

Utilization of Evaluation Results in Policy-Making and Administration: A Challenge to Political Science Research

*Hellmut Wollmann**

UDK 35.07:32.019.52
323.01

Original scientific paper / izvorni znanstveni rad

Received / primljeno: 26. 5. 2016.

Accepted / prihvaćeno: 9. 9. 2016.

The paper deals with utilization of evaluation results in policy-making and public administration. The article is written in five steps. First, major variants of evaluation are presented with a particular eye to their respective utilization potential. Second, different concepts are sketched that have been forwarded to capture the utilization of (social) science-generated knowledge by political, administrative, and social actors. Third, looking at Germany, Switzerland and the European Union pertinent research findings are discussed. Parts four and five contain summary, conclusions and some remarks on ensuring research needs. The political rationality and the underlying political will of the decision-makers prevail in crucial political decisions while in less important decisions evaluation-generated evidence does show some effect. However, the stock of evaluation-generated knowledge about the effects of policies has

* Professor Emerituss Helmut Wollmann, retired from Humboldt-Universität in Berlin, Germany (profesor emeritus Sveučilišta Humboldt u Berlinu, Njemačka, e-mail: hellmut.wollmann@rz.hu-berlin.de).

seen an exponential growth, which makes the call for evidence-based policy-making and for utilizing the available evaluation findings all the more mandatory and urgent.

Keywords: evaluation, policy-making, public administration, utilization of evaluation results

1. Introduction

The utilization of (social) science-generated knowledge by political and social actors has been advocated and hailed for a long time as a crucial means and resource for the improvement of policy-making in modern times.¹ Max Weber spoke of the “intellectualist rationalization through science”² (Weber, 1922, p. 593) and Harold Lasswell proclaimed “policy sciences” as an all-out effort to mobilize the societal and interdisciplinary knowledge to promote political and social progress (Lasswell, 1951, p. 3; see also Wagner, Wittrock & Wollmann, 1991). The idea of a “(social) scientification” of the political and societal activities has been voiced by Donald Campbell in his call for “reforms as experiments” (Campbell, 1969; see also Hellstern & Wollmann, 1983; Danielson, 2007) which aimed at guiding policy-making by scientifically accompanied and evaluated “social experiments”. Moreover, the maxim of “evidence-based policy-making” that was advanced by the New Labour Government in the U.K. in the late 1990s (Cabinet Office, 1999)³ and has since gained (almost inflationary) international currency is targeted at grounding political decision-making on empirical evidence.

The evaluation of public policies, programmes and measures, which first emerged as a policy strategy in the US during the 1960s and has since been pursued in many countries (Furubo & Sendahl, 2002; Wollmann, 2003, 2003c), has come to provide an enormous body of empirical evaluative findings on the successes and failures of policies. This rapidly expanding fundus

¹ This article is a significantly revised version of a previously published piece (in German language; see Wollmann 2014), which was also translated into Chinese and published in Journal of Fujian Administration Institute, 2016, spring issue.

² „Intellektualistische Rationalisierung durch Wissenschaft und wissenschaftlich orientierte Technik“

³ Government “must produce policies that really deal with problems that are forward looking and are shaped by evidence rather than by response to short-term pressures: that tackle causes not symptoms.”

of internationally and nationally available policy-relevant knowledge renders the question whether, when and how such knowledge has been utilized (or not) in policy-making and policy implementation ever more urgent.

In the pursuit of this guiding question, the article will come in five steps. First, major variants of evaluation will be briefly presented with a particular eye on their respective “utilization potential”. Second, different concepts that have been forwarded to capture the utilization of (social) science-generated knowledge by political, administrative and social actors will be sketched. Third, pertinent research findings will be discussed looking at Germany, Switzerland, and the European Union as “cases in point”. Fourth, a summary and conclusion will be given. Fifth, some remarks on ensuing research needs will be made.

2. Variants of Policy Evaluation and Their Respective Utilization Potential

At the outset, major variants of evaluation (Wollmann, 2007, p. 393; 2003a) and their respective utilization potential are outlined.

2.1. Ex-post Evaluation

Ex-post evaluation is the classical variety of evaluation used to assess the goal attainment of policies and measures once they have been terminated. Typically, such evaluations are faced with two crucial conceptual and methodological problems. First, the policy goals in terms of intended consequences need to be conceptualised by defining appropriate and, if possible, measurable indicators. At the same time, unintended consequences have to be taken into consideration. Second, it must be determined whether the observed effects have been caused by the policy or measure concerned, which is a methodologically complex and demanding question.

Because of the research skills and capacity which the conduct of (fully-fledged) ex-post policy evaluations requires, they are usually carried out by (external) researchers by way of commissioned/contractual research (on contractual research see Wollmann, 2002). Once the (commissioned) evaluator has completed the study and delivered the results, it is, as a rule, entirely up to the (commissioning) political actors or agency to decide whether and how the results are going to be used.

2.2. Ex-ante Evaluation

Ex-ante evaluation is directed at (in advance) assessing the effects and consequences of an envisaged policy, programme or measure, whereby possible different (alternative) paths of action (scenarios) are to be anticipated and explored. In an ex-ante evaluation it is, as a rule, left to the respective political actor to decide whether and how the ex-ante evaluation is used.

The ex-ante logic is also characteristic of the various types of impact (pre-) assessment, such as the environmental impact assessment, the legal regulation impact assessment (in Germany: *Gesetzesfolgenabschätzung*, see Konzendorf, 2009) and the regulatory impact assessment. Moreover, also hinging on the ex-ante logic, the ex-ante cost benefit analysis aims at ascertaining both the achievement (benefits) and the costs of an envisaged policy or measure (possibly also pursuing alternative paths of action). The simultaneous assessment of costs and benefits of a measure allows weighing the pros and cons and providing a “balance sheet” thereof.

The comprehensive evaluation system which the European Union has used for the evaluation of its structural fund programmes since the mid-1990s is a somewhat exceptional variety of ex-ante assessment. Within each of the five-year programme phases, a sequence of *ex-ante*, *ongoing* (in the EU terminology: intermediate) and *ex-post* evaluation steps is prescribed and it is conceptually and procedurally mandated to feed the results of the ex-ante evaluation into the subsequent programme stages.

2.3. Ongoing Evaluation

Ongoing evaluation, which sets in as soon as the policy or measure in question starts to be implemented, has the purpose to ascertain (interim) effects generated by the respective policy and measure. A crucial task of ongoing evaluation is to transmit (feedback) the interim results to the policy-makers and/or project operators concerned in order to enable them to rectify and modify the related policy design or implementation process. Hence, ongoing evaluation is meant to enable and foster the utilization of pertinent information.

However, it is as a rule left to the operator of the programme or measure to decide whether and which use is made of the incoming feedback information. By contrast, in the interventionist and participatory varieties of ongoing evaluation the “evaluate” is actively involved in the mutual learning and utilization process (see below 2.4.)

Benchmarking. Benchmarking may be seen as a variety of ongoing evaluation in which, after periodical or non-periodical monitoring of the relevant changes and effects, the results are put into comparative perspective, be it intra-organisationally or inter-organisationally (for the development of benchmarking in Germany see Korte, 2004; Hollenrieder, 2004; for an international overview see Kuhlmann, 2010; Jäkel & Kuhlmann, 2014; Kuhlmann & Wollmann, 2014: 226 ff.). The decision whether and how to use benchmarking again depends on the respective operative unit. It should be mentioned that in the reform of German federalism in 2009 benchmarking received a constitutional recognition (in article 91b of Federal Constitution) as a comparison-related tool to “assess and promote the performance of administration”.

New Public Management-inspired performance management. A remarkable move towards institutionalizing the crucial feedback loop can be seen in the (indicator-based) performance management system that is part and parcel of the New Public Management (and its German derivate “New Steering Model” – *Neues Steuerungsmodell*) concept. It hinges on the premise that (indicator-based) information on the current performance, including achievement and cost data, is (steadily) reported back (feedback) to the responsible operators and actors. Insofar as such indicator-based controlling is designed to function primarily intra-administratively, i.e. within the respective administrative unit, it can be seen as a form of self-evaluation which, echoing Niklas Luhmann’s concept of self-referentiality (see Luhmann, 1993), is designed to link the intra-administrative actors with the flow of feedback information directly and steadily. In order to connect the feedback loop with the extra-administrative world, an indicator-based reporting system meant to inform the political decision-makers (for instance parliament and local council) as well as the general political public about the state of affairs is put in place (on Germany see Bogumil et al., 2007, p. 303; Kuhlmann et al., 2008). Because of assigning the feedback loop such as central evaluative function and of procedurally integrating it into the evaluation system, NPM has been acknowledged by some as inaugurating a “new wave” of evaluation (Wollmann, 2003a).

2.4. Interactive, participatory and user-focused evaluation

Within ongoing or accompanying evaluation one can distinguish between a primarily analytical modality in which the evaluator remains detached and distanced from the evaluated operation in order to assert the inde-

pendence and objectivity of the evaluation process and its findings (Wollmann, 2007, p. 394). By contrast, the interventionist modality of ongoing/accompanying evaluation implies that the evaluator is assumed to actively engage in the implementation process in order to rectify shortcomings and give advice. Such interventionist orientation approximates the (social science) concept of action research (Wollmann, 2007, p. 394). While this participatory mode of evaluation jeopardizes and impairs the objectivity of its findings, it enhances the potential of mutual learning and the sharing of insights.

Similar strategies to bridge the hiatus between the role of the evaluator and that of the “evaluatee” have been advanced – with different accents on the role of the evaluatee – as interactive (Balthasar, 2012), user-related (Patton, 1997) or empowerment-related (Fetterman et al., 1996) modes of evaluation. They have in common that the role of the evaluatee is enhanced in the generation and utilization of evaluation results (Wollmann, 2013, p. 92).

3. Concepts of Utilizing Research-Generated Knowledge

Dating back to the 1970s, knowledge utilization research has developed in the USA as a social science research field, addressing the question whether, how, and why social science-generated knowledge has (or has not) been used and applied in the political, administrative and social practice. Carol H. Weiss was perhaps the most prominent and influential scholar in this field (Weiss, 1979; see also Wollmann, 2009, pp. 392 ff.). During the 1980s, the research focus in Germany was taken up by a research group funded by the German Research Council (*Deutsche Forschungsgemeinschaft*) and coordinated by Ulrich Beck (for the concept and results of this research group see Beck & Bonß, 1989).

In the course of discussion on knowledge utilization, different approaches have evolved which can be grouped along the distinction between political and scientific rationality. Ideally speaking, the former may be seen as characteristic of the politico-administrative world whose actors are essentially driven by their will to gain and retain power and to realize their own and their followers’ interests. By contrast, the scientific rationality can be regarded as specific of the science system and its members being (ideal-typically) committed to an independent, objective, and impartial search for truth.

In an early phase of that debate the concept of the two worlds prevailed (Caplan et al., 1975) being both based on and driven by profoundly different and essentially incompatible rationalities and logics.

Within the range of concepts on the relations between the world of politics and the world of science one view perceives the decision-making in the political world as determined by the political will of the ruling majority and by its resolve to enforce their and their clientele's interests while discarding scientific evidence. This political rationality-bound mode of decision-making corresponds with the decisionist model proposed by Jürgen Habermas (Habermas, 1968).

Closely bordering the political rationality-dominated decision-making model, the concept of the instrumental knowledge utilization is based on the assumption that the key political decisions (on policy goals, resource allocation etc.) are determined politically, i.e. by political rationality, while social science research-generated knowledge is turned to by the political decision-makers to provide advice, information and means how to attain the politically pre-decided goals (Alkin, 2005). A similar instrumental concept and connotation of ancillary knowledge utilization resonates in the engineering and problem-solving models (Weiss, 1979).

Similarly, the concept of the symbolic use of knowledge tends towards the dominance of political rationality as it assumes the relevant politico-administrative actors tap and use pertinent (social) science-generated knowledge and expertise in order to confirm and (post factum) legitimize pre-decided decisions (see also Balthasar, 2009, p. 493; for a more recent discussion and distinction of various uses of evaluation results with ample references see Hojland, 2015a, pp. 54-64).

In a similar vein, tactical use of research-generated knowledge is guided by political rationality when political actors resort to commissioning research with the purpose of buying time and sitting out current conflicts (Weiss, 1979).

By contrast, the scientific rationality prevails in concepts according to which political decision-making is (largely) guided by (social) science-generated knowledge and evidence. A striking example can be found in the concept of and call for "reforms as experiments" as voiced and propagated by Donald Campbell (Campbell, 1969). His concept of experimental policy-making is based on the idea that the decision to adopt a certain policy or measure should finally be taken only after it was exposed to an "experimental real test" and to a rigorous scientific evaluation (for an overview see Hellstern & Wollmann, 1983; Danielson, 2007). The principle

that political decisions should be essentially guided by scientific expertise corresponds to Jürgen Habermas' technocratic model (Habermas, 1968). Another school of thought ascribes certain superiority and enlightenment potential to the scientific rationality *vis-à-vis* the political world. Hence, even when and if (evaluation) research-generated insights are not immediately and directly accepted and translated into political decisions, they may – in the form of “data, ideas, arguments” (Weiss, 1991) – make their way and trickle into the decision-making arena and actor constellation through multifarious diffusion paths and information networks, such as parliamentary hearings, professional workshops, publications, mass media, informal contacts etc. (Krautzberger & Wollmann, 1988), before they finally arrive in a possibly changed (Caplan, 1983) form.

Finally, one should mention a kind of compromise concept in which the possible contradiction and conflict between the political and scientific rationality is bridged by a dialogue or pragmatic model (Habermas, 1968). It largely falls in line with the perception that in the “real world” of evaluation, particularly in its various ongoing variants, the evaluation process unfolds as an interaction between the evaluator and the evaluatee and, consequently, as a mutual cognitive and learning process.

4. Some Empirical Findings

The following chapter presents brief accounts of the research that has been conducted on the utilization of evaluation with Germany, Switzerland and the European Union (EU) singled out as cases in point.

Since the late 1960s, Germany has been one of the frontrunners in the expansion of evaluation research among European countries and continued to rank among the evaluation-intensive countries (Furubo & Sandahl, 2002; Wollmann, 2003b, 2003c; Widmer, Beywl & Fabian, 2009). In Switzerland since the late 1990s, policy evaluation has experienced a strong expansion which has lifted its evaluation profile to a leading position among European countries (Widmer & Beywl, 2009, p. 515). Thus, both Germany and Switzerland should lend themselves to serve as instructive examples of the use of evaluation. On the level of the European Union (EU) evaluation has, since 1995, become centrally enshrined in the seven-year programming cycle prescribed for its structural funds with systematic sequence of ex-ante, intermediate (ongoing) and ex-post evaluation (Leeuw, 2006, p. 72; Smismans, 2015, pp. 10 ff.). The importance

of the use of evaluation has been highlighted in the Commission's Communication (2007) Responding to Strategic Needs: Reinforcing the Use of Evaluation.⁴

4.1. Germany

More recently strong impulses to further promote evaluation have come, *inter alia*, from labour market policy and education policy (particularly under the OECD's pressure conduct PISA-type surveys and assessments) (for an overview see the policy-related chapters in Stockmann, 2006). In the Federal Budgetary Regulation (*Bundeshaushaltsordnung*) of 2001, it is stipulated that evaluations (in the terminology of that provision: "success controls", *Erfolgskontrollen*) should be carried out on (all) "financially relevant measures" (*Erfolgskontrollen finanzwirksamer Maßnahmen*). At the same time, the Federal Budgetary Regulation prescribes that, besides carrying out success control on financially pertinent measures, studies should be conducted as to whether the results of such success controls are implemented (*umsetzen*) by the ministry or agency concerned.

Measured by the large (and ever expanding) volume of policy evaluations that have been conducted (and are being conducted) on the federal and *Länder* levels and considering that since 2001 the studies on implementation (*Umsetzung*) of evaluation results have been given legal recognition and salience, the number and scope of available pertinent studies and data is, at least on the basis of our analysis, remarkably (and surprisingly) scarce.

In 1989, president of the Federal Court of Audit (*Bundesrechnungshof*) – in his function as Federal Mandatee for the (Economic) Efficiency in Administration (*Bundesbeauftragter für Wirtschaftlichkeit in der Verwaltung*) commissioned a comprehensive study directed at evaluating the implementation of the aforementioned Federal Budgetary Regulation of 2001 on two crucial scores: whether success controls were conducted by the ministries concerned, and second, whether the results of such success controls were implemented (*umsetzen*).

With regard to the question whether evaluation (success control) results were implemented, the pertinent investigation conducted in 1989

⁴ http://ec.europa.eu/smart-regulation/evaluation/docs/eval_comm_sec_2007_213_en.pdf

reached the conclusion that the utilization rate was small (*gering*) (Bundesbeauftragter, 1989, p. 30; Stockmann, 2006, p.33).

Some ten years later, president of the Federal Court of Audit commissioned a follow-up investigation. Again the assessment of the utilization rate was largely negative. The implementation of the results of success control (e.g. correction, continuation, or termination of a measure) with the exception of few ministries varied greatly, but was in sum small (*gering*). In some cases, measures were continued or terminated in defiance of the pertinent results of “success control without explaining why” (Bundesbeauftragter, 1998, p. 27).

Only the Federal Ministry of International Economic Cooperation was positively set off in various aspects. For one it was pointed out that “insights into the causes of failed measures are taken notice of in follow-up projects”, including “learning by and from mistakes” (Bundesbeauftragter, 1998, p. 28). Moreover, it was accentuated that “in selected ... projects 10 to 15 years after completion a follow-up ex post evaluation (ex post ex post) was conducted in order to assess the lasting (*nachhaltig*) success of a measure” (Bundesbeauftragter, 1989, p. 24).

The follow-up study of 1998 also addressed the question whether evaluative success controls (*Erfolgskontrolle*), as required by the Regulation of 2001, were carried out. It was stated that success controls failed to be conducted especially in politically sensitive cases. “In the cases in which it was foreseeable that the political head (*Leitung*) (of the ministry) could be (negatively) affected (*berührt*) by failure of a measure, the officials in charge of success control anticipated conflicts with the political head which they wanted to avoid.” (Bundesbeauftragter, 1998, p. 33)

The Federal Ministry of Economic Cooperation, which got a positive rating in the aforementioned report, has been standing out among federal agencies in other regards thanks to its initiatives and innovations in the field of evaluation. As early as at the beginning of the 1970s this ministry started to build up and institutionalize an evaluation system (Stockmann, 2006b; Zintl, 2009). It was the first federal ministry to create an intra-ministerial evaluation and inspection unit which was from the outset assigned the task to foster the utilization of evaluation results by feeding them back into ongoing decision-making as well as into ministry-related vocational training (Lorz, 1984, p. 293). While in the initial phase evaluation studies were largely directed at individual development and aid projects, the ministry’s evaluation strategy has, for a number of years, increasingly aimed at accumulating and synthesizing the collected evalua-

tion knowledge and information – by way of cross-cutting evaluations (so called meta evaluations). General recommendations and criteria should be gleaned from the latter to serve as guidance for the selection, planning, and conduct of future development projects (Stockmann, 2006b, p. 380). In addition, in a more recent move, evaluation activities have addressed the long-term effects of development projects that were hitherto often neglected as the focus of evaluation was directed on short-term effects and results. In sum, the evaluation strategies and activities pursued by the Federal Ministry of Economic Cooperation have proved exemplary of a ministry's handling of its evaluation efforts as well as of making evaluation results available to the public.

Over the years, the academic (university based) social science and political science research community has exhibited hardly any interest in the study of the utilization of social science-generated knowledge, leave alone evaluation-generated findings. It is true that during the 1980s a research consortium that was coordinated by Ulrich Beck and funded by the German Research Council (*Deutsche Forschungsgemeinschaft*) conducted an empirical research on the utilization (Beck & Bonß, 1989). However, there has not been any noteworthy major university-based research project in this field since.

It should be noted, though, that quite recently a group of university-based social scientists embarked upon a (still small-scale) project which is set to study the utilization of research findings in German federal ministries (Spiel & Bergsmann, 2009). On the basis of a questionnaire (sent out to a very small sample of high-level ministerial officials – with a low response rate, see Spiel & Bergsmann, 2009, p. 483) it was concluded that the “discrepancy between the importance which the utilization of evaluation results has and its practice is very large” (Spiel & Bergsmann, 2009, p. 465; see also Bergsmann & Spiel 2009, Spiel & Schober 2015).

