undertaking without any added value. It is hardly likely that those involved will allow themselves to be won over a second time to an evaluation which was so ineffective (cf. Stockmann 2012a: 212).

This problem of course also occurs if evaluations are not conducted with the necessary expertise and fail to come up with any utilizable findings for that reason. Paradoxically, this risk is becoming greater and greater with the increasing popularity of evaluation, in other words with more and more 'laymen', i.e. insufficiently qualified experts, using the instrument. The situation is becoming even more conducive to this trend, with handbooks being written in 'cookbook style' and crash courses for programme managers suggesting that evaluation can be used by just about anyone. This belief is upheld particularly strongly if there is insufficient money available for the evaluations to be conducted professionally by experts: "That is, while mandates for evaluation often exist, the money to hire formally trained evaluators often doesn't exist" (Datta 2006: 430).

Evaluations which are conducted unprofessionally and in which professional standards are ignored, and evaluations which have no political consequences because their findings are not integrated in decision-making processes, are inappropriate for improving the legitimacy and credibility of policy. Moreover, they also undermine the credibility of the value of evaluation itself. Accordingly, not only the way evaluations are conducted needs to be professionalized, but also the way people deal with them.

In addition, the evaluation boom may lead to the fact that the training of evaluators cannot keep up with the demands and requirements of the clients. If evaluations are no longer sufficiently carried out in a professional manner and appropriate as far as the quality is concerned because there are not enough qualified evaluators, it can happen that the results of evaluations do not meet the expectations and needs of the sponsors.

This inevitably leads to frustration on both sides. The sponsors could turn away

from the instrument of evaluation after such a negative experience and instead seek other methods for the generation of evidence for political control. Or in the worst case: they could dismiss the idea of rational policy making completely.

Evaluation is not the only instrument that is used for gaining political evidence. There is also a variety of other instruments: See for example: accreditation, auditing, organizational development, policy analysis, performance management, action research, supervision, consultancy (cf. Dahler-Larsen 2012). If the task of evaluation is not clearly distinguished from other instruments, it is running the risk of losing its profile, its specific characteristics. Evaluation disappears into the arbitrariness of a variety of instruments. Evaluation is everything and nothing.

Evaluation is increasingly threatened by the fact that it disappears in a sea of consultings. In many countries, science is not sufficiently concerned with the topic of evaluation. Thereby theories and methods of the discipline might dry up. That means science gets lost in the field of evaluation. Sponsors want to be informed quickly, accurately and decision-related. This way the aspect of the research, the theory and method development fall behind. However, this is not the task of consultings, which are increasingly charged with the implementation of evaluations. Therefore, science has to fill this gap and become more involved in the field of evaluation.

V.   CONCLUSIONS

Even soothsayers cannot successfully predict the future. Scientists have also frequently been known to make inaccurate predictions. Nevertheless, it is helpful to offer some well-founded statements about possible future developments so as to prevent ourselves being taken completely by surprise by what occurs in the future. The current article used a series of indicators to show that evaluation is in a historic growth phase. Among the three social functions of evaluation, it is above all in

the context of the management of measures, projects and programs that evaluation is required, so as to increase their sustainability, effectiveness, and efficiency. Within the framework of democratic governance, evaluation is used to produce political legitimacy, and to make politics more transparent and credible. Evaluation has so far been utilized least in its enlightenment function. It is precisely in reflexive, modern societies, which are organized according to sustainable development paradigms, that the critical questioning of development goals and the recording and evaluation of the unintended consequences of goal-directed action are essential.

In the article, it is ascertained that the extent to which evaluation has been institutionalized to achieve these three goals varies considerably. Most governmental and non-governmental organizations have firmly established evaluation as a management function. As an instrument of democratic governance, however, evaluation is less strongly institutionalized. The weakest institutionalized structures are to be found when it comes to the enlightenment function. What are needed here, above all, are independent institutions with their own budget, which could operate freely without administrative constraints and agendas.

