



Evaluation use in evaluation systems – the case of the European Commission

Evaluation

2014, Vol. 20(4) 428–446

© The Author(s) 2014

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/1356389014550562

evi.sagepub.com**Steven Højlund**

Copenhagen Business School, Denmark

Abstract

This article investigates the European Union's evaluation system and its conduciveness to evaluation use. Taking the European Commission's LIFE programme as its case, the article makes an empirical contribution to an emerging focus in the literature on the importance of organization and institutions when analyzing evaluation use. By focusing on the European Union's evaluation system the article finds that evaluation use mainly takes place in the European Commission and less so in the European Parliament and the European Council. The main explanatory factors enabling evaluation use relate to the system's formalization of evaluation implementation and use; these factors ensure evaluation quality, timeliness and capacity in the Commission. At the same time, however, the system's formalization also impedes evaluation use, reducing the direct influence of evaluations on policy-making and effectively 'de-politicizing' programme evaluations and largely limiting their use to the level of programme management.

Keywords

European Commission, evaluation systems, evaluation use, LIFE programme, programme management

Introduction

In the last 30 years, evaluation practices have spread and become common practice in most OECD countries. In parallel with the spread of evaluation practices, national and international organizations have to a large extent institutionalized and ritualized evaluation practices into what has been termed 'evaluation systems'. Prior research has hypothesized about the implications of evaluation systems on evaluation use but the phenomenon still needs more empirical investigation (Furubo, 2006; Leeuw and Furubo, 2008; Rist and Stame, 2006).

Corresponding author:

Steven Højlund, Department of Business and Politics, Copenhagen Business School, Steen Blichers Vej 22, 2000 Frederiksberg, Denmark.

Email: sho.dbp@cbs.dk

This article investigates evaluation use in the European Union's (EU) evaluation system where evaluation practices have been institutionalized over several decades, particularly in the European Commission. Thereby the article aims to contribute empirically to the emerging focus in the evaluation literature on contextual organizational factors explaining evaluation use (Højlund, 2014). It does so by focusing on the evaluation system understood as the institutionalization of evaluation practices in the EU organizational bodies – in particular the Commission. The evaluation system thus becomes the systemic setting and context in which evaluation use is analyzed. The underlying assumption is therefore that the attributes of the evaluation system can explain the way evaluations are used in this particular system. Thereby the article relies on newer theoretical contributions on evaluation systems (e.g. Leeuw and Furubo, 2008) as well as a more general introduction of organizational theory into the theoretical landscape of evaluation use (Højlund, 2014).

The main contribution of the article is to improve our understanding of the implications that an evaluation system has for evaluation use. Formal and informal organizational practices both impede and enable the use of evaluation. Despite some evaluation use by policy-makers, the article finds that most use takes place at the level of programme management in the Commission. Thus evaluation use at the programme level tends to be instrumental, strategic, legitimizing and informational, whereas policy-makers use evaluations strategically and to get information, albeit to a much lesser extent overall.

This article reports five overall findings. First, the strong formalization of evaluation practices in the system enables 'findings use' but impedes 'process use'. Evaluations are thus typically used after their completion and not during their implementation, due to the Commission's stress on the independence of the evaluator. Second, significant findings use typically takes place at 'decision points' every seventh year in the programming phase. Other uses do take place at the programme management level, but instrumental use that affects the programme or other policies is typically indirect as evaluations feed into impact assessments and ex-ante evaluations of the new programme. Third, evaluations have little overall relevance for policy-makers and programme management alike. In particular, evaluations are not relevant for policy-makers outside the Commission due to competing information and the technical nature of evaluations. Fourth and for the above reasons, programme evaluations are 'de-politicized' and generally they are not something policy-makers participate in nor have any use for. Fifth and finally, the 'de-politicization' represents a paradox since it is the European Parliament and Member States in the European Council that required the Commission to set up the evaluation system and also demand evaluations to be a part of the legal basis of programmes such as LIFE. But this article shows that policy-makers rarely use the evaluations, while at the same time the Commission is burdened by the evaluations and numerous other internal and external assessments and audit studies.

The article is organized as follows. In the first sections, evaluation systems and evaluation use are defined and discussed. Then a section follows describing the methodology used in the analysis as well as the analysis itself. Finally, a conclusion is followed by reflections on potential extensions of the research on evaluation systems and evaluation use.

Evaluation systems

The discussion on evaluation systems took a leap forward with the book *From Studies to Streams* edited by Ray C. Rist and Nicoletta Stame in 2006. Several subsequent studies picked up the baton (Imam et al., 2007; Leeuw and Furubo, 2008; Williams and Imam, 2008) improving our

conceptual understanding of the phenomenon. The literature on evaluative systems relate to a broader focus in the evaluation literature on evaluation as a phenomena understood as contingent on complex societal contexts such as institutions, norms and power (Dahler-Larsen, 2012; Van der Knaap, 1995). Particularly, Peter Dahler-Larsen has used institutional organizational theory to explain the phenomena of evaluation and adaptation of evaluative practices by public organizations. Only recently has the same theoretical framework been used explicitly to explain the phenomena of evaluation use (Højlund, 2014).

This article builds on the fundamental idea that institutions and organization determine evaluation use. The focus in the article is on evaluation systems because an evaluation system is composed of several organizational entities that to some degree share formal and informal evaluative practices and norms (i.e. a shared evaluation institution). Leeuw and Furubo (2008) stress the following four elements constituting an evaluation system:

1. Participants in the evaluation system share a common understanding of the objectives of evaluation and the means by which the objectives are attained.
2. The evaluation system is institutionalized formally in at least one organizational structure, in which it is separated from the operational structure of this organization. Hence, the system has at least one formal institutionalized organizational element (e.g. 'an evaluation unit') that typically is in charge of planning, tendering, implementing, quality-checking and following-up on evaluations.
3. Evaluation systems are permanent in the sense that their setup has no time-limitation. Moreover, evaluations are undertaken continuously and systematically and in relation to previous and future evaluations.
4. In the evaluation system, evaluations are organized and planned so that they relate to the cycle of activities of the organization or the evaluand (e.g. budget or policy cycle).

Based on the four elements above and other contributions, a definition of an evaluation system could be summarized as follows: an evaluation system is permanent and systematic formal and informal evaluation practices taking place and institutionalized in several interdependent organizational entities with the purpose of informing decision making and securing oversight.