Finally, an (earlier) study should be mentioned which, in a kind of a case study, dealt with the utilization of evaluation-generated information in the legislative activities of a federal ministry (that is, the Federal Ministry of Urban Development (*Bundesbauministerium*) (Krautzberger & Wollmann, 1988). Based on anecdotal (rather than systematic) evidence (including personal participation of one of the authors in the respective legislative process), the study sheds some light on the complex setting and the course of legislative decision-making in which the evaluation-generated information flow is only one of several relevant information channels to which the responsible legislative actors are exposed and which provide

them with often conflicting (interest-laden) information and data. Regarding the substance of decision-making, the study argues that evaluation-generated knowledge hardly exerts noticeable influence on legislative decisions that rank high on the political agenda of the acting government and its ministers, thus letting the political rationality prevail. Yet, when it comes to settling minor issues in the elaboration of legislative drafts, evaluation-generated information, instrumental knowledge, and advice often do play a relevant role. The study also gives some insight into the peculiarities of the ministerial and parliamentary legislative process as relevant evaluation-generated information may not find direct access to legislative decisions, but may go through delays and detours, can be stored and filed in ministerial archives or minds, before being put to use, perhaps in an altered or converted form, if and when a (legislative) window of opportunity finally opens.

Finally, we should mention an inter-university research initiative directed at studying the generation, diffusion and application/utilization of research knowledge in different fields of public policy and administration. The initiative coordinated by Sabine Kuhlmann and institutionally based at the University of Potsdam aims at receiving a major collective grant from the German Science Foundation (Kuhlmann et al., 2016). If it launches successfully, the project could link up with the research project on knowledge utilization that was coordinated by Ulrich Beck (also funded by the German Science Foundation) during the mid-1980s (Beck & Bonß, 1989) and which has not seen a thematically pertinent follow-up research effort since.

4.2. Switzerland

Since the late 1990s, policy evaluation in Switzerland has experienced a strong expansion which has led it to the leading position among European countries (Widmer & Beywl, 2009, p. 515). This development was significantly promoted in 1999 by the adoption of Article 170 of the Federal Constitution in which the federal Parliament (*Bundesversammlung*) is mandated to “ensure that the effectiveness of the measures undertaken by the Federation be reviewed” (Mader, 2009, p. 53). Since then the evaluation function has been laid down in a large number of legal provisions, such as evaluation clauses, parliamentary supervision of administrative operations, federal finance control (*Eidgenössische Finanzkontrolle*). On the federal level, some 500 evaluation studies were carried out between 1999 and 2002 (Balthasar, 2007; Mader, 2009, p. 60; for a recent overview see Balthasar, 2015).

The use of evaluation results by federal administration was addressed in a study undertaken by *Andreas Balthasar* in 2008 (Balthasar, 2009). Methodologically and conceptually, it was conducted as a survey among the federal officials in charge of evaluation in the federal administration. The result of the survey was that two thirds of federal officials concerned indicated that the utilization rate was high or fairly high (*hoch oder eher hoch*) (Balthasar, 2009, p. 497). In about a half of the evaluations with a high or fairly high utilization rate, the utilization mode was judged as directly implementing the recommendations given by the evaluators (Balthasar, 2009, p. 498). To explain the comparatively high utilization rate Balthasar argues, in referring to Michael Patton's process-related utilization concept, that "a high utilization rate can be expected if and when the evaluator closely cooperated with the evaluatee" whence he derives the notion of the role of the evaluator as being a "critical friend". By contrast, Balthasar does not find "any evidence that the chance of knowledge utilization improves when, in the institutionalization of the evaluation function, the distance and independence between evaluator and evaluatee are given priority" (Balthasar, 2008, p. 243). The salience which evaluation (and its utilization) has gained in the Swiss political system is shown in the creation of a major research project called *SynEval* in January 2013. The project has been financed through the Swiss National Science Foundation and is designed to "address the questions how policy evaluation in Switzerland is influenced by the Swiss political system and how policy evaluation in turn influences the Swiss political system".⁵ One of the main *SynEval* projects focuses on "parliaments and evaluations" and on the question of how evaluation influences the parliamentary legislative, oversight, and budgetary functions (Eberli et al., 2014).

In the pursuit of this project, in the spring of 2014, an online survey was conducted among members of parliament at the national and regional (cantonal) levels to explore their experience with evaluation. In total some 1,570 (or 55.3 per cent) out of 2,841 members of parliament participated in the online survey. The results provide quite a differentiated picture of the utilization practice (for details see Eberli et. al, 2014). Most respondents assess evaluation positively and indicate that in their parliamentary practice they make a rather frequent use of evaluation findings in the full range of their parliamentary (legislative, oversight and budgetary) tasks.

⁵ See: <http://www.syneval.ch/index.php/en/about-syneval>. Institutionally, the project is anchored mainly in the Department of Political Science of the University of Zürich and coordinated by Thomas Widmer.

4.3. European Union

In 1995, the European Union introduced an evaluation system for its structural funds which hinged on a sequence of seven-year programming phases with each of the phases comprising the evaluation cycle of ex-ante, intermediate (on-going) and ex-post evaluation (Leeuw, 2006, p. 72; Stern, 2009; Smismans, 2015). The complexity and scope of the EU's evaluation system is enhanced by the provision that evaluation is designed to be carried out at both the level of the European Commission and that of the Member States.

At the EU level each Directorate-General is held to evaluate (or commissioned to evaluate) the policy and funding programmes within its respective jurisdiction. For instance, during the funding period of 1996-2000 some 470 evaluation studies were commissioned (European Commission, 2001, p. 29). The "Annual Evaluation Reviews" give an account of the current evaluation projects (see e.g. European Commission, 2001, 2010). At the national level, the Member States (in Germany, because of its federal structure, this applies first of all to the regional States, *Länder*) are in turn obliged to conduct evaluations on the structural funding they receive. They, too, have to follow the three-phase evaluation cycle (ex-ante, intermediate, and ex-post) pursuing the five-year rhythm (European Commission, 2001, p. 239; 2010, p. 1). In all the countries that receive European funding this has resulted in the emergence of evaluation research "landscape" to carry out (and to be financed through) these evaluation projects (for Germany see Toepel & Schwab, 2005).

The following account will draw on an empirical study on the "utilization of evaluation results in the Commission" ordered by the European Commission from an independent research team (EPEC)⁶ (EPEC, 2005a, 2005b).

The study examines the use of evaluations during the period 2002-2004. The focus was one the use of evaluation *within* the Commission, i.e. by its Directorates General and the respective operative units. The study focused on the use of evaluations in EU-funded programmes and policies and covered all types of evaluation: ex ante, intermediate, final, and ex post (EPEC, 2005a, p. 1).

The (conceptually sophisticated and empirically penetrating) study concluded "that evaluation is highly influential in the design and implementation of interventions ... The most common contribution of evaluation

⁶ Composed of K. Williams, B. de Laat, G. Bastian and E. Stern.

work lies in the improvement of interventions. However, there are policy instruments of the Commission where evaluation is not yet so common and influence is also less evident, especially in the area of legislation ... (Evaluations) are less influential in the setting of political priorities or choosing between different options per se. However, there are examples such as ex ante evaluation of Structural Fund programmes where political priorities have been influenced by evaluation at both the national and Commission levels (Cohesion policy). Finally, all evidence shows that overall allocation of EU resources is certainly *not* determined on efficiency grounds, but much more by political decision-making. The study shows especially that such decisions are generally very little informed by evaluation. Evaluation has however led to changes in resource allocation *within* interventions (especially expenditure programmes), in changes in eligibility criteria and incremental improvements in operational efficiency – the latter clearly resulting from evaluations exposing inefficiencies” (EPEC, 2005a, p. iii).

Moreover, the EPEC report notes that the prescribed sequence of evaluation steps (ex-ante, intermediate, and ex-post) has been routine for quite some time.⁷ It is observed that the evaluation knowledge gathered in the precedent programme phase has been utilized in the subsequent programme phase, no matter whether related to ex ante, intermediate, or ex post evaluation in the narrow sense.⁸ The steering groups, which play a significant role during the entire duration of programme cycles, along with representatives of the respective General Direction, comprise external actors who are involved in the programme implementation (EC, 2010, p. 36). This actor participation has been deemed a crucial precondition for the utilization of interim-evaluation information⁹ – in line with (Patton’s) process-related utilization concept.

Possibly responding to recommendations by the EPEC report, the European Commission subsequently embarked on a reorientation of its evaluation policy to improve the use of the evaluation process (Smismans,

⁷ “All the evidence gathered in this study shows that the EC evaluation has become highly integrated into programme management, with results being routinely used by programme managers” (EPEC, 2005a, p. 20).

⁸ See similarly EC, 2001, p. 32: “In practice intermediate and ex post evaluation carried out for the previous periods are often used in the preparation of a programme renewal. In fact, the distinction between ex ante, intermediate and ex post evaluations is sometimes artificial”.

⁹ “Involvement of potential users is important for the uptake of evaluation results” (EPEC, 2005a, p. 44).

2015, p. 10). A first step was set in the 2007 Commission Communication “Responding to Strategic Needs: Reinforcing the use of evaluation”¹⁰ and later the evaluation function was placed in the broader Better Regulation and Smart Regulation agenda (for details and references see Smismans, 2015, p. 11).

In a conceptually and methodologically sophisticated study on the use of evaluation in the EU’s programme undertaken by *Steven Hoyland* in 2015 (Hoyland, 2015)¹¹ three programme areas were singled out for an empirical analysis. First, the investigation was directed at “how an evaluation system is implemented and how the issues of accountability and learning are reflected in the practices implemented” (p. 74). Second, it was meant to “analyse the use of four evaluations of the Commission’s Programme for the Environment and Climate Action (LIFE) over a ten-year period. “The LIFE programme was chosen because it is a ‘classic’ centrally-managed EU expenditure programme ... The programme has experienced a full Commission evaluation cycle (ex-ante, mid-term, final, and ex-post) and therefore represents a complete picture of evaluation use over an entire policy cycle as well as an entire evaluation cycle. Moreover, by looking at four evaluations over time, we get a more diversified idea of the effects of the evaluation system on the use of evaluation” (p. 75). Third, Hoyland’s study addressed “a collective of three cases ... (that) are of similar budget sizes for the programme cycle 2007-2013” (for details see p. 76).

Summarising, Hoyland concludes, *inter alia*, (pp. 83 ff.) that “the scope for evaluation use is considerably limited by the formal structures of the evaluation system. However, within the narrow framework of the evaluation system, evaluations are produced to be used and to increase learning in the Commission as well as outside the Commission (p. 86) ... The evaluations are used in a number of ways, ranging from making small-scale instrumental changes to the programme to more conceptual learning about the programme”.

Hoyland’s conceptually sophisticated and empirically well-grounded study points at the potentials as well as limitations of the use of evaluation in the EU’s programme cycle.

¹⁰ See footnote 4.

¹¹ In the frame of a PhD thesis at the Copenhagen Business School.

5. Summary and Conclusions

5.1. Further Expansion of Evaluation Activities

In recent years, evaluation approaches and activities have expanded further still. This holds true for the classical ex-post evaluation of policies, programmes, and measures on which our discussion in this article has focused in picking Germany, Switzerland, and the EU as cases in point. But this expansion applies to other approaches and variants of evaluation as well, not least thanks to the advances of indicator-based monitoring, benchmarking and performance management procedures and tools for which the generation and feed-back of evaluative results is pivotal.

Thus, the stock of evaluation-generated information and knowledge about the effects of policies and measures has seen an all but exponential growth which makes the call for evidence-based policy-making and for utilizing the available evaluation findings all the more mandatory and urgent.

4.2. Scope and Limits of Evaluation-Generated Knowledge Utilization

In focusing on fields and examples of classical ex-post evaluation in Germany, Switzerland, and the EU, we arrived at the conclusion that the rate of utilization of evaluation-generated knowledge has so far turned out to be mostly limited. On top of that, the referred-to studies suggest that, insofar as evaluation findings have been implemented, they have not pertained to the politically crucial decisions of policy-makers, but to minor (operational etc.) ones instead. Drawing on our earlier conceptual debate it can be said that regarding crucial decisions, the political rationality and the underlying political will of the decision-makers prevail while concerning less important decisions evaluation-generated evidence does show some effect and, hence, a dose of scientific rationality comes to the fore.

However, this assessment and interpretation arguably need some modification on sundry scores. First, as to the conclusiveness of the interpretation of the referred-to studies it should be borne in mind that they were largely guided by a conceptual lens which analytically recognises only the cases of immediately effectuated utilization of evaluation results. Hence, they tend to ignore the modes and events of utilization that may occur later in the course of the decision-making process, such as “sleeper” effects. The referred-to sketch on the legislative activities unfolding inside

a German federal ministry (Krautzberger & Wollmann, 1988) hinted at the complicated process of relevant (and possibly conflicting) information finding access (sort of trickling) into the decision-making process – in line with the conceptual utilization scheme with a dose of (scientific rationality-transmitted) “enlightenment” (Weiss, 1989).

Second, the range of the fields of ex-post evaluation discussed in this article has an admittedly limited coverage. Major policy fields (and their evaluation repertoire) have not been included in the present account. This applies not least to the field of education and university reforms on which, in the wake of the OECD’s PISA initiatives, nationally and internationally large-scale surveys and assessments have been conducted which certainly deserve attention, not least under the aspect of knowledge utilization.

Another nationally and internationally advancing evaluative track that looks promising for further discussion and research is the benchmarking scheme (Kuhlmann, 2010; Jäkel & Kuhlmann, 2013). Empirical evidence shows that benchmarking has enhanced the operational and financial transparency of administrative activities and has induced what is called “benchlearning”. However, such overture to more transparency on the operational and finance performance has so far been often limited to the intra-administrative world whereas the political decision-makers (in the local council etc.) and the general public have so far not been (or insufficiently) connected (Hollender, 2003, pp. 158 ff.).

5. Research Desiderata

The politically, administratively and socially relevant evaluation-generated knowledge has continued to grow in an all but exponential rate. It has been propelled on different evaluation tracks, be that ex-post and ex-ante evaluation, monitoring, or ongoing evaluation schemes.

Whereas the empirical evidence on the use of evaluation still appears fragmentary and ad hoc rather than systematic and comprehensive, some noteworthy advances have recently been made particularly in Switzerland and in the EU. Yet, the need is palpable to have further research on the question as to why and how evaluation results have been utilized (and why not).

University-based political/social science research and researchers have shown remarkable (and regrettable) restraint in paying attention to the issue of the utilization of evaluation-generated knowledge in the political, administrative, and social practice. Except for the emergence and blos-

soming of knowledge utilization research in the US during the 1970s and a short-lived upsurge in Germany during the 1980s (Beck & Bonß, 1989), there has since been no noteworthy engagement of political/social science researchers in this subject matter. However, the afore-mentioned major research consortium which is currently being formed at the University of Potsdam (Kuhlmann, 2016) promises, if coming to pass, to fill this gap. Political/social science research should move to close this research gap on two scores.

Firstly, the research should be resumed and pursued to empirically explore knowledge utilization on the different tracks and fields of evaluation.

Secondly, to overcome the fragmentation of research results and knowledge, the research should move forward and attempt, somewhat reminiscent of Harold Lasswell's vision of policy science (Lasswell, 1951), to systematize and generalize the findings on why and how (and why not) pertinent research-generated knowledge is utilized.

In pursuit of this agenda political/social science research should, normatively speaking, make it a point, in the tension between political rationality and scientific rationality, to advocate the latter and thus, in Aaron Wildavsky's famous book title, to "speak truth to power" (Wildavsky, 1979).

References

- Alkin, M. C. (2005). Utilization of evaluation. In S. Mathison (ed.), *Encyclopedia of Evaluation* (pp. 234-236). Thousand Oaks, USA: Sage.
- Balthasar, A. (2012). Fremd- und Selbstevaluation kombinieren: der 'Critical Friend Approach' als Option. *Zeitschrift für Evaluation*, 11(2), 173-192.
- Balthasar, A. (2008) Der Einfluss unterschiedlicher Instrumentalisierungsmodelle auf die Verwendung von Evaluationen: Erfahrungen aus der Schweiz. *Zeitschrift für Evaluation*, 7(2), 243-271.
- Balthasar, A. (2007). *Institutionelle Verankerung und Verwendung von Evaluationen. Praxis und Verwendung von Evaluationen in der Schweizer Bundesverwaltung*. Zürich, CH: Rüegger.
- Balthasar, A. (2009). Evaluationen in der Schweiz. Verbreitung und Verwendung. In T. Widmer, W. Beywl, & F. Carlo (eds.), *Evaluation: Ein systematisches Handbuch* (pp. 486-498). Wiesbaden, DE: VS Verlag.
- Balthasar, A. (2015). L'utilisation de l'évaluation par l'administration fédérale. In K. Horber-Papazian (ed), *Regards croisés sur l'évaluation en Suisse* (pp. 115-132). Lausanne, CH: Presses polytechniques et universitaires romandes.
- Balthasar, A. & Rieder, S. (2009). Wo ist evidenzbasierte Politik möglich? Die Verbreitung von Evaluationen auf kantonaler Ebene. In A. Vatter, F. Varo-

- ne, & F. Sager (eds.), *Demokratie als Leidenschaft: Planung, Entscheidung und Vollzug in der schweizerischen Demokratie* (pp. 403–429). Bern (CH), Stuttgart (DE), Wien (Austria): Haupt Verlag.
- Beck, U., & Bonß, W. (eds.) (1989). *Verwendung, Verwandlung, Verwissenschaftlichung*. Frankfurt M. (DE): Suhrkamp.
- Beywl, W. & Widmer, T., (2009). Einleitung. In T. Widmer, W. Beywl, & F. Carlo (eds.), *Evaluation: Ein systematisches Handbuch*. Wiesbaden (DE): VS Verlag.
- Bergsmann, E. & Spiel, C. (2009). Context, standards and utilisation of evaluation in Austrian federal ministries. In A. Fouquet & L. Measson (Eds.), *L'évaluation des politiques publiques en Europe. Cultures et futurs/Policy and programme evaluation in Europe. Cultures and Prospects* (pp. 187-198). Grenoble, FR: L'Harmattan.
- Bogumil, J., Grohs, S., Kuhlmann, S., & Ohm A. (2007). Zehn Jahre Neues Steuerungsmodell. Sigma Berlin Bundesbeauftragter für die Wirtschaftlichkeit der Verwaltung, 1989, Erfolgskontrolle finanzwirksamer Maßnahmen in der öffentlichen Verwaltung, Gutachten.
- Bundesbeauftragter für die Wirtschaftlichkeit der Verwaltung (1998). Erfolgskontrolle finanzwirksamer Maßnahmen in der öffentlichen Verwaltung, Gutachten, 2. Aufl.
- Cabinet Office (1999). *Modernising Government, White Paper*. London, UK: Cabinet Office.
- Campbell, Donald T. (1969). Reforms as experiments. *American Psychologist*, 24, 409-429
- Caplan, N. (1983). Knowledge conversion and utilization. In B. Holzner, K. Knorr, & H. Strasser (eds.), *Realizing social science knowledge* (pp. 50-77). Wien, Austria: Physica Verlag.
- Caplan, N., Morrison, A., & Stambaugh, R. (1975). *The use of social science knowledge in policy decisions at the national level*. Ann Arbor, USA: University of Michigan
- Danielson, C. (2007). Social experiments and public policy. In F. Fischer, G. J. Miller, & M. S. Sidney (eds.), *Handbook of Public Policy Analysis* (pp. 381-392). Boca Raton: CRC Press, Taylor & Francis Group.
- Derlien, H. (1990). Genesis and structure of evaluation efforts in comparative perspective. In C. Rist Ray (ed.), *Program Evaluation and the Management of Government* (pp. 147-176). New Brunswick (UK), London (UK): Transaction.
- Eberli, D., Bundi, P., Frey, K., & Widmer, T. (2014). *Befragung Parlamente und Evaluationen, Ergebnisbericht*. Retrieved from http://www.zora.uzh.ch/103842/1/Ergebnisbericht_ParEval.pdf
- EPEC (European Policy Evaluation Consortium) (2005a). *Study on the use of evaluation results in the commission, Final report, dossier no. 1, Synthesis report and annexes*. Retrieved from http://ec.europa.eu/smart-regulation/evaluation/docs/eval_results_comm_2005_en.pdf
- EPEC (European Policy Evaluation Consortium) (2005b). *Study on the use of evaluation results in the commission, Final report, dossier no. 2, case study reports*.

- Retrieved from http://ec.europa.eu/smart-regulation/evaluation/docs/eval_results_comm_case_2005_en.pdf
- European Commission (2001). *Annual evaluation review 2000*. Retrieved from http://ec.europa.eu/dgs/secretariat_general/evaluation/docs/eval_review_2000_en.pdf
- European Commission (2010). *Annual evaluation review 2009*. Retrieved from http://ec.europa.eu/dgs/secretariat_general/evaluation/docs/final_aer_2009_en.pdf
- Fabian, C. (2009). Nutzung der Evaluationsfunktion im Vergleich. In: T. Widmer, W. Beywl, & F. Carlo (eds.), *Evaluation: Ein systematisches Handbuch* (pp. 498-505). Wiesbaden, DE: VS Verlag.
- Furubo, J., Sandahl, R. (2002). A diffusion perspective on global development in evaluation. In J. Furubo, Ray C. Rist, & R. Sandahl (eds.), *International Atlas of Evaluation* (pp. 1-26). New Brunswick, UK: Transaction.
- Habermas, J. (1968). Verwissenschaftlichte Politik und öffentliche Meinung. In Habermas, J., *Technik und Wissenschaft als Ideologie* (pp. 20-41). Frankfurt, DE: Suhrkamp.
- Hellstern, G., & Wollmann, H. (eds.) (1983). *Experimentelle Politik – Reformstrohfeuer oder Lernstrategie*. Opladen, DE: Westdeutscher Verlag.
- Hollenrieder, J. (2004). Benchmarking für soziale Dienstleistungen im Rahmen der Arbeitsmarkt- und Sozialpolitik. In S. Kuhlmann, J. Bogumil, & H. Wollmann (eds.), *Leistungsmessung und –vergleich in Politik und Verwaltung* (pp. 150-175). Wiesbaden, DE: VS Verlag.
- Hojlund, S. (2014a). Evaluation use in evaluation systems – the case of the European Commission, PhD thesis, Copenhagen Business School. Retrieved from: http://openarchive.cbs.dk/bitstream/handle/10398/9117/Steven_H%C3%B8jlund.pdf?sequence=1
- Hojlund, S. (2014b) Evaluation use in evaluation systems – the case of the European Commission. *Evaluation*, 20, 428-442, preprint version retrieved from http://openarchive.cbs.dk/bitstream/handle/10398/9117/Steven_H%C3%B8jlund.pdf?sequence=1, pp. 158 ff.
- Jäkel, T., & Kuhlmann, S. (2014). Lernen von den Besten? Steuerung und Nutzung von Leistungsvergleichen in europäischen Verwaltungen. In S. Kropp, & S. Kuhlmann (eds.), *Wissen und Expertise in Politik und Verwaltung, der modern Staat* (Sonderheft 1/2013) (pp. 141-160). Opladen, DE: Verlag B. Budrich.
- Konzendorf, G. (2009). Institutionelle Einbettung der Evaluationsfunktion in Politik und Verwaltung in Deutschland. In: T. Widmer, W. Beywl, & F. Carlo (eds.) *Evaluation* (pp. 27-33). Wiesbaden, DE: VS Verlag.
- Korte, R. (2004). Ziele, Aufwand und lokaler Nutzen von Kennzahlen und Kennzahlvergleichen für Kommunen aus der Sicht interkommunaler Vergleichsringe. In S. Kuhlmann, J. Bogumil, & H. Wollmann (eds.), *Leistungsmessung und –vergleich in Politik und Verwaltung* (pp. 123-137). Wiesbaden, DE: VS Verlag.
- Krautzberger, M. & Wollmann, H. (1988). Verwendung sozialwissenschaftlichen Wissens in der Gesetzgebung. *Zeitschrift für Rechtssoziologie*, 2, 188-205.

- Kuhlmann, S. (2004). Interkommunaler Leistungsvergleich in Deutschland. In Kuhlmann, S., J. Bogumil, & H. Wollmann (eds.), *Leistungsmessung und –vergleich in Politik und Verwaltung* (pp. 94-122). Wiesbaden, DE: VS Verlag.
- Kuhlmann, S. (2010). Performance measurement in European local governments: a comparative analysis of reform experiences in Great Britain, France, Sweden and Germany. *International Review of Administrative Sciences*, 76(2), 1-15.
- Kuhlmann, S., Bogumil, J. & Wollmann, H. (eds.) (2004). *Leistungsmessung und –vergleich in Politik und Verwaltung*. Wiesbaden, DE: VS Verlag.
- Kuhlmann, S., Bogumil, J., & Grohs, S. (2008). Evaluating administrative modernization in German local government: success or failure of the “New Steering Model”. *Public Administrative Review*, 68(5), 851-863.
- Kuhlmann, S., & Wollmann, H. (2014). *Introduction to comparative public administration, administrative systems and reforms in Europe*. Cheltenham, UK: Edward Elgar.
- Kuhlmann, S. et al. (2016). *Das Wissen der Verwaltung, Antrag zur Einrichtung einer DFG-Forschergruppe (Entwurf)*. Potsdam, DE: Universität Potsdam, unpublished ms.
- Lasswell, H. D. (1951). The policy orientation. In D. Lerner, H. D. Lasswell (eds.), *The Policy Sciences* (pp. 3-15). Stanford, USA: Stanford University Press.
- Leeuw, Frans L. (2006). Evaluation in Europa. In R. Stockmann (ed.), *Evaluationsforschung* (3. Aufl.) (pp. 64-84). Münster, DE: Waxmann.
- Lorz, Rainer E. (1984). Das Inspektionsreferat des Bundesministeriums für wirtschaftliche Zusammenarbeit. In G. Hellstern, & H. Wollmann (eds.), *Handbuch zur Evaluierungsforschung*, (Bd. 1) (pp. 289-301). Opladen, DE: Westdeutscher Verlag.
- Luhmann, N. (1993). *Soziale Systeme*. Frankfurt, DE: Suhrkamp.
- Mader, L. (2009). Die institutionelle Einbettung der Evaluationsfunktion in der Schweiz. In T. Widmer, W. Beywl, & F. Carlo (eds.), *Evaluation* (pp. 201-220). Wiesbaden, DE: VS Verlag.
- Patton, M. Q. (1997). *Utilization-focused evaluation* (3rd ed). Thousand Oaks, USA, Sage.
- Smismans, S. (2015). Policy evaluation in the EU: The challenges of linking ex ante and ex post appraisal, *European Journal of Risk Regulation*, 1, 6-26.
- Spiel, C., & Bergsmann, E. (2009a). Zur Nutzung der Evaluationsfunktion. Partielle Bestandsaufnahme und Pilotstudie in österreichischen und deutschen Bundesministerien. In T. Widmer, W. Beywl, & F. Carlo (eds.), *Evaluation* (pp. 478-485). Wiesbaden, DE: VS Verlag.
- Spiel, C., & Schober, B. (2015). Der Beitrag wissenschaftlicher Evaluation zur Evidenzbasierung in Politik und Praxis. In V. Hennefeld, W. Meyer, & S. Silvestrini (eds.), *Nachhaltige Evaluation? Auftragsforschung zwischen Praxis und Wissenschaft* (pp. 41-52). Münster, DE: Waxmann Verlag.
- Stern, E. (2009). Evaluation policy in the European Union and its institutions. *New Directions for Evaluation*, 123, 67-90

- Stockmann, R. (ed.) (2006). *Evaluationsforschung* (3. Aufl.). Münster, DE: Waxmann.
- Stockmann, R. (2006a). Evaluation in Deutschland. In R. Stockmann (ed.), *Evaluationsforschung* (3. Aufl.) (pp. 15-46). Münster, DE: Waxmann.
- Stockmann, R. (2006b). Evaluation staatlicher Entwicklungspolitik. In: R. Stockmann (ed.), *Evaluationsforschung* (3. Aufl.) (pp. 278-416). Münster, DE: Waxmann.
- Toepel, K., Schwab, O. (2005). Evaluation in der europäischen Strukturpolitik. *Zeitschrift für Evaluation*, 1, 63-74
- Wagner, P., C. H. Weiss, B. Wittrock, & Wollmann, H. (1991). The policy orientation: Legacy and perspective. In: P. Wagner, C. H. Weiss, B. Wittrock, & H. Wollmann (eds.), *Social Science and Modern States* (pp. 2-19). Cambridge, UK: Cambridge University Press.
- Weber, M. (1922). *Wirtschaft und Gesellschaft*. Tübingen, DE: Mohr Siebeck.
- Weiss, Carol H. (1979). The many meanings of research utilization. *Public Administration Review*, 39(5), 563-574.
- Weiss, Carol H. (1991). Policy research: data, ideas or arguments. In P. Wagner, C. H. Weiss, B. Wittrock, & H. Wollmann (eds.), *Social Sciences and Modern States* (pp. 290-306). Cambridge, UK: Cambridge University Press.
- Widmer, T., W. Beywl, & F. Carlo (eds.) (2009). *Evaluation: Ein systematisches Handbuch*. Wiesbaden, DE: VS Verlag.
- Widmer, T., & F. Leeuw (2009). Die institutionelle Einbettung der Evaluationsfunktion: Deutschland, Österreich und Schweiz im Vergleich. In T. Widmer, W. Beywl, & F. Carlo (eds.), *Evaluation: Ein systematisches Handbuch* (pp. 64-71). Wiesbaden, DE: VS Verlag.
- Wildavsky, A. (1979). *Speaking truth to power*. Little Brown.
- Wollmann, H. (2002). Contractual research and policy knowledge. In *International Encyclopedia of the Social and Behavioral Sciences*, 5, 11574-11578.
- Wollmann, H. (ed) (2003). *Evaluation in public sector reform*. Cheltenham, UK: Edward Elgar.
- Wollmann, H. (2003a). Evaluation in public sector reform: Towards a "new wave" of evaluation. In H. Wollmann (ed.), *Evaluation in Public Sector Reform* (pp. 1-11). Cheltenham, UK: Edward Elgar.
- Wollmann, H. (2003b). Evaluation and public sector reform in Germany: Leaps and lags. In H. Wollmann (ed.), *Evaluation in Public Sector Reform* (pp. 118-132). Cheltenham, UK: Edward Elgar.
- Wollmann, H. (2003c). Evaluation in public sector reform. Trends, potentials and limits. In H. Wollmann (ed.), *Evaluation in Public Sector Reform* (pp. 231-258). Cheltenham, UK: Edward Elgar.
- Wollmann, H. (2007). Policy evaluation and evaluation research. In F. Fischer, G. J. Miller, & M. S. Sidney (eds.), *Handbook of Public Policy Analysis* (pp. 393-404). CRC Press. Boca Raton etc.

- Wollmann, H. (2009). Kontrolle in Politik und Verwaltung: Evaluation, Controlling und Wissensnutzung. In K. Schubert, N. Brandelow (eds.), *Lehrbuch der Politikfeldanalyse 2.0* (2. Aufl.) (pp. 379-400). München, DE: Oldenbourg Verlag.
- Wollmann, H. (2013). Evaluation in Politik und Verwaltung, Ansätze, Ergebnisse, Verwendung. In: C. Reichard, & E. Schröter (eds.), *Zur Organisation öffentlicher Aufgaben* (pp. 362-386. Leverkusen, DE: Verlag B. Budrich.
- Wollmann, H. (2014). Zur (Nicht-) Verwendung von Evaluationsergebnissen in Politik und Verwaltung, Eine vernachlässigte Fragestellung der Evaluationsforschung. In: S. Kropp, & S. Kuhlmann (eds.), *Wissen und Expertise in Politik und Verwaltung, der moderne staat* (Sonderheft 1/2013) (pp. 87-102). Leverkusen, DE: Verlag B. Budrich. .
- Zintl, M. (2009). Evaluierung in der deutschen Entwicklungszusammenarbeit. In T. Widmer, W. Beywl, & F. Carlo (eds.), *Evaluation: Ein systematisches Handbuch* (pp. 245-255). Wiesbaden, DE: VS Verlag.

UTILIZATION OF EVALUATION RESULTS IN POLICY-MAKING AND ADMINISTRATION – A CHALLENGE TO POLITICAL SCIENCE RESEARCH

Summary

The paper deals with utilization of evaluation results in policy-making and public administration. In the pursuit of this guiding question, the article is written in five steps. First, major variants of evaluation are presented with a particular eye to their respective utilization potential. Second, different concepts are sketched that have been forwarded to capture the utilization of (social) science-generated knowledge by political, administrative, and social actors. Third, looking at Germany, Switzerland and the European Union as cases in point pertinent research findings are discussed. Parts four and five contain summary, conclusions and some remarks on ensuing research needs. The political rationality and the underlying political will of the decision-makers prevail in crucial political decisions while in less important decisions evaluation-generated evidence does show some effect and a dose of scientific rationality comes to the fore. However, the stock of evaluation-generated information and knowledge about the effects of policies and measures has seen an exponential growth, which makes the call for evidence-based policy-making and for utilizing the available evaluation findings all the more mandatory and urgent.

Keywords: evaluation, policy-making, public administration, utilization of evaluation results

KORIŠTENJE REZULTATA EVALUACIJSKIH STUDIJA U KREIRANJU JAVNIH POLITIKA I JAVNOJ UPRAVI – ISTRAŽIVAČKI IZAZOV U POLITIČKIM ZNANOSTIMA

Sažetak

Rad se bavi korištenjem evaluacijskih studija u kreiranju javnih politika i javnoj upravi. Tekst je organiziran u pet dijelova. U prvom su dijelu opisane glavne varijante evaluacije s posebnim naglaskom na njihovu potencijalnu iskoristivost. Drugi se dio bavi različitim konceptima razvijenim da bi se utvrdilo kako politički, upravni i društveni akteri koriste znanje nastalo u društvenim znanostima. U trećem se dijelu raspravlja o važnim nalazima istraživanja u Njemačkoj, Švicarskoj te u tijelima Europske unije. Četvrti i peti dio sadrže glavne nalaze, zaključke te neke napomene o potrebama budućih istraživanja. Politička racionalnost i politička volja donositelja političkih odluka na kojoj se ta racionalnost temelji prevladavaju kad je riječ o najvažnijim odlukama, dok rezultati evaluacija i određena mjera znanstvene racionalnosti dolaze do izražaja kada je riječ o manje važnim odlukama. Količina informacija i znanja o učincima javnih politika i mjera dobivenih evaluacijom eksponencijalno raste, što zahtjeva za kreiranjem javnih politika temeljem dokaza i za korištenjem raspoloživih evaluacijskih nalaza čini to više hitnim i obveznim.

Ključne riječi: evaluacija, kreiranje javnih politika, javna uprava, korištenje rezultata evaluacije