Accepting the thesis that evaluation is above all policy driven, the question was then examined how social context factors and, under their influence, political demands on evaluation will change in the next few years, and to what extent this could modify the significance of the three functions of evaluation. It was found that the legitimacy function and the enlightenment function could gain in importance.

The final part of the article discusses the dangers that could negatively influence both the use and the further expansion of evaluation in the processes of politics and civil society.

Evaluation may well continue to flourish, but it has to recognize the signs of the times and pay attention to current challenges; if not, it can quickly lose importance – as happened once before. To prevent this from happening, it is necessary…

- to respond to social trends and the new requirements of the clients to evaluate,
- that evaluation is not frozen in routine so that the usefulness of evaluation is preserved,
- to have sufficient education and training opportunities created so that the quality of evaluation does not suffer,
- that evaluation sharpens its profile and elaborates its strengths compared with instruments of political control,
- that evaluation does not lose its scientificness and just turns into a technique. For this, the discipline of teaching and research in universities must be strengthened.

When evaluation faces these challenges in the 21st Century and does not remain self-pleased in the current social esteem, then it can also apply for the future that these are great days for evaluators.

VI. BIBLIOGRAPHY

Altschuld, J. W. (1990): The Certification of Evaluators: Highlights from a Report Submitted to the Board of Directors of the American Evaluation Association. In: American Journal of Evaluation, Vol. 20, pp. 481-493.

Alvesson, M.; Karreman, D. (2001): Odd Couple: Making Sense of the Curious Concept of Knowledge Management. In: Journal of Management Studies, 38(7), S. 995-1018.

Amelingmeyer, J. (2004): Wissensmanagement. Analyse und Gestaltung der Wissensbasis von Unternehmen. Wiesbaden: DUV.

Becerra-Fernandez, I.; Leidner, D. E. (eds. 2008): Knowledge Management – An Evolutionary View. Advances in Management Information Systems, Vol. 12. Armonk, NY: M. E. Sharpe.

Beywl, W.; Harich, K. (2007): University-Based Continuing Education in Evaluation. The Baseline in Europe. In: Evaluation, 13(1), pp. 121-134.

Borrmann, A. (1999): Erfolgskontrolle in der Deutschen

Entwicklungszusammenarbeit. Baden-Baden: Nomos.

Borrmann, A.; Gleich, A. v.; Holthus, M.; Shams, R. (2001): Reform der Erfolgskontrolle in der deutschen Entwicklungszusammenarbeit: eine Zwischenbilanz. Veröffentlichungen des Hamburgischen Welt-Wirtschafts-Archivs (HWWA), Vol. 63. Baden-Baden: Nomos.

Borrmann, A.; Stockmann, R. (2009): Evaluation in der deutschen Entwicklungszusammenarbeit. Vol. 1 Systemanalysen, Vol. 2 Fallstudien. Studie im Auftrag des Bundesministeriums für Wirtschaftliche Zusammenarbeit und Entwicklung – BMZ. Vols. 8 & 9 'Sozialwissenschaftliche Evaluationsforschung'. Münster: Waxmann.

Brandt, T. (2009): Evaluation in Deutschland. Professionalisierungsstand und -perspektiven. ,Sozialwissenschaftliche Evaluationsforschung', Vol. 7, Münster: Waxmann.

Campbell, D. T. (1969): Reform as Experiments. In: American Psychologist, 24(4), pp. 409-429.

Chelimsky, E. (2008): A Clash of Cultures – Improving the 'Fit' Between Evaluative Independence and the Political Requirements of a Democratic Society. In: American Journal of Evaluation, 29(4), S.400-415.

Dahler-Larsen, P. (2006): Evaluation after Disenchantment? Five Issues Shaping the Role of Evaluation in Society. In: Shaw, I.F.; Greene, J.C.; Melvin M. (Ed.): Handbook of Evaluation. Policies, Programs and Practices. London: Sage, pp. 141-160.

Dahler-Larsen, P. (2011): The Evaluation Society. Stanford University Press: Stanford.

Dahler-Larsen, P. (2012): The Changing Role of Evaluation in a Changing Society. Lecture presented at the conference 'The Future of Evaluation' at the Center for Evaluation, Saarland University.

Datta, L. (2006): The Practice of Evaluation: Challenges and New Directions. In: Shaw, I. F.; Greene, J. C.; Melvin, M. (Hg.): The Sage Handbook of Evaluation. S.419-438.

EU (2010): Documents and Links on evaluation. Available: http://ec.europa.eu/dgs/secretariat_general/evaluation/documents_en.htm (15.07.2013).

Fitzpatrick, J. L.; Sanders, J. R.; Worthen, B. R. (2004): Program Evaluation. Alternative Approaches and Practical Guidelines. 3rd edition. Boston et al.: Pearson.

Furubo, J.-E.; Rist, R. C.; Sandahl, R. (2002) (Ed.): International Atlas of Evaluation. New Brunswick: Transaction Publishers.

Götz, K.; Schmid, M. (2004): Praxis des Wissensmanagement. München: Vahlen.

Haun, M. (2005): Handbuch Wissensmanagement. Grundlagen und Umsetzung, Systeme und Praxisbeispiele. Berlin: Springer.

Hwalek, M. (2011): The Stars are Aligned for Disruptive Change in Evaluation Practice. In: American Journal of Evaluation, 32 (4), pp. 582-586.

Ipe, M. (2003): Knowledge Sharing in Organizations: A Conceptual Framework. In: Human Resource Development Review 2/4, S. 337-359.

Jacob, S.; Boisvert Y. (2010): To Be or Not to Be a Profession: pros, Cons and challenges for Evaluation. In: Evaluation, 16(4), pp. 349-369.

Jones, S. C.; Worthen, B. R. (1990): AEA Members' Opinions Concerning Evaluator Certification. In: American Journal of Evaluation, Vol. 20, pp. 495-506.

Kistler, S. J. (2011): Technology, Social Networking, and the Evaluation Community. In: American Journal of Evaluation, 32 (4), pp. 567-572.

Kuhlmann, S.; Bogumil, J.; Wollmann, H. (Ed.) (2004): Leistungsmessung und –vergleich in Politik und Verwaltung: Konzepte und Praxis (Stadtforschung aktuell Bd. 96). Wiesbaden: Verlag für Sozialwissenschaften.

Labin, S. N. (2011): Shaping the Future: An Integrative Methodological Agenda. In: American Journal of Evaluation, 32 (4), pp. 572-578.

Leeuw, F. L.; Toulemonde, J.; Brouwers, A. (1999): Evaluation activities in Europe: a quick scan of the market. In: Evaluation, 5(4), pp. 487-496.

Leeuw, F.L. (2009): Evaluation – a Booming Business but is it Adding Value? In: Evaluation Journal of Australasia, 9(1), S.3-9.

Leeuw, F. L. (2012): Evaluation in the Future: On Internet and a Few Other Developments. Lecture presented at the conference 'The Future of Evaluation' at the Center for Evaluation, Saarland University.

Leeuw, F. L.; Leeuw, B. (2012): Cyber society and digital policies: Challenges to evaluation? In: Evaluation, 18(1), pp. 111-127.

Meyer, W.; Stockmann, R. (2006): Comment on the Paper: An Evaluation Tree for Europe. Vortragsmanuskript zur EES-Tagung: Evaluation in Society: Critical Connection.

Picciotto, R. (2002): Evaluation in the World Bank: Antecedents, Methods, and Instruments. In: Furubo, J.-E.; Rist, R. C.; Sandahl, R. (Ed.): International Atlas of Evaluation. New Brunswick: Transaction Publishers.

Picciotto, R. (2011): The Logic of Evaluation Professionalism. In: Evaluation, 17(2), pp. 165-180.

Pollitt, C. (1998): Evaluation in Europe: Boom or Bubble? In: Evaluation, 4(2), pp. 214-224.

Preskill, H. (2008): Evaluation's second act – a spotlight on learning. In: American Journal of Evaluation, 29(2), S.127-138.

Rist, R.; Stame, N. (2006): From Studies to Streams – Managing Evaluative Systems. New Brunswick: Transaction Publishers.

Rugh, J. (2011): The Expending World Scene in Evaluation. In: American Journal of Evaluation, 32(4), pp. 586-593.

Schedler, K.; Proeller, I. (2003): New Public Management (2., überarbeitete Auflage). Bern: Haupt.

Smith, M. F. (1990): Should AEA Begin a Process for Restricting Membership in the Profession of Evaluation? In: American Journal of Evaluation, Vol. 20, pp. 521-531.

Smith, N. L.; Brandon, P. R. (2011): If Not to Predict, At Least to Envision, Evaluation's Future. In: American Journal of Evaluation, 32(4), pp. 565-566.

Smith, N. L.; Brandon, P. R.; Hwalek, M.; Kistler, S. J.; Labin, S. N.; Rugh, J.; Thomas, V.; Yarnall, L. (2011): Looking Ahead: The Future of Evaluation. In: American Journal of Evaluation, 32(4), pp. 565-599.

Speer, S;, Jacob, S.; Furubo, J.-E. (2013): Different Paths for Institutionalizing Evaluation: Updating the International Atlas of Evaluation 10 Years Later. In: Evaluation – The International Journal of Theory, Research and Practice [forthcoming].

Stockmann, R. (1997): The Sustainability of Development Cooperation. Baden-Baden: Nomos.

Stockmann, R. (2006) (Ed.): Evaluationsforschung: Grundlagen und ausgewählte Forschungsfelder. 3rd edition. Münster et al.: Waxmann.

Stockmann, R. (2008): Evaluation and Quality Development: Principles of Impact-Based Quality Management. Frankfurt et al.: Peter Lang.

Stockmann, R. (2009): Evaluación y Desarrollo de la Calidad. San Jose: Universidad de Costa Rica.

Stockmann, R. (ed. 2011): A Practitioner Handbook on Evaluation. Cheltenham / Northampton: Edward Elgar.

Stockmann, R. (ed. 2011): Manual de Evaluación: Una guía práctica de procedimientos. San José: Editorial. Universidad de Costa Rica.

Stockmann, R. (2012a): Evaluation und Gesellschaft. In: Strobl, R.; Lobermeier, O.; Heitmeyer, W. (Hg.): Evaluation von Programmen und Projekten für eine demokratische Kultur. Wiesbaden: Springer VS. S.195-221.

Stockmann, R. (2012b):Von der Idee zur Institution. Institut für Deutsche Entwicklungsevaluierung gegründet (From an idea to an institution. Institute for German Development Evaluation founded). In: Zeitschrift für Evaluation, 11(1), pp. 85-93.

Stockmann, R. & Meyer, W. (2013): Functions, Concepts and Methods for Evaluation Research. Basingstoke: Palgrave Macmillan (in print; planned publication date: fall 2013).

Vedung, E. (2010): Four Waves of Evaluation Diffusion. In. Evaluation, 16(3), pp. 263-277.

Weiss, C.H. (1998): Evaluation – Methods for Studying Programs and Policies. New Jersey: Prentice Hall.

Willke, H. (2004): Einführung in das systematische Wissensmanagement. Heidelberg: Carl-Auer-Systeme-Verlag.

Winkler, K. (2004): Wissensmanagementprozesse in face-to-face und virtuellen communities. Konzepte, Gestaltungsprinzipien und Erfolgsfaktoren. Berlin: Logos.