In relation to evaluation use, evaluation systems are generally assumed to have a negating effect on information- and knowledge-use in policy-making (Leeuw and Furubo, 2008; Pollitt et al., 1999). Previous studies suggest that evaluative knowledge tends to be made relevant primarily for administrators and not for policy-makers and that use in administrations will be linked to procedural assurance and legitimization of the organization rather than used to inform policy-making (see also Furubo, 2006; Langley, 1998). The purpose of this article is to continue the research on evaluation systems' effect on evaluation use and provide empirical evidence where presently there is little.

The EU organizational setup constitutes a very good example of an evaluation system (Stern, 2009). Evaluation is an integral part of the activity-based management and budgeting system of the Commission and thus formally related to decision making regarding EU budgetary allocations. The system's core consists of the European Commission (the Commission), the European Parliament (EP) and the European Council (the Council). As the EU executive body, the Commission is also responsible for commissioning, implementing and disseminating evaluations of EU programmes and policies. The Commission has a legal obligation to evaluate programmes and policies as stipulated in the Commission's management policies as well as the

legal basis of the programmes and policies. For this reason, the Commission has institutionalized evaluation practices over the last 30 years in each Directorate General (DG) through evaluation policies, guidelines and standards. In the DGs, designated evaluation units supervise and guide evaluation activity with support from the Secretariat-General. The evaluation units are subject to internal audits as they are described in the Internal Control Standards of the Commission services. It is the Commission that undertakes most evaluations in the system, but the EP and Member States do also carry out or commission evaluations, usually subject to EU evaluation standards and supervised by the Commission (in the case of Member States). About 80 percent of all evaluations in the Commission are externalized to consultants or groups of experts (Commission, 2007) and the consultancies are thus also a part of the evaluation system.

Evaluation use

In the 1960s, scholars started to question whether knowledge is used to inform policy-makers in order to improve policy (Lazarsfeld et al., 1967; Weiss and Bucuvalas, 1980). The answer to this question was partly negative and the situation was referred to by some scholars as a 'utilization crisis' (Floden and Weiner, 1978; Patton, 1997). In the cases where evaluation information was actually used, evaluation research conceptualized use-categories, which have not changed significantly over the years (Leviton, 2003). Four main types of evaluation use emerged: instrumental-, conceptual-, process- and symbolic use. These four categories are still used as the basis for most research, though newer and more elaborate conceptual frameworks have been suggested (Henry and Mark, 2003; Kirkhart, 2000; Weiss, 1998).

In the wake of the disenchantment associated with the scarce evidence of use from evaluations, the literature instead asked why evaluations were used or not used (Cousins and Leithwood, 1986; Leviton and Hughes, 1981). Studies focused on factors related to the attributes of the evaluation (e.g. methodology, quality, relevance of findings etc.) or the immediate contextual factors pertaining to the organization in which the evaluation is implemented (e.g. political climate, timing of the evaluation relative to decision making etc.). These categories were empirically informed from the late 1970s and onwards (see, for example, Leviton and Hughes, 1981). This article leans on the broad definition of evaluation use provided by Johnson et al. (2009: 378): 'any application of evaluation processes, products, or findings to produce an effect.' This definition captures the variety of use types applied in this article (see Research question and design).

In relation to the interest in evaluation systems, Furubo (2006: 151) suggests that the literature could still benefit from a better understanding of the effects that evaluation systems have on evaluation use. In general, it seems that most evidence on evaluation use is still linked to single ad hoc evaluation studies rather than systematic evaluation information and does not specifically address the evaluation system. It is on this topic that this article makes its contribution. Similarly, only very few studies in the evaluation literature take into account organizational explanations when analyzing evaluation use (Højlund, 2014).

Research question and design

This article investigates whether *evaluation systems are conducive to evaluation use*. In order to properly answer this question, three sub-questions are proposed: 1) *how* are evaluations used in evaluation systems? 2) *who* uses evaluation findings in evaluation systems? 3) *why* do – or do not – evaluation systems support the use of evaluation findings?

Table 1. The 10 evaluation use types.

Process use	Findings use
(evaluation use during the process of evaluation; typically use of preliminary results etc.)	(evaluation use after the evaluation process has ended; typically use of the findings and recommendations of a report)
<ul style="list-style-type: none"> - Instrumental: The evaluation findings are used to change the evaluand or the conditions that it is working under. - Conceptual: The evaluation is used to gain new conceptual knowledge. - Symbolic: The evaluation is used to legitimize the organization that is responsible for the evaluand. - Information: The evaluation is used to acquire information. - Strategic: The evaluation is used for advocacy. 	<ul style="list-style-type: none"> - Instrumental: The evaluation findings are used to change the evaluand or the conditions that it is working under. - Conceptual: The evaluation is used to gain new conceptual knowledge. - Legitimizing: The evaluation is used to legitimize the evaluand. - Information: The evaluation is used to acquire information. - Strategic: The evaluation is used for advocacy.

Consequently, the *dependent factors* are evaluation uses. Considering evaluation use, the article distinguishes between 10 different types of evaluation use organized under two headings: ‘findings uses’ (instrumental, conceptual, legitimizing, information and strategic) and ‘process uses’ (instrumental, conceptual, symbolic, information and strategic) (Alkin and Taut, 2003; Leviton, 2003; Leviton and Hughes, 1981; Weiss, 1998). Table 1 below gives an overview of the 10 types of evaluation use in the analysis.

The 10 use categories are informed by existing literature on evaluation use. Hence, Alkin and Taut (2003) proposed the conceptual division between findings use and process use as they recognized that process use (use of the evaluation *during* the evaluation process) was not a type of use in itself as it could both be instrumental, conceptual and legitimizing (e.g. evaluation is legitimizing the organization).¹ In addition to instrumental-, conceptual-, legitimizing- and symbolic uses, the evaluation literature has also proposed two other categories of uses that relate to the use of evaluation understood as simply a source of information – a type of use that often precedes other use forms (Alkin and Stecher, 1983; Finne et al., 1995). An instance of information use would be to use evaluation information in a presentation or simply reading the evaluation to acquire knowledge. ‘Information use’ can take place both before and after the completion of the evaluation and is thus both related to ‘findings use’ and ‘process use’.

Finally, scholars have pointed to a fifth type of use often referred to as ‘strategic use’. Strategic use is distinguished from symbolic and legitimizing use types as it is not related to securing organizational or programme legitimacy, but rather to advocacy in relation to decision- or policy-making (Pröpper, 1987: cited in Van der Knaap, 1995: 211; Weiss, 1992). Strategic use needs to be included because legitimizing use, originally proposed by Rich (1977), does not appropriately cover the strategic and political use of arguments found in evaluations and used to justify political arguments and decisions. Legitimizing use is the evaluating organization justifying the programme or policy that is evaluated. However, in an evaluation system there are more actors involved, who have an interest in using the evaluation as a source of legitimacy to back their positions and political arguments. This type of use I call ‘strategic use’ as it does not necessarily have to be related to legitimizing the programme (legitimizing use)

or the justification of the evaluating organization (symbolic use). Instead it is related to other issues, such as when facts from the evaluation are used to back a certain position in the renegotiation of a new programme.

The overall *independent factor* is the context of the EU evaluation system. However, to better understand the processes in play, the analysis contains intermediate explanatory factors providing for a more detailed understanding of barriers and enablers of evaluation use in the evaluation system. Here, the article relies on the conceptual framework of Cousins and Leithwood (1986) and Johnson et al. (2009). They refer to 12 specific factors that can influence evaluation use. These factors are divided into two categories. The first category is 'evaluation implementation' (a. evaluation quality, b. evaluation credibility, c. evaluation relevance, d. communication quality, e. evaluation findings, f. timeliness), and the second one is 'decision or policy setting' (a. information needs, b. decision characteristics, c. political climate, d. competing information, e. personal characteristics, f. receptiveness to evaluation). The first category relates to traits about the evaluation in question. The second relates to factors linked to the organizational decision making and other contextual factors not directly linked with the evaluation.

Data and methodology

This article analyzes the use of four evaluations of the Commission's Programme for the Environment and Climate Action (LIFE) over a 10-year period (2003–13). The case is thus the EU's LIFE programme. Case studies like this one are common to the evaluation literature and mirror the fact that interventions and their evaluations are often uniquely tied to a particular organizational or systemic context as is the case here (Easterby-Smith et al., 2000). The EU evaluation system is a well-constituted evaluation system that matches the definition of an evaluation system as described earlier.

The choice of the LIFE programme as the case was made because of data availability and because the LIFE programme has experienced a full Commission evaluation cycle (ex-ante, midterm, final and ex-post) and therefore it represents a complete picture of evaluation use over an entire policy cycle as well as an entire evaluation cycle. Further, evaluation use in the Commission has been given little attention by researchers so far (see Bienias and Iwona, 2009; Zito and Schout, 2009) with the exception of two Commission-sponsored reports (Laat, 2005; Williams et al., 2002). This is unfortunate, because the Commission is important in terms of spreading evaluation practices in Europe (Furubo et al., 2002; Toulemonde, 2000).

The analysis is based on 16 semi-structured in-depth interviews and eight follow-up interviews. The informants were sampled purposefully according to relevance and availability and consisted primarily of staff from the Directorate General for the Environment (DG ENV), consultants that performed the evaluations, representatives of Members of the EP's Committee for the environment (ENVI-Committee) and Council members (Ritchie et al., 2003). In addition, 36 background interviews were conducted with Commission staff working in other DGs on other EU programmes to qualify the information and understand the evaluation system. The analysis also included relevant documents such as the four retrospective evaluations of the LIFE programme (Midterm, 2003; ex-post, 2009; midterm, 2010; final, 2013) and several other documents including DG ENV presentations to the Committee of Regions and the EP, internal Commission documents, the combined ex-ante and impact assessment (IA) along with explanatory policy fiches and Commission position papers for the new LIFE programme 2014–20.

The methodology applied in the article was based on the principles of qualitative content analysis (Mayring, 2000; Schreier, 2012) and the actual coding and analysis of data was carried out using the NVIVO software package (Bazeley, 2013). The first 16 semi-structured interviews were analyzed with a view to existing conceptual frameworks developed in the evaluation literature and described earlier. The eight follow-up interviews were conducted to check for saturation. The semi-structured interview guides gave the interviewees flexibility to elaborate on evaluation use and explanatory factors in relation to the evaluation in question and to the extent the interviewee found it relevant. Coder reliability was sought by using the existing conceptual frameworks and subsequently running three rounds of coding on the interview data (Kohlbacher, 2006; Mayring, 2004). Further, the credibility of the findings was strengthened by a prolonged engagement in the field, conducting interviews in four consecutive waves over a period of one year. Findings were triangulated and validated with document data and follow-up interviews and interpretations were checked against interview data. Interviewees were debriefed and had the opportunity to comment on the findings of the article, and peers with comprehensive knowledge on the subject gave important comments on the draft article before submission. Finally, the researcher has several years of experience with evaluation of EU programmes including work as an evaluator on the ex-post evaluation of LIFE.

Analysis

The analysis is divided into three sections in answer to the three research questions. The first section is dedicated to the use of four LIFE evaluations produced between 2003 and 2013. The second section treats the explanations for evaluation use and the final section in the analysis summarizes who the users of the LIFE evaluations are. In each section the findings are recapped at the end of the section.

Uses of LIFE evaluations

This section contains an analysis of the process uses as well as the findings uses of the LIFE evaluations. Table 1 in the supplementary data (available at <http://evi.sagepub.com/supplemental>) summarizes the distribution of the qualitative codes on three groups of interviewees and provides examples of interviewee quotes. Most notable is the lack of process use, but there are several interesting patterns in the findings uses as well that are described and summarized below.

Process use. The interview data rarely contains references to process use. The few accounts of process use in the data concerns mainly strategic use. In the evaluation process, key stakeholders have the opportunity during the evaluation implementation to influence evaluation findings by coordinating answers to interview questions or raising particular issues of concern in interviews. The impact from this can be directed towards short-term decision making within the Commission as well as towards programme change. Additionally, the evaluation system makes a strong link between evaluations inducing evaluators to build on previous findings. Evaluations are therefore also used strategically in the long run as issues raised in consecutive evaluations gain prominence in decision making. Also, symbolic use of evaluations were found to play a role in the evaluation process as DG ENV as well as other DGs are concerned about reducing negative findings about their organizations in the evaluation.

Apart from these few instances of evaluation process use, there were no accounts of process uses in the data. There are two main reasons for this finding. First, the evaluation process is carried out mainly by the external evaluator in relative seclusion from potential users in the Commission or in other parts of the evaluation system. The evaluation process is typically managed by one desk officer in DG ENV, who is the liaison between the evaluator and the Steering Committee that oversees the evaluation at regular intervals (five to seven meetings during the evaluation process). This process is standard in the Commission and is meant to secure the independence of the evaluation as well as the proper and efficient evaluation execution. However, it also limits use in the evaluation process, because the potential users are rarely directly involved in the evaluation activity. Second, the evaluation findings can rarely be put directly to use during the evaluation process in DG ENV. Whether process findings are instrumental or symbolic, the use would normally require the evaluation to be finalized in order for them to be used for **instrumental use** and symbolically as well. DG ENV is expected to evaluate as it is stipulated in the LIFE Regulation. Flagging evaluation activity as a symbolic act during the evaluation process to gain external legitimacy in the system is therefore not necessary, as evaluation is expected by the organizational environment. One interviewee from the Commission put it this way: ‘the legitimacy [from evaluations] is by now almost automatic . . .’. However, failing to evaluate or delaying evaluation activity would be perceived negatively by other actors in the EU evaluation system.

The LIFE evaluations analyzed in this article took place every two to three years. Therefore the **conceptual use** of LIFE evaluations related to methods and programme-related concepts was limited, as evaluation practice and findings were repetitive. The staff implementing the evaluation were also very knowledgeable about the LIFE programme and therefore there is also little **information use** during the evaluation implementation.

Findings use. LIFE evaluations are mainly used **conceptually** in the Commission, whereas policy-makers do not mention conceptual use in the interviews. Evaluations are used conceptually to inform discussions in DG ENV about evaluation methodology, programme indicators and the purpose of evaluation and programme impacts. The Commission’s own ideas and opinions about the programme are clarified and sharpened by evaluations. This includes findings that might go against the conventional wisdom of the programme management such as the feasibility of indicator systems across the heterogeneous project portfolio of the LIFE programme.

Instrumental use is linked intrinsically to programme management in the two LIFE units in DG ENV. Instrumental use takes place to a large extent based on the evaluation’s recommendations through follow-ups and most recommendations are addressed after evaluation implementation. Examples of instrumental use could be work optimization or improved internal communication between the LIFE units and the thematic units in DG ENV. However, fundamental changes to the programme administration are rare because they require changes to the LIFE Regulation or the Commission’s Financial Regulation. Also, the programme has been optimized over twenty years and thus several interviewees argue that most workable options have already been tried out.

On the other hand, the data suggests that policy-makers do not use LIFE evaluations instrumentally. That is, evaluations do not directly inform policy-making on the policy-level through, for example, the evaluation’s recommendations. Policy-makers do not in general read the LIFE evaluations, so evaluations feed into the policy cycle through instrumental use on the programme management level, where evaluations are the knowledge base for policy development

in, for example, ex-ante evaluations and IAs. Thus, the effect of evaluations on policy is indirect. Additionally, evaluations are also used to update the evaluation system based on knowledge and experience linked to evaluation practices. This information is collected by the Commission's Secretariat General. As a result of changes to evaluation practice, a final evaluation is no longer a requirement in the new LIFE Regulation (2014–20).

Information use was referred to by all interviewees. NGOs and potential beneficiaries as well as people new to the programme use evaluations to get an overview and understanding of the way the programme works. MEPs use the Commission's presentations but also the evaluation documents needed to get up to speed with the LIFE programme. When a decision has to be made regarding LIFE, they read the executive summary and browse or search keywords in the document. Also, MEPs and particularly the LIFE units in DG ENV use evaluations for presentations and to communicate about LIFE. Evaluations typically contain aggregated information and facts about the programme as well as graphs and figures, which can readily be used for presentation purposes. However, evaluations seldom contain information that is completely new to the programme management. Finally, evaluations are systematically referred to in evaluation tender material and so evaluation findings and recommendations are used by other evaluators in subsequent evaluations.

Information use is linked closely to legitimizing and strategic uses. Using evaluations to legitimize the programme is a common practice, mentioned both by programme management and policy-makers. **Legitimizing use** is most common in relation to informational practices such as when DG ENV is required to report evaluation findings or make presentations to the EP, the Council or the LIFE Committee (consisting of Member State representatives). The extent to which these activities are related to legitimizing use depends very much on the timing of the evaluation. Flagging that the Commission is doing a good job is important. One key informant in DG ENV described it in the following way: 'our evaluation reports and results are used and when we have them, it is something that people are happy about, because they prove that we in most cases are doing a good job.' Legitimizing use, however, is not the most common use type because many of the recommendations are programme-specific and not orientated towards legitimizing the programme on a political level. Interviewees specified that the evaluation system carries legitimacy by default and therefore one LIFE evaluation does not add significantly to the legitimacy of the programme. Rather, it is the EU evaluation system that gives legitimacy to the LIFE programme and DG ENV. Finally, the LIFE programme is very popular among stakeholders and not very contested. Evaluations of larger EU expenditure programmes as well as policies may be considerably more contested resulting in more legitimizing evaluation use.

Strategic use of the LIFE evaluations coincided with legitimizing use, particularly when DG ENV presents evaluation results to the EP and the Committee of Regions in relation to the new LIFE programme proposal. In addition to this, DG ENV uses the evaluations as a reference document to raise issues and problems about the programme that they are aware of, but which are easier for them to communicate with an independent evaluation. This is done externally to the EP and Council but also internally to other units in the DG. Internally, other units in DG ENV needed to be convinced about certain directions of the programme. The Members of Parliament (MEPs) and Member States used the LIFE evaluations as a basis for arguments in the negotiation for the new LIFE programme and in questions to the Commission. The MEPs can base opinions on evaluation data or findings. Occasionally, lobbyists raise issues to MEPs stemming from an evaluation that the MEP might in turn take up with the Commission.

Summary of uses of LIFE evaluations. The main finding of the analysis of evaluation use is that process use almost never takes place. On the other hand, several instances of findings use were discovered. The data suggests that after the LIFE evaluations were implemented, instrumental- and information uses were the most common, followed by strategic use. Legitimizing and conceptual uses were mentioned less often by interviewees. At the level of programme managers, instrumental use happens to adjust the programme within the limitations of the LIFE Regulation and the Commission's Financial Regulation. Also, instances of legitimizing and strategic uses were found in the data related to programme managers.

Regarding policy-makers, evaluations were used to acquire information (information use) and to advocate certain policy issues or positions (strategic use). Data suggests that programme evaluations are not used in a direct instrumental way by policy-makers for the purpose of making significant programme changes. Rather, they are used indirectly to support IAs and policy fiches in the negotiations for the new programme at the end of each programme cycle. Evaluators, who have an informed outsider's look at things, mainly refer to information use and instrumental use of the evaluations.

Explaining evaluation use of LIFE evaluations

This section contains the analysis of factors explaining evaluation use observed from the LIFE evaluations. Table 2 in the supplementary data (available at <http://evi.sagepub.com/supplemental>) gives an overview of the factors that explain evaluation use. The table summarizes the distribution of the qualitative codes for three groups of interviewees and provides examples of interviewee quotes. The findings from the data are described and summarized below.

Decision and policy setting. Regarding the explanatory factors relating to *Decision and policy setting*, the EU evaluation system has institutionalized a high level of **receptiveness and commitment** to evaluation in the Commission and also in DG ENV. Evaluation practice is considered part of the administrative practice in DG ENV including obligations towards the other organizations in the system, which again creates commitment. Overall, the interviews and desk research point to an organizational willingness to use evaluations. However, receptiveness in relation to LIFE programme evaluations is largely limited to the desk officers involved in programme management within the Commission. Interest and knowledge of LIFE evaluations in the EP and Council is very low, mainly due to low relevance of evaluation information to the work of policy-makers.

There is a lot of **competing information** regarding LIFE evaluations. This is one factor that impedes evaluation use both for programme managers and policy-makers. Competing information primarily includes previous evaluations of LIFE (either produced by the Commission or, for example, by the European Court of Auditors). But the experience that programme managers have managing the LIFE programme should also be considered as competing information. It is simply very difficult for consultants to bring about new information, new findings or produce knowledge that the experienced programme staff do not already have. Finally, evaluations are relatively broad in scope, covering the entire programme. In terms of use, specific studies about parts of the programme are more focused on key contested areas, which according to interviewees make them more likely to be used than the broadly-scoped evaluations. At the same time, however, interviewees consider LIFE evaluations to be the best general source of assessment information about the LIFE programme.

Table 2. Explanatory factors.

Explanatory factors	Explanation
Decision and policy setting	
Commitment and receptiveness	Staff commitment, receptiveness, responsiveness etc. to evaluations, evaluation procedures and practices.
Competing information	The influence of other reports, studies and prior knowledge.
Decision characteristics	The influence of the procedures and practices of decision-making including the barriers and enabling factors that are related to decision-making (e.g. timing, legal framework etc.)
Information needs	The influence of new information on the performance and of the organization.
Personal characteristics	The influence of the involved person's personalities and experience.
Political climate	The influence of the saliency of an issue, political or public focus.
Evaluation implementation	
Timeliness	Timeliness of reports and other deliverables in the evaluation implementation.
Credibility	Perceived credibility of the evaluation overall as well as findings.
Evaluation quality	Overall perception of evaluation quality, soundness of methods and methodology.
Findings	Saliency of findings, conclusions and recommendations.
Relevance	Overall relevance of the evaluation including its methodology, methods and evaluation questions.
Communication quality	Quality of communication in the evaluation deliverables.

According to many interviewees, the **decision characteristics** of the EU institutions do to a large extent impede direct evaluation use, particularly on the level of programme management. The seven-year budget period in the EU makes substantial and incremental programme changes difficult. Also, political decisions about the programme are complex and depend on several political actors (EP and Council) potentially with very diverse political interests and priorities. Nevertheless, evaluation use is an integrated part of the evaluation system every seventh year when the new programme is being prepared. An IA must draw on available information, including evaluations, and the importance of the budget decision draws attention to all available information on LIFE. Thus the EU policy cycle – which the evaluation system is designed to feed information – both enables and impedes evaluation use.

In relation to evaluations of the LIFE programme, the programme management also has **information needs**, despite the fact that the programme management is by far the most knowledgeable in relation to the programme. Nevertheless, some specific parts of the programme might be unknown even to the programme management, in which case an evaluation can be used to shed light on the issue. In the EP and the Council, the need for evaluation information is much lower, mainly because policy-makers need information that communicates key information about issues related to decision making. This is rarely considered to be the case for programme evaluations of the LIFE programme.

Interviewees did not often mention **personal characteristics** in relation to evaluation use. DG ENV has substantial experience with evaluation over several decades. As the practice of evaluation is highly institutionalized, the LIFE evaluations were carried out and supervised according to the Commission's general evaluation policy and standards. Both the Commission

and the external evaluators had sufficient capacity to manage and carry out the evaluations. The data suggests that the institutionalized practice of the evaluation system has a positive effect on evaluation use because staff are trained and used to working with evaluations, which in turn influences the quality of the evaluations positively.

Few interviewees mentioned the **political climate** as an explanatory factor in relation to evaluation use. Evaluations are more likely to be used to justify positions, if the programme or an issue is contested politically (Cousins and Leithwood, 1986). But as the LIFE programme is an overall popular and well-run programme, the political climate is not important for the use of evaluation findings in this case. As long as evaluations are not very negative, evaluations are used mainly as information in the IAs. Hence, evaluations are not brought into policy-making directly but only as secondary information after IAs and fiches in the programming of the new LIFE programme.

Evaluation Implementation. Considering the explanatory factors linked to *Evaluation implementation*, the interviewees mentioned these factors considerably less than the decision and policy setting. Very few interviewees mentioned **timeliness** as an important factor. Similar to the personal characteristics, the evaluation system produces evaluations with a high degree of professionalism and timeliness. Evaluation management in the Commission is supervised by the evaluation unit and the capacity to manage evaluations is highly developed in the organization. At the same time, evaluators are carefully selected through tendering. Due to the institutionalization of evaluation practice in the evaluation system, timeliness is less of an issue. However, timeliness is important for the evaluation system because of the policy cycle, and therefore securing timely delivery of evaluations has a positive impact on evaluation use.

The **credibility** of evaluations is considered crucial for the use of evaluations in the Commission. To several of the interviewees in DG ENV, the credibility of Commission evaluations is related to the credibility of the Commission itself as well as the high quality of evaluation work and independence of evaluators. The high credibility of Commission evaluations enables the justificatory uses of the LIFE evaluations towards the EP and Council, particularly in the renegotiation of a new programme. This is induced by the way the evaluation system is constructed, with a stress on external evaluation and thus a relative independence in the evaluation implementation.

The situation is similar in relation to **evaluation quality**. The evaluation quality is perceived by interviewees to be high and is also considered important for evaluation use. Contrary to the enabling factor played by the evaluation quality in the evaluation system, interviewees see evaluation **findings** as constrained by the Commission's evaluation guidelines and standards. The Commission's evaluation tenders include evaluation questions as well as a relatively rigid format for the evaluation process including a specific number of meetings with the Steering Committee. Some interviewees argue that this procedural format leaves little room for innovation during the evaluation process as well as in evaluation findings because both the process and the findings are scoped and framed relatively narrowly by the tender material which is more or less standardized in the common evaluation guidelines. In turn, this impedes evaluation use as evaluators are constrained in the process and findings seldom are surprising or completely innovative to the Commission. The lack of innovative recommendations is also due to knowledge-asymmetry between the evaluator and programme management. This is linked to the high complexity of the LIFE programme and the decision characteristics mentioned earlier, which often render innovative recommendations useless due to legal constraints on the Commission that can render innovative recommendations useless.

The **relevance** of evaluations is considered by interviewees to be very important for evaluation use of the LIFE programme. In the EU evaluation system, the relevance of evaluations is closely linked to the timing of the evaluation relative to decision-situations in which the evaluation can be used. That is crucial both on the political level as well as the administrative level in the Commission. If the timing of the evaluation does not match the programme cycle, then it can have several implications. If the evaluation is too early, there is little new data available and the evaluation will be less usable compared to the previous evaluation (partly the case in the 2010 midterm evaluation). If the evaluation is too late, then the evaluation findings are of little or no use (the case in the 2012 final evaluation). Also, an evaluation can be timed too close to other evaluations and thus leave no time for follow-up before the next evaluation starts. Mistiming of evaluations relative to decisions or other evaluations is an impeding factor relative to evaluation use.

LIFE evaluations were not perceived to be relevant directly for policy-making as the information contained in evaluations is too general and relates exclusively to programme implementation and less to the overall political rationales behind the programme. Also, evaluations are considered too technical and detailed by policy-makers. Evaluations are therefore not directly relevant for policy-makers, except when they serve as information feeding into the Commission's preparations for a new programme at the end of every programme cycle.

Finally, the data suggests that the standard evaluation implementation process of the Commission also secures a high **communication quality** of evaluations often through an iterative work process involving several parallel quality checks in the final phases of the evaluation implementation.

Summary of factors explaining evaluation use in LIFE evaluations. The results of the data analysis show that decision and policy settings are far more important than factors related to evaluation implementation. In relation to the decision and policy setting, the most prevalent impeding explanatory factors on evaluation use are competing information and decision characteristics. Particularly for policy-makers, there is a lot of information and very little time to digest it. For programme management, both are major obstacles to the use of evaluation findings and recommendations and substantial changes to the programme are difficult to make. Also the political climate is impeding evaluation in this case, as the LIFE programme is a small and not very contested, hence the Commission's need to prove accountable and the EP's urge to check the Commission are reduced relative to other programmes.

At the same time, explanatory factors related to decision and policy settings also enable evaluation use. The personal characteristics are enabling use, as evaluation training as well as tendering evaluations make the right people available for the job both internally and externally. Evaluation use is also enabled by high receptiveness and commitment to evaluation in the Commission. The advantages of the system's institutionalized practice is mainly relevant for the Commission and less so for other organizations in the system such as the EP and Council. However, all actors in the system share the need for information, which is a major enabling factor both in the Commission as well as outside the Commission.

Table 3 provides a summary of the explanatory factors' effect on use as they are reported by interviewees.

In relation to evaluation implementation, the main findings of the analysis are that the evaluation system reduces the most negative effects of several of the explanatory factors linked to evaluation implementation. Hence, timeliness, credibility, evaluation quality and communication quality are all factors that are supported by routinized and systematized institutions of

Table 3. Overview of the effect of the explanatory factors.

Explanatory factors	Enabling/impeding
Decision and policy settings	
Commitment and receptiveness	Enabling
Competing information	Impeding
Decision characteristics	Impeding
Information needs	Enabling
Personal characteristics	Enabling
Political climate	Impeding
Evaluation implementation	
Timeliness	Enabling
Credibility	Enabling
Evaluation quality	Enabling
Findings	Impeding
Relevance	Impeding
Communication quality	Enabling

practice in the Commission. Also, these advantages of the system's institutionalized practices are mainly relevant for the Commission and less so for other organizations in the EU system such as the EP and Council.

As just explained, the institutionalized practices of the evaluation system enable evaluation use. However, the rigidity of the system's practices also impedes evaluation use by controlling and thereby reducing the innovation and possible outcome of evaluation findings. The relevance of evaluations is limited to policy-makers in terms of content and to programme managers because of evaluation timing, which is not always optimal for evaluation use. The relevance of evaluations is limited to policy-makers in terms of content and to programme managers because of the timing. Evaluation findings and relevance are the impeding factors related to the evaluation implementation.

Users of LIFE evaluations

The main users of LIFE programme evaluations are the programme management in DG ENV. Evaluation use by other actors depends mainly on the relevance (timing relative to a decision) and the demand for information. Generally, evaluations are not used directly for policy making outside the Commission; i.e. the programme management will use evaluations to prepare the new LIFE programme and adjust its implementation. Thus LIFE programme evaluations are supporting documents in the programme proposals that decision-makers outside the Commission read and use every seventh year in the policy cycle. Therefore evaluation use for policy making is mainly indirect outside the Commission. Rather, evaluations outside the Commission are used for information purposes and sometimes for advocacy (strategic use) by for example NGOs.

Table 4 summarizes the findings from the above sections including the explanatory factors and use types that were mentioned most by interviewees. Explanatory factors not included in the table were those perceived to be less significant by interviewees. These include evaluation quality, timeliness, credibility, personal characteristics etc.

Table 4. Summary of most mentioned evaluation use types and explanatory factors per user group.

	Programme management	Policy-makers
Most relevant explanatory factors		
Enabling	- Commitment and receptiveness - Information need	- Information need
Impeding	- Decision characteristics - Findings - Relevance	- Competing information - Commitment and receptiveness - Relevance
Most common use types		
Process use	No process use	No process use
Findings use	- Instrumental - Legitimizing use - Strategic use	- Information use - Strategic use

Commitment and receptiveness to programme evaluations is an enabling factor mostly relevant to the programme management, as evaluations are too detailed and not directly relevant for policy-makers and other stakeholders. However, most stakeholders to the LIFE programme express a need for knowledge. Information need is thus an enabling factor for most evaluation users.

Regarding factors impeding evaluation use, interviewees mentioned mostly decision characteristics, relevance and findings and competing information. For programme management it is the decision characteristics of the programme cycle and the EU decision-making structure that impede evaluation use. Given that programme managers are the most knowledgeable with regard to the programme, institutionalized evaluation practices limit the relevance and innovation of findings, which could induce evaluation use. For policy-makers overload of information, lack of receptiveness to evaluation and the low relevance of evaluations are the main obstacles to evaluation use.

Conclusion

The aim of this article is to find answers to how, who and why – or why not – an evaluation system affects the utilization of evaluations. Throughout the analysis it has been demonstrated that in the case of the LIFE programme, the EU evaluation system is conducive to evaluation use, while at the same time also impeding use in several ways. The article finds that the evaluation system is conducive to instrumental-, strategic- and legitimizing types of evaluation use on the level of programme management. On the level of policy-makers, the evaluation system is conducive to strategic- and information types of uses. Regarding the users of evaluation findings, the instrumental use of LIFE evaluations is linked intrinsically with programme management in the two LIFE units in DG ENV while being far less used by policy-makers in the EP and the Council. While the system does enable uses, it also impedes process use of evaluations during the evaluation implementation.

With that in mind, the decision characteristics are key in order to understand the use of the LIFE evaluations and in particular the absence of process use. Decision characteristics are contextual relating to the legal- and organizational setup as well as the formal and informal practices of the evaluation system. The Commission decision characteristics largely determine

the effect that other explanatory factors have on evaluation use. In particular, the alignment of programme evaluation to the Commission's budget- and policy cycle has several consequences for evaluation use. These consequences of Commission decision characteristics are elaborated in the five main findings below.

1. Findings use over process use

The EU evaluation system is primarily designed to feed information into the EU decision-making procedure every seventh year before the beginning of a new policy cycle. Commission staff working with programme evaluation are managing the evaluations with the main objective of satisfying the evaluation obligation by securing the timeliness, quality and independence of the final output. To that end, the evaluation standards including the guidelines and terms of reference specify in great detail how the evaluation process should be executed in such a way that process use – as envisaged by for example Patton (1997) – is not enabled. As will be elaborated below, the limited process use is a consequence of choices made deliberately to secure and improve findings use for decision making, in particular 'decision points'. The loss of process use is thus a direct consequence of policy-making practices in the EU political system as a whole as well as the work practices decided for and by the Commission.

2. Findings use in 'decision-points'

The policy-cycle's decision-points enable findings use because the evaluation system is designed to deliver feedback into a particular decision-point at the time for programme renegotiating. In the case of the LIFE programme, this decision-point enables findings uses including instrumental-, legitimizing and strategic uses. These types of findings use are enabled because the decision point allows for potential programme change and because the Commission needs programme legitimacy and overall accountability when the focus of policy-makers in the EP and Council is on the programme. It is at these points in time that LIFE evaluations will be used indirectly (through for example IAs) to change the programme. The evaluation system is designed to deliver independent quality evaluations that are timely, well communicated, independent and credible in order to secure the legitimacy of the programme and the strategic position of the Commission in such negotiations. The institutionalization of evaluation practices including the highly standardized evaluation process, staff training and guidelines almost completely negates process uses as explained previously. However, they also enable a smoother execution of evaluation processes and ultimately deliver a more standardized evaluation product with minimal 'risks', as potential negative influences from personal characteristics of evaluators and staff, organizational deficiencies etc. are largely avoided.

3. Low evaluation relevance regarding significant programme change

Due to mistiming and competing information, LIFE evaluations are seldom relevant to potential users in decision making. The relevance of evaluations is affected negatively and directly by decision characteristics. Sometimes, evaluations are mistimed in relation to decision events such as the preparation of an IA or the proposal for a new programme. Further, the evaluation system generates competing information, including first and foremost many evaluations on the LIFE programme as well as other studies and audits. Altogether, this reduces the relevance of individual evaluations and impedes the instrumental use of them. Also, the legal structures

that govern the Commission's work (Financial Regulation etc.) reduce the possibilities of innovative and surprising recommendations because consultants' recommendations are steered first of all by the evaluation questions in the tender material that is typically in accordance with what the Commission can influence within the scope of its legal competences. For example would potential improvements to the LIFE Regulation be downplayed relative to recommendations that could readily be applied by the programme units in DG ENV.

4. The 'de-politicization' of programme evaluations

Following the three previous points, instrumental evaluation use is mainly limited to managers of the LIFE programme. Also, evaluations are rarely used in evaluation implementation and only rarely used directly for policy-making (for a similar finding see Laat, 2005). Due to mistiming of evaluations, other competing information and non-innovative findings, evaluations' relevance for decision making is very limited. These two findings imply a de facto de-politicization of programme evaluations in the EU evaluation system, where evaluation information conforms to the administrative context of programme management in the Commission instead of the political context of policy-makers.

5. The paradoxical evaluation system

The 'de-politicization' represents a paradox since it is the EP and the EU Member States that compelled the Commission to develop its evaluation practices in the first place. But this article shows that policy-makers rarely use the evaluations while at the same time the Commission also does not maximize the utility of evaluations. Also, the Commission allocates considerable resources to evaluations and numerous other internal and external assessments and audits. This paradox is probably explained best by turning to the complex nature of the inter-institutional context of the EU evaluation system and EU evaluations (see for example Dahler-Larsen, 2012; Højlund, 2014). To do that would be too much to cover at this point and more evidence is also needed to support such an analysis. But hopefully this article will spawn more interest in other aspects of the EU's evaluation system, such as the birth of systematic policy-evaluations (Stern, 2009). These often-contested and highly political evaluations have other attributes than the Commission's programme evaluations and therefore they are also likely to be used in different ways than we have seen is the case with the Commission's programme evaluations.

Funding

The PhD project is supported by the Danish Government's Industrial PhD programme, COWI A/S as well as the COWI-Foundation.

Note

1. At this point it should be noted that these well-known categories are all ex-post to evaluation implementation and therefore do not include the effects and evaluation that exist ex-ante as a consequence of evaluation anticipation (for example, redressing or window-dressing before the evaluator starts working). This use type was not included in the analysis because data on such uses is difficult to collect up to one decade after the evaluation was finalized. Moreover, the data collection allowed for accounts of ex-ante uses by asking several open questions, but no examples were given by the interviewees.

References

- Alkin M and Taut S (2003) Unbundling evaluation use. *Studies in Educational Evaluation* 29: 1–12.
- Alkin MC and Stecher B (1983) Evaluation in context: information use in elementary school decision making. *Studies in Educational Evaluation* 9: 23–32.
- Bazeley P (2013) *Qualitative Data Analysis with NVIVO*. London: SAGE.
- Bienias S and Iwona L (2009) *Evaluation Systems in the Visegrad Member States*. Warsaw: Polish Ministry of Regional Development.
- Commission E (2007) Evalueringsaktiviteter & -resultater på tværs af Europa-Kommissions tjenestegrene. In: Jacobsen S (ed.) *Dansk Evalueringsselskabs konference*. Kolding, Denmark.
- Cousins JB and Leithwood KA (1986) Current empirical research on evaluation utilization. *Review of Educational Research* 56: 331–64.
- Dahler-Larsen P (2012) *The Evaluation Society*. Stanford, CA: Stanford University Press.
- Easterby-Smith M, Crossan M and Nicolini D (2000) Organizational learning: debates past, present and future. *Journal of Management Studies* 37: 783–96.
- Finne H, Levine M and Nilssen T (1995) Trailing research: a model for useful program evaluation. *Evaluation* 1: 11–31.
- Floden RE and Weiner SS (1978) Rationality to ritual: the multiple roles of evaluation in governmental processes. *Policy Sciences* 9: 9–18.
- Furubo J-E (2006) Why evaluations sometimes can't be used – and why they shouldn't. In: Rist R and Stame N (eds) *From Studies to Streams*. New Brunswick, NJ: Transaction Publishers, 147–65.
- Furubo J-E, Rist RC and Sandahl R (2002) *International Atlas of Evaluation*. London: Transaction Publishers.
- Henry GT and Mark MM (2003) Beyond use: understanding evaluation's influence on attitudes and actions. *American Journal of Evaluation* 24: 293–314.
- Højlund S (2014) Evaluation use in the organisational context – changing focus to improve theory. *Evaluation* 20: 26–43.
- Imam I, LaGoy A and Williams B (2007) Introduction. In: Williams B and Imam I (eds) *Systems Concepts in Evaluation: An Expert Anthology*. Point Reyes, CA: EdgePress of Inverness.
- Johnson K, Greenesid LO, Toal SA, et al. (2009) Research on evaluation use: a review of the empirical literature from 1986 to 2005. *American Journal of Evaluation* 30: 377–410.
- Kirkhart KE (2000) Reconceptualizing evaluation use: an integrated theory of influence. *New Directions for Evaluation* 88: 5–23.
- Kohlbacher F (2006) The use of qualitative content analysis in case study research. *Forum: Qualitative Social Research* 7(1). URL: <http://www.qualitative-research.net/index.php/fqs/article/view/75/154>
- Laat Bd (2005) Study on the use of evaluation results in the commission. Paris: Technopolis, 31 May.
- Langley A (1998) In search of rationality – the purpose behind the use of formal analysis in organizations. In: Maanen JV (ed.) *Qualitative Studies of Organizations*. Thousand Oaks, CA: SAGE, 51–90.
- Lazarsfeld PF, Sewell WH and Wilensky HL (1967) *The Uses of Sociology*. New York: Basic Books.
- Leeuw FL and Furubo J-E (2008) Evaluation systems : what are they and why study them? *Evaluation* 14: 157–69.
- Leviton LC (2003) Evaluation use: advances, challenges and applications. *American Journal of Evaluation* 24: 525–35.
- Leviton LC and Hughes EFX (1981) Research on the utilization of evaluations: a review and synthesis. *Evaluation Review* 5: 525–48.
- Mayring P (2000) Qualitative content analysis. *Forum: Qualitative Social Research* 1(2). URL: <http://www.qualitative-research.net/index.php/fqs/article/view/1089/2385>
- Mayring P (2004) Qualitative content analysis. In: Flick U, Von Kardorff E and Steinke I (eds) *A Companion to Qualitative Research*. London: SAGE, 266–9.

- Patton MQ (1997) *Utilization-Focused Evaluation: The New Century Text*, Thousand Oaks, CA: SAGE.
- Pollitt C, Girre X, Lonsdale J, et al. (1999) *Performance Audit and Public Management in Five Countries*. Oxford: Oxford University Press.
- Pröpper IMAM (1987) Beleidsevaluatie als argumentatie. *Beleidswetenschap* 2: 113–36.
- Rich RF (1977) Uses of social science information by federal bureaucrats: knowledge for action versus knowledge for understanding. In: Weiss CH (ed.) *using Social Research in Public Policy Making*. Lexington, MA: Lexington Books.
- Rist R and Stame N (2006) *From Studies to Streams: Managing Evaluative Systems*. London: Transaction Publishers.
- Ritchie J, Lewis J and Elam G (2003) Designing and selecting samples. In: Ritchie J and Lewis J (eds) *Qualitative Research Practice - A Guide for Social Science Students and Researchers*. London: SAGE.
- Schreier M (2012) *Qualitative Content Analysis in Practice*. Thousand Oaks, CA: SAGE.
- Stern E (2009) Evaluation policy in the european union and its institutions. *New Directions for Evaluation* 123: 67–85.
- Toulemonde J (2000) Evaluation culture(s) in Europe: differences and convergence between national practices. *Vierteljahrshefte zur Wirtschaftsforschung* 69: 350–7.
- Van der Knaap P (1995) Policy evaluation and learning: feedback, enlightenment or argumentation? *Evaluation* 1: 189–216.
- Weiss CH (1992) *Organizations for Policy Analysis – Helping Government Think*. Newbury Park, CA: SAGE.
- Weiss CH (1998) Have we learned anything new about the use of evaluation. *American Journal of Evaluation* 19: 21–33.
- Weiss CH and Bucuvalas M (1980) *Social Science Research and Decision Making*. New York: Columbia University Press.
- Williams B and Imam I (2008) Evaluation systems: what are they and why study them? *Evaluation* 14: 157–69.
- Williams K, Laat Bd and Stern E (2002) *The Use of Evaluation in the Commission Services*. Paris: Technopolis, 8 October.
- Zito A and Schout A (2009) Learning theory reconsidered: EU integration theories and learning. *Journal of European Public Policy* 16: 1103–23.

Steven Højlund is a PhD candidate at the Copenhagen Business School and is currently affiliated with Stanford University as a visiting scholar. Steven has more than five years of professional experience with evaluation for the European Commission.