

Evaluation capacity: enabler or exponent of evaluation culture?

De Peuter Bart & Pattyn Valérie, Public Management Institute K.U.Leuven (Belgium)

1. Introduction

When reviewing the evaluation literature, one can hardly contest the existence of a growing consensus on the necessity of a mature *evaluation capacity* in the light of good governance. In addition, there seems to be agreement on the beneficial value of having a mature *evaluation culture*. However, when focusing more in detail on the essential characteristics of both 'phenomena', it is difficult keeping the overview over the multitude of opinions circulating. The analysis is even more complicated if one tries to shed light on the *interrelationship* of both concepts. Surprisingly, both 'discourses' are indeed seldom explicitly integrated. Varying explanations may hold for this, ranging from the conviction that both concepts mean the same, to the assumption that both variables develop independently from each other.

In this paper the authors aim to bring some clarity about the interrelationship of the two concepts. More concretely, the following questions will be addressed: Does an evaluation capacity constitute a precondition for the development of an evaluation culture? Or do we instead need to see evaluation culture rather as a prerequisite and capacity building as its exponent? Or do both phenomena presume other underlying determinants?

The paper in a first part discusses the methodological approach which guided the research (section 2). Based on this methodology, the authors present in a following section their analysis of the existing discourse on the two constructs (section 3). The trends and commonalities brought to the surface are subsequently linked to the different stages of the generic evaluation process (section 4). The coupling to the evaluation process finally forms the basis to bring more insight in the position of evaluation capacity and culture vis-à-vis each other (section 5).

By exploring the theoretical meaning of both concepts, and their translation into indicators, the paper also intends to contribute to a consensus-formation on the polyphony of conceptualizations circulating. Without denying the 'social constructivism' of 'meaning', a common understanding of evaluation capacity and evaluation culture would be beneficial to both theorists and practitioners. It would facilitate a common language among evaluation stakeholders and provide a clear way to communicate (Taut, 2007; Compton et al., 2007).

2. Methodological approach

In order to reveal the mutual relationship between evaluation capacity and evaluation culture, it is *a priori* considered necessary having a clear understanding of both concepts separately. A critical analysis of the written discourse on both topics formed hereto the basis. By deliberately using an inductive approach, it was intended to respect as much as possible the divergent views and conceptualizations which could be detected in the different sorts of documents studied, without imposing a specific bias beforehand.

For the purposes of the research, it was decided to 'screen' a rich variety of different sources (n=16), relevant for the domain of 'policy evaluation'. As such, the research focused on (1) articles of the most cited journals in the field (Evaluation, New Directions for Evaluation, American Journal of Evaluation); (2) books/chapters in books of renowned 'evaluation scholars'; and (3) some of the most well-known evaluation checklists. The study initially also included an analysis of evaluation glossaries and standards of evaluation networks. Yet, it appeared that only a minor number of glossaries comprised the terms evaluation capacity or culture. It was further noticed that those glossaries mentioning (one of) the terms provided only very general descriptions. As the latter were not very useful for further analysis, glossaries were excluded from the final investigation. Evaluation standards were neither considered useful, as they never explicitly comprised definitions or descriptions of evaluation capacity and/or evaluation culture.

Obviously restricted by language barriers, the analysis primarily concerned English language sources, though coming from a rich international group of scholars. Every document was in detail analyzed with the aim to distil relevant indicators with which the phenomena are referred to. The analysis proceeded until no new indicator was encountered. The objective is to come to a number of categories of indicators with general relevance. Indicators which only concerned e.g. developing countries were deleted from the analysis. The systematic comparison of the different indicators per author constituted the basis to derive the consensus on the essence of every construct.

To bring some clarity in the mutual relationship between evaluation culture and evaluation capacity, it was further investigated whether the indicators revealed could be linked with certain stages of the evaluation process. Do capacity and culture refer to the same stages of the evaluation process, or do they each rather relate with some specific stage(s) in the evaluation process?

3. Views on evaluation capacity and culture: convergence or Babel-like confusion?

3.1 Terminological diversity

When systematically screening sources on relevant indicators for evaluation capacity and culture, the amalgam of existing notions and nuances related to the concepts under study is striking. We could roughly notice a difference between the following notions, though none of them is applied in a systematic and consistent way across the different sources.

- Evaluation capacity versus evaluation capability (see e.g. Williams, 2001)
- Evaluation capacity building versus evaluation capacity development (see e.g. Horton, 1999)
- Professional evaluation versus evaluation capacity building (see e.g. Baizerman et al., 2005)
- Evaluation capacity versus evaluation practise (see e.g. Boyle et al., 1999)

In order not to be biased by one particular author, and given the inconsistent use of these nuances over different sources, the research was deliberately not restricted to one specific category of studies. Instead, all relevant sources were listed, notwithstanding their particular approach.

Similarly, given the incoherent use of evaluation culture and evaluation capacity, it was considered more appropriate not clustering the analysis from the very beginning into two separate 'lists' of indicators. On the contrary, one single (giant) table was produced, including

the overview of indicators for both constructs. The indicators derived from the document analysis initially concerned all levels of detail. This complicated the research to some extent, but also guaranteed not losing any relevant data by excluding particular criteria. An overview of all 251 indicators identified is listed in Annex 1.

3.2 Different sources, different indicators

The rich heterogeneity over the number of sources is hence intentionally taken into consideration. Every particular source approaches evaluation capacity and/or evaluation culture from a particular perspective.

For the purpose to compare the evaluation maturity of 21 countries and international organisations, *Furubo and Sandahl* developed a 9-items benchmarking tool. The instrument intends to describe two aspects of evaluation maturity: the degree to which an evaluative praxis is established as an integral part of the politico-administrative system; and the actual spread and pluralism of the evaluative culture and its openness to new ideas and impulses (Furubo et al., 2002).

Varone et al. developed an index of institutionalization of evaluation, suitable to apply at the national level. The index basically revolves around two dimensions: existence of formal organisations and existence of an epistemic community. Their instrument correlates in several respects with the Furubo & Sandahl's 'evaluation maturity index'. No country seems to demonstrate an intensive institutionalization process without an accompanying increase in maturity. The inverse is nevertheless not always valid. There are countries who developed a mature evaluation culture, without displaying a high degree of institutionalization (Varone et al., 2006)

Based on the World Bank's experience in supporting governments developing monitoring and evaluation capacity, *Mackay* (Mackay, 2002) prioritizes demand-fostering as pivotal in evaluation capacity building. A purely-supply oriented focus would be ineffective to guarantee the necessary long-term commitment to evaluation capacity. Emphasizing a context-adapted capacity building strategy, the Operations Development Department has developed a checklist for a good-quality level ECB. Evaluation capacity building is further represented as a 'results chain', with performance indicators indicated at each subsequent stage (activities, output and outcome).

In another renowned publication, *Mackay* lists a number of potential barriers to building evaluation systems. Although the publication mainly concerns developing countries, the list of hindrances is a valuable source for other countries as well. Central in his vision is the plea for a comprehensive approach in building evaluation capacity, in synergy with performance measurement and management (Mackay, 1999).

Guerrero, also affiliated to the World Bank, sketches the variables perceived essential for institutionalizing evaluation, based on experience in three developing countries (Colombia, China, and Indonesia). The supply and demand framework constitutes the main 'road map' (Guerrero, 1999). Like other World Bank evaluators, Guerrero emphasizes the need for a country-adapted approach, depending on each country's position in the demand/supply matrix.

The *European Commission* in their widely used 'Evalsed Guidelines' considers evaluation capacity a priori as a multifaceted notion. Consequently, they emphasize the necessary development of evaluation capacity at different levels, reinforcing each other (individual, organisational, inter-organisational and societal level). For every level, they identify a number of key indicators, which should be present in an ideal 'evaluation capacity'

situation. Governments engaging in evaluation capacity development in a strategic way should, according to the European Commission's point of view, deal with four key areas in particular: evaluation architecture, evaluation demand, evaluation supply, and institutionalizing evaluations (European Commission, 2008).

Based on a case study of a government's five-year capacity building development path, also *McDonald et al.* have constructed a list of practical themes to be considered as recommendations for building evaluation capability. Evaluation capability is not only understood as self-evaluation by programme managers and staff; but also refers to the effective use of evaluations and monitoring (McDonald et al., 2003). A fishing metaphor is used to illustrate this argument.

Lahey, experienced in evaluation in the Canadian federal public service, also emphasizes to distinguish between developing the capacity to 'do' evaluations and developing the capacity to 'use' evaluations. Based on his expertise, he has developed an 'evaluation infrastructure checklist' with both cultural and operational key factors crucial for effective evaluations. (Lahey, 2004).

The *United States General Accounting Office* investigated five agencies which all demonstrated evaluation capacity. Based on their empirical findings, they derived four key elements of evaluation capacity. Noteworthy is that evaluation culture has explicitly been identified as a sub element of evaluation capacity (US GAO, 2003).

As stipulated above, *Boyle, Lemaire and Rist* conceive evaluation capacity as clearly different from evaluation practice. While the latter refers to the actual 'doing of evaluations', the former is to be seen as the necessary precondition, or the hardware without which practice is unthinkable. Similar to the European Commission (2008), Boyle et al. have enumerated seven key issues that governments must consider wishing to enhance the demand and supply of evaluations (Boyle et al., 1994).

Analogously, *Baizerman, Compton, and Stockdill* call to make a clear distinction between evaluation capacity building on the one hand and the practise of professional program evaluation on the other (Baizerman et al., 2005).

As part of the Evaluation Checklist Project, the University of Michigan's Evaluation Center has developed two reviewed checklists with immediate relevance to our research question. A first checklist is the one by *Volkov and King* on "Building Organizational Evaluation Capacity" (Volkov & King, 2007), which provides guidance for incorporating evaluation routinely into the life of an organization. This checklist complements the "Institutionalizing evaluation checklist" of *Stufflebeam* (Stufflebeam, 2002) which is – according to the author- also of utmost use for fields as Evaluation Capacity Building and Evaluation Capacity Development.

King (2007) addresses the issue of intentional process use of evaluation for evaluation capacity building. The author uses the CIPP model of *Stufflebeam* which emphasizes four aspects of program evaluation: context, input, process en product, and applies it to evaluation capacity building as evaluandum. In this way the CIPP model suggests variables to consider in using process use to build evaluation capacity.

Bustelo (2006) discusses the potential role of evaluation standards or guidelines in developing an evaluation culture in Spain. In her analysis, the author states indicators of a professional consolidation of a discipline and a number of reasons for the absence of an evaluation culture in Spain. The latter can indeed be regarded as critical success factors for evaluation culture from her point of view.

Dabelstein (2007) reports some lessons learned with evaluation capacity development as experienced by the DAC Working Party on Aid Evaluation. In particular, he refers to the

findings of a survey of 1996 on this topic, which in his conviction are still valid. The lessons, Dabelstein argues, generally apply to both donor and recipient agencies in the context of (evaluation of) development aid policy.

3.3 Toward unity behind diversity

Although concise, the overview of sources clearly illustrates the variety of different interpretations in the discourse about evaluation capacity and evaluation culture. Having listed the various indicators identified in literature, an attempt was made to bring some clarity in the many opinions and levels of analysis. It was strived to come to a common denominator for every 'cluster of a particular type of indicators'. This exercise yielded the following 17 categories of indicators (table 1). The table also clarifies the specific rationale behind every 'label'. Table 1 also contains a code for each indicator category. Annex 1 includes for each of the 251 indicators the code of the corresponding indicator category it has been assigned to.

Category	Code	Quid?
1. Awareness of intrinsic value of evaluation	AWE	- Policy makers have a positive attitude about evaluation
2. Context-dependency	CTX	- Evaluators and evaluation capacity builders take specific environmental circumstances into account
3. Coupling with policy and management decisions	CPM	- Evaluations are integrated into policy making - Evaluations are linked with management instruments - Evaluations are timely generated to make it possible incorporating them into the decision-making process
4. Coverage of the policy spectrum	CPS	- Evaluations take place at each stage of the programming cycle - Evaluations are executed in a wide number of policy domains - A regular flow of evaluations is undertaken
5. Data collection mechanisms	DCM	- The availability of data sources is guaranteed through well-functioning data collection systems
6. Demand	DEM	- A genuine evaluation demand exists (without referring to a specific 'driver' for demand)
7. Diffusion and feedback mechanisms	DIF	- Procedures to accumulate evaluation findings - A smooth dissemination of reports - The existence of well-functioning communication channels to spread the evaluations
8. Financial resources	FR	- The necessary financial sources are available to conduct evaluations
9. HR capacity (internal / external)	HR	- Sufficient human resources are available to perform an evaluation (at the organisation level; as well as in the broader evaluation market)
10. Legal embedding	LE	- Existence of a legal framework for evaluations
11. Networking	NET	- Networking with evaluation stakeholders in the framework of evaluation associations, etc. The label also clusters references to 'a 'Community of Practice'.
12. Organisational anchoring of evaluation function	ORG	- The evaluation function is structurally embedded in the organisation (centralized/decentralized) - The evaluation function is structurally embedded in a country's governing system (within the legal and/or executive power)
13. Ownership	OWN	- Relevant stakeholders are engaged in the evaluation process - Socialization of evaluation in the organisation
14. Political commitment	POL	- There exists genuine political guidance to help developing the evaluation function - The leadership is supportive for developing the evaluation function
15. Quality instruments	QLI	- Mechanisms to secure the quality of the evaluation process exist
16. Skills to perform evaluation	SK	- The necessary technical skills to successfully perform an evaluation are available or are being developed (through trainings, etc.)
17. Use	USE	- The effective use of evaluations

Table 1: Categories of indicators for evaluation capacity and culture

The following table gives the overview of indicators per author, irrespective of their citation frequency.

Categories of indicators for evaluation capacity and culture		Source (by number) ¹															total	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15		16
ORG	Organisational anchoring of evaluation function	x		x		x	x		x	x	x	x	x	x	x	x	x	13
SK	Skills to perform evaluation	x	x	x	x	x	x		x	x	x		x	x			x	13
AVE	Awareness of intrinsic value of evaluation		x	x	x			x		x	x		x	x	x	x	x	10
DCM	Data collection mechanisms	x	x	x	x			x			x		x	x	x			9
NET	Networking						x	x		x	x		x	x	x	x	x	9
QLI	Quality instruments			x	x	x		x		x	x		x	x	x			9
CPM	Coupling with policy and management decisions	x	x	x	x	x				x					x	x	x	9
DEM	Demand		x	x	x	x		x							x	x	x	8
DIF	Diffusion and feedback mechanisms		x			x			x	x			x	x	x	x	x	8
FR	Financial resources	x		x		x				x	x				x	x	x	8
POL	Political commitment	x		x	x			x					x	x	x	x	x	8
USE	Use	x			x		x		x	x	x					x	x	8
CTX	Context-dependency			x	x	x	x		x		x					x		7
HR	HR volume		x	x	x	x					x			x			x	7
LE	Legal embedding	x		x	x			x	x								x	6
CPS	Coverage of the policy spectrum	x			x	x	x										x	5
OWN	Ownership		x			x									x	x		4

Table 2: Indicators for evaluation capacity and culture grouped per category for each source

¹ The numbers respectively refer to the following sources: (1) Boyle, Lemaire and Rist, 1999; (2) Mackay, 1999; (3) Guerrero, 1999; (4) Mackay, 2002; (5) Stufflebeam, 2002; (6) Furubo and Sandahl, 2002; (7) United States General Accounting Office, 2003; (8) Dabelstein, 2003; (9) McDonald, Rogers and Kefford, 2003; (10) Lahey, 2004; (11) Baizerman, Compton and Stockdill, 2005; (12) Varone, Jacob and De Winter, 2005; (13) Bustelo, 2006; (14) Volkov & King, 2007; (15) King, 2007; (16) European Commission, 2008.

As can be derived from table 2, none of the indicators is mentioned by all authors. This might be partially clarified by the sometimes particular scope of the documents studied (see above). Nevertheless, despite their variety and the inherent limitations of the exercise, some relevant trends can be identified. 13 out of the 16 different sources seem to agree on the necessity of having the required 'skills to perform an evaluation' and the structured 'anchoring of the evaluation function within the organisation' to speak of evaluation capacity and/or evaluation culture. More than half of the authors also mention 'awareness of the intrinsic value of evaluation' (n=10); 'data collection mechanism' (n=9), 'networking' (n=9); 'quality instruments' (n=9) and the coupling with policy and management decisions as essential characteristics (n=9). It is further remarkable that only a quarter of the documents stipulates the requirement of having 'ownership' of an evaluation in discussing evaluation capacity and/or evaluation culture (n=4).

The picture, however, needs to be nuanced to some degree, if we add the following table, mentioning the exact number of citations for every category of indicators, per author.

While an equal number of authors for instance places 'organisational anchoring of the evaluation function' and 'skills to perform evaluations' central in its discourse (table 2); the former de facto refers to 37 out of the total number of 251 initial indicators. The latter corresponds instead to a 'fewer' amount of 33 indicators (table 3). Similar observations can be made for other indicators. The comparison of both tables offers hence more solidity to claim that most discourse concerns the issue of 'anchoring the evaluation function'. It can in addition be said that the frequency of citations under the 'veil' of every category to some extent also forms an indicator itself for the richness of sub-indicators circulating on a particular theme.

Categories of indicators for evaluation capacity and culture		Sources (by number, see footnote 1)																Total indicators per category
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
ORG	Organisational anchoring of evaluation function	1	0	2	0	1	3	0	1	1	5	4	3	1	8	6	1	37
SK	Skills to perform evaluation	3	1	2	2	4	1	2	0	2	5	1	0	3	5	0	2	33
AVE	Awareness of intrinsic value of evaluation	0	2	1	1	0	0	1	0	0	5	1	0	1	5	1	3	21
NET	Networking	0	0	0	0	0	1	1	0	1	3	0	1	3	4	1	5	20
QLI	Quality instruments	0	0	1	1	5	0	1	0	1	4	0	2	2	2	0	0	19
CPM	Coupling with policy and management decisions	1	1	2	3	1	0	0	0	0	1	0	0	0	2	1	2	14
POL	Political commitment	2	1	1	2	0	0	0	1	0	0	0	0	1	1	3	1	13
USE	Use	1	0	0	4	0	0	1	0	2	1	1	0	0	0	1	2	13
DIF	Diffusion and feedback mechanisms	0	1	0	0	1	0	0	0	1	1	0	0	1	4	1	2	12
DCM	Data collection mechanisms	1	1	1	3	0	0	1	0	0	0	1	0	1	1	1	0	11
FR	Financial resources	1	0	1	0	1	0	0	0	0	2	1	0	0	3	1	1	11
HR	HR volume	0	1	1	2	1	0	0	0	0	0	2	0	0	1	0	3	11
CPS	Coverage of the policy spectrum	1	0	0	1	1	3	0	0	0	0	0	0	0	0	0	3	9
CTX	Context-dependency	0	0	1	2	1	1	0	0	1	0	1	0	0	0	1	0	8
DEM	Demand	0	1	1	1	1	0	0	1	0	0	0	0	0	1	1	1	8
LE	Legal embedding	1	0	1	1	0	0	0	1	1	0	0	0	0	0	0	1	6
OWN	Ownership	0	1	0	0	1	0	0	0	0	0	0	0	0	2	1	0	5
Total indicators per source:		12	9	15	23	17	9	7	4	10	27	12	6	13	37	18	27	251

Table 3: Breakdown of indicators for evaluation culture and capacity by category and source

4. Clustering of indicators and determinants

In the former section we have looked for a way to find some unity behind the diversity concerning the indicators which are used in literature to describe and define the concepts of evaluation capacity, evaluation culture and related terms such as evaluation practice and evaluation maturity. This first attempt to regroup the huge variety of aspects and components mentioned by different types of sources resulted in 17 categories of indicators. Although not completely neutral, this classification provides a useful working instrument to handle the differences and commonalities across definitions and accents in descriptions.

In this section we take our analysis two steps further toward the answer on the central question on the interrelation between evaluation culture and evaluation capacity. In a first step, the indicator categories will be classified into two refined clusters of evaluation culture and evaluation capacity (section 4.1). In a second stage, we introduce a process perspective in order to examine the interrelatedness between evaluation culture and capacity (section 4.2).

4.1 Delineation of the concepts evaluation culture and capacity

If we want to discern between evaluation culture related indicators and evaluation capacity related indicators in order to take the analysis a step further, we need a clear and transparent definition of both concepts. This exercise is not evident since the main conclusion from the literature research was exactly that both terms are not used and defined in a consistent way and, consequently, show considerable overlap.

We refer to the general concept of 'culture' to define *evaluation culture*. Of course, there exists much debate on the best delineation of the term 'culture'. Fleshing out the different standpoints in this respect would lead us too far within the limits of this paper. We take instead the definition of Davis (1984) which contains components about which there seems to exist a broad consensus that they are relevant to describe the core meaning of culture. This author states that "culture is the pattern of shared beliefs and values that give members of an institution meaning, and provide them with the rules for behaviour in their organisation" (quoted in Martin, 2002). We consequently consider evaluation culture as the pattern of shared beliefs and values of policy makers and evaluators which provide them with rules for behaviour that lead towards a practice of evaluation.

As far as *evaluation capacity* is concerned, we associate this concept in the reminder of our analysis with more operational aspects and components which are deemed necessary for conducting an evaluation. In that sense, evaluation capacity is strongly linked to the evaluation practice itself. They both relate to personnel related issues as well as technical facilities and instruments in support of evaluation.

Applying this rationale to delineate the concepts of evaluation culture and capacity on the 17 categories results in the following table. Most of the indicator categories can easily be assigned to either the culture cluster or the capacity cluster.

The cultural cluster contains the awareness for evaluation, the context-dependency of evaluation, the (existence of) a genuine demand for evaluation, the (degree of) networking activities, the ownership of evaluations, political commitment to evaluation, and the ways and extent of effective utilisation of evaluation results.

The capacity cluster includes mechanisms for data collection and for diffusion and feedback, financial and human resources, as well as skills to perform evaluation and the presence of instruments to guarantee the quality of evaluation.

There are however 4 categories which are more difficult to classify either in the culture cluster or in the capacity cluster. In fact, all four can be associated with both clusters.

The coupling of evaluation with policy and management (category 3) decisions can be seen as a function of the awareness or political commitment in support policy evaluation. On the other hand one could translate this category into more technical procedures which ensure that evaluation is carried out timely and results feed into the policy-making process. If we verify the initial long list of single indicators, the latter translation is mostly stated. Therefore, we have assigned this category to the capacity cluster, but put it between brackets to indicate the room for interpretation.

The same remark holds for the legal embedding of evaluation (category 10), albeit that we would rather assign this category to the culture cluster considering the legal embedding as a consequence of cultural drivers.

The coverage of the policy spectrum (category 4) refers to the volume of demand across policy sectors and policy processes. Of course, this degree of coverage will be influenced by the presence of capacity within different policy sectors, but also by the commitment and awareness of the added value of evaluation by the respective policy-makers within different policy areas. The descriptions in the long list of indicators do not bring much more clarity in this regard. Because evaluations can be carried out in circumstances with suboptimal capacity, we consider the degree of coverage rather as a cultural fact.

The organisational anchoring of the evaluation function (category 12) can be considered as a deliberate choice motivated by the conviction of the added value of evaluation. On the other hand, the anchoring exercise itself breaks down in a number of capacity related aspects such as the structural positioning of the evaluation function, the provision of resources and the installation of mechanisms for data collection and feedback. Therefore, we consider this category primarily as belonging to the capacity cluster.

Category	Culture cluster	Capacity cluster
1. Awareness of intrinsic value of evaluation	X	
2. Context-dependency	X	
3. Coupling with policy and management decisions		(X)
4. Coverage of the policy spectrum	(X)	
5. Data collection mechanisms		X
6. Demand	X	
7. Diffusion and feedback mechanisms		X
8. Financial resources		X
9. HR capacity (internal / external)		X
10. Legal embedding	(X)	
11. Networking	X	
12. Organisational anchoring of evaluation function		(X)
13. Ownership	X	
14. Political commitment	X	
15. Quality instruments		X
16. Skills to perform evaluation		X
17. Use	X	

Table 4: Indicator categories assigned to a culture and capacity cluster

4.2 The evaluation process as a dynamic framework

Let us reconsider the central question for this paper: Is evaluation capacity an enabler or an exponent of evaluation culture? In other words, is the existence of an evaluation capacity needed to develop an evaluation culture, or do we have to think otherwise and see evaluation culture rather as a prerequisite for building evaluation capacity? This question indeed assumes that there may be playing a trigger effect between both elements. Instead of going so far in our assumptions to expect a straightforward causal relationship, we rather suspect that one aspect can be regarded as contributing to the development of the other.

In the previous sections we reduced the variety of indicators mentioned in literature for both concepts and regrouped them into two clusters: evaluation culture and evaluation capacity. In order to position evaluation culture and capacity vis-à-vis each other, we need a framework to position the respective indicator categories of each cluster. Ideally this framework introduces a dynamic perspective to analyse the relationships between both clusters.

The generic scheme of an evaluation process excellently serves the purposes of our exercise. Scriven (1980) discerns 4 standard steps which have to be taken in each evaluation: (a) structuring, i.e. clarifying the evaluative questions and criteria, (b) observing (c) analyzing and (d) judging. When we look at an evaluation process from a somewhat broader view, we can add a preceding step where a decision is made to evaluate a measure, project or programme. Further, evaluations are not an end in itself but are conducted in order to be used in the decision-making process. Therefore, we can add an extra step following the stages Scriven has stated.

Thus, we come to a process with six steps:

- (1) deciding to evaluate
- (2) structuring / planning the evaluation
- (3) observing
- (4) analyzing
- (5) judging
- (6) utilization

These steps can be regrouped into 3 phases, i.e. (1) deciding to evaluate, (2) conducting the evaluation and the (3) utilization / influence of an evaluation. We will use this scheme to assign the 17 indicator categories from the culture and capacity cluster.

4.3 Attribution of the indicators to the standards steps in the evaluation process

The next table shows the result of connecting the indicator categories from the culture and capacity cluster to the standard steps of the evaluation process. For each indicator the respective cluster is indicated between brackets.

To the first stage of the evaluation process – the decision to evaluate- five indicator categories can be attributed. In order to decide to evaluate a policy measure, project or programme, we can logically presume that awareness of the intrinsic value of evaluation enhances the chances that policy makers will opt for evaluation to underpin their decisions

and commission specific evaluations. Political commitment to support evaluation is another variable which contributes to decisions initiating evaluations. Also the mere volume of demand is obviously connected to the stage of deciding to evaluate. The same holds for the coverage of the policy spectrum; the extent of evaluations commissioned in each policy domain. We also categorised the indicator of context-dependency under the stage of deciding to evaluate. The extent to which it is decided to evaluate will vary according to country-, organisation- or policy sector-specific factors.

To the second stage of the evaluation process – conducting evaluations- six indicator categories can be assigned. In order to carry out a specific evaluation, not only financial and human resources are needed. Evaluators also need to have the necessary skills to evaluate policy. Evaluation also relies on input from different information sources. Besides data collection mechanisms, also arrangements to diffuse information during the evaluation process (e.g. between commissioner and evaluator) need to be in place. Quality instruments are developed with regard to the process of evaluating itself and its different steps of structuring, data collection, analysis and judgement.

To the last stage of utilization and influence we have attributed two indicator categories. Ownership can be regarded as an important variable and condition to enhance the chance that evaluation results are used or will have an influence on decisions of policy-makers. The use itself was also identified as a common indicator after the literature screening.

Four out of the 17 categories of indicators are difficult to assign to one of the steps of the evaluation process. The legal embedding of evaluation procedures and requirement relates perhaps most to the stage of deciding to evaluate. However, regulations on evaluation and an evaluation framework may also contain prescriptive clauses on the process of evaluation or with regard to the utilisation of evaluation results. It is for example not unimaginable that evaluation conclusions must be included in a note underpinning a specific kind of policy decision, as part of good governance practice. The organisational anchoring has often been mentioned in the light of capacity building. Intra-organisational capacity can relate both to supply and demand of evaluation and additionally also to the approach on utilisation of evaluation results. The coupling of evaluation with policy and management decisions can be addressed at the stage of deciding an evaluation but also at the end of the evaluation process itself when results feed into the decision-making process, in the stage of utilisation. The last category, networking, can be linked to all different steps in the elaborated evaluation process; the exchange of ideas, expertise and experience can contribute to better decisions and planning for evaluation, enhance the quality of performing an evaluation as well as support the utilisation of evaluation findings.

These four categories of indicators do, in our view, not exclusively relate to one particular stage in the evaluation process. Therefore, we have classified them separate line at the bottom of the table. The double arrows indicate that they are linked with the evaluation process as a whole.

Deciding to evaluate	Conducting the evaluation				Utilization / influence
	Structuring	Observing	Analyzing	Judging	
<ul style="list-style-type: none"> • Awareness of intrinsic value of evaluation (CUL) • Political commitment (CUL) • Demand (CUL) • Coverage of the policy spectrum (CUL) • Context-dependency (CUL) 	<ul style="list-style-type: none"> • Financial resources (CAP) • HR capacity (internal / external) (CAP) • Skills to perform evaluation (CAP) • Data collection mechanisms (CAP) • Diffusion and feedback mechanisms (CAP) • Quality-instruments (CAP) 				<ul style="list-style-type: none"> • Ownership (CUL) • Use (CUL)
<p>Legal embedding (CUL)</p> <p>← Organisational anchoring (CAP) →</p> <p>Coupling with policy and management decisions (CAP)</p> <p>Networking (CUL)</p>					

Table 5: Indicator categories assigned to the standard steps of a generic evaluation process

5. Evaluation capacity: enabler of exponent of evaluation culture?

In this final section, we zoom out again and examine the distribution of indicator categories from both the evaluation culture cluster and the evaluation capacity cluster throughout the generic evaluation process. We come now to the core of our analysis in order to answer the research question: is evaluation capacity an enabler or exponent of evaluation culture?

5.1 Evaluation culture-driven demand and use of evaluation

The distribution of indicator categories shows that all five categories linked to the step of deciding to evaluate stem from the evaluation culture cluster. The extent of demand itself and its coverage of the policy spectrum are foremost culturally determined. Awareness of the value of evaluation and political commitment are also cultural variables which can be seen as drivers for demand. The country- or organisation-specific context has a cultural connotation too. Also this element can be introduced to explain extent of an existent demand of evaluation.

Analogously, the two indicator categories classified under the utilisation and influence step of the evaluation process are cultural variables: the extent to which evaluation findings are really used in policy-making is also dependent on the general politico-administrative culture. Ownership of evaluation results has to be built during the evaluation process and will in its turn contribute to the use of evaluation conclusions and recommendations.

We conclude that both demand and use of evaluation are predominantly driven by factors relating to evaluation culture.

5.2 Capacity-driven performance of evaluation

To the step of conducting evaluation, i.e. the technical core part of the evaluation process, six indicator categories are assigned which all stem from the evaluation capacity cluster. Besides resources and professional skills, mechanisms and instruments for data collection and information diffusion are needed, as well as for quality control.

We conclude that the process of performing an evaluation relies predominantly on conditions which relate to (technical) evaluation capacity.

5.3 Evaluation capacity vis-à-vis evaluation culture

The two former synthetic conclusions bring us to the answer on the interrelatedness of evaluation capacity and culture. The 17 indicator categories refer to specific indicators which in literature are mentioned with regard to the concepts of evaluation culture, capacity, practice, maturity and institutionalisation. We have regrouped these categories into an evaluation culture cluster and a technical evaluation capacity cluster. Thereafter we have assigned the categories to the different steps in a generic evaluation process, bringing in a process perspective. From the distribution of the indicator categories we can formulate the following conclusions on the interrelatedness of evaluation capacity and culture.

There are some cultural factors which can be regarded as prerequisites for capacity components to be built up in order to perform evaluations. Political commitment and awareness of the added value of evaluation will trigger demand for evaluation. Demand on its part needs supply and thus technical capacity to carry out an evaluation. Hence, the need for capacity building will increase when demand increases.

In turn evaluation capacity will have an influence on the use of evaluation. The amount of resources, the presence of quality control instruments and mechanisms to feed the evaluation findings into the policy process will contribute the extent of evaluation use. Ownership is also an important cultural element which contributes to evaluation use.

Given these observations, we can conclude that evaluation capacity is both an exponent and an enabler of evaluation culture. Nevertheless, at the very start of evaluation finding its way into the policy process, we can presume that some elements of evaluation culture must be in place as a trigger for evaluation capacity building. Over time, the building of evaluation capacity may strengthen the evaluation culture.

A remaining question is then how we can interpret the position and role of the four indicator categories or elements which we could not easily assign to one particular stage in the evaluation process. This does not mean that they are static in themselves, but they can be considered as variables which enable both evaluation capacity and culture to develop. Legal embedding, organisation anchoring and the coupling of evaluation with policy and managerial decisions as well as networking provides a structural foundation for evaluation culture to spread and for evaluation capacity building. In fact, these four elements can be regarded as levers for the institutionalisation of evaluation either on country level or organisational level or within a specific policy area.

The overall conclusions are reflected in the figure below. However, we do not conclude that the arrows are to be seen as unconditional causal relations, but rather as enabling relations.

To conclude, we want to stress the fact that this paper is only a first attempt to create some clarity on the interrelation of evaluation culture and capacity. We recognize that our exercise has limitations with regard to the sources selected for the literature screening and the choices made throughout the analysis. At the same time, this is a plea for more in-depth and broader research on this topic in order to verify the approach used in this paper and to fine-tune the conclusions we have made so far.

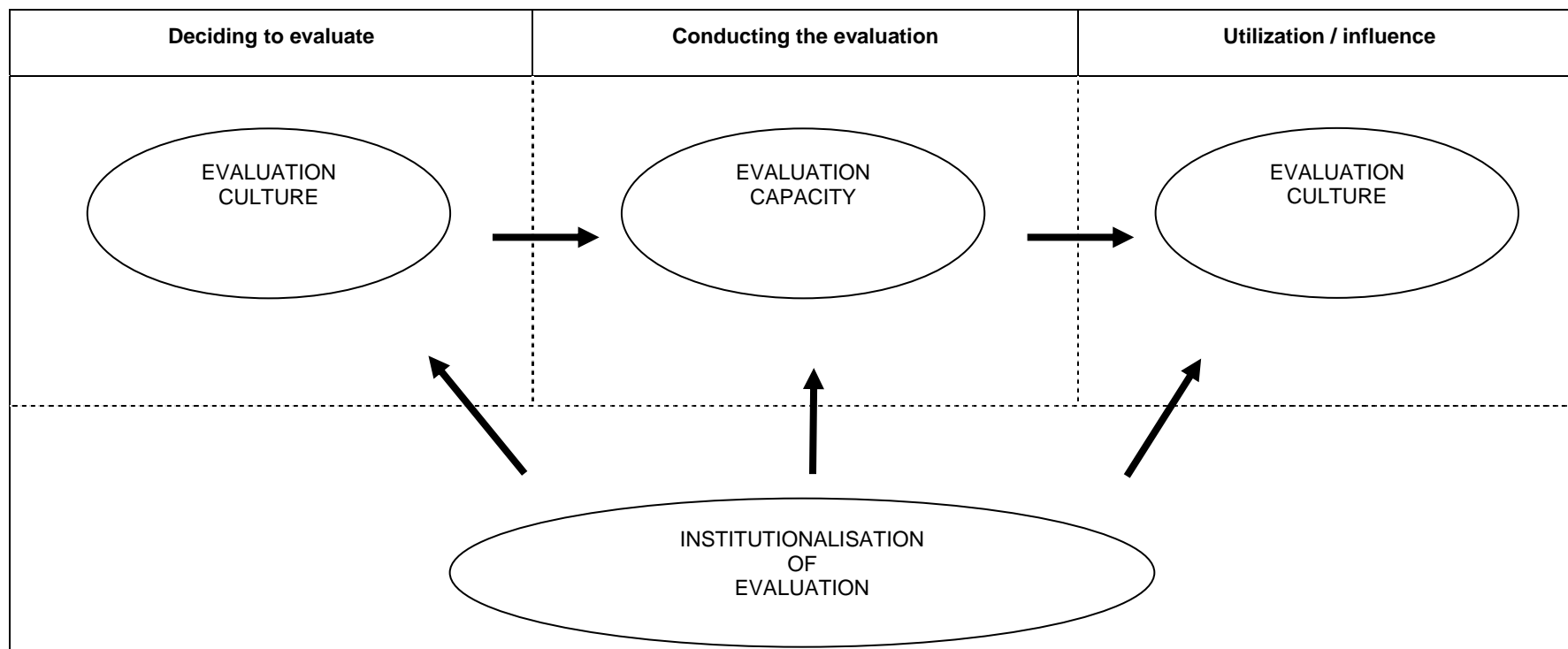


Figure 1: The interrelation between evaluation culture and capacity and institutionalisation grafted on the generic evaluation process

Bibliography

Books

Baizerman, M., Compton, Donald W., & Stockdill, Stacey Huffle (2005) "Capacity building", in S. Mathison (Ed.), *Encyclopedia of evaluation*, Thousand Oaks, CA: Sage, pp. 38-39.

Boyle, Richard, Donald Lemaire, and Ray C Rist (1999). "Introduction: Building evaluation capacity", in Richard Boyle and Donald Lemaire (eds.) *Building effective evaluation capacity. Lessons from practice*. New Brunswick: Transaction publishers, pp. 1-19.

Furubo, Jan Eric and Rolf Sandahl (2002). "Introduction. A diffusion perspective on global developments in evaluation » in Jan Eric Furubo, Ray C. Rist and Rolf Sandahl (2002) *International Atlas of Evaluation*, Transaction publishers, pp. 1-23.

Checklists

Stufflebeam, Daniel L. (2002). *Institutionalizing evaluation checklist*. Evaluation Checklists Project, Western Michigan University, The Evaluation Center, www.wmich.edu/evalctr/checklists

Volkov, Boris B. and King, Jean A. (2007). *A checklist for building organizational evaluation capacity*, Evaluation Checklists Project, Western Michigan University, The Evaluation Center, www.wmich.edu/evalctr/checklists

Articles

Bustelo, Maria (2006). "The Potential Role of Standards and Guidelines in the Development of an Evaluation Culture in Spain" in: *Evaluation*, vol. 12(4), pp. 437-453.

Compton, Donald W., and Baizerman, Michael (2007). "Defining evaluation capacity building", in *The American Journal of Evaluation*, 28 (1), pp. 118-119.

Dabelstein, Niels (2003). "Evaluation Capacity Development: Lessons Learned" in: *Evaluation*, vol. 9(3), pp. 365-369.

Horton, Douglas (2002). "Planning, implementing and evaluating capacity development", in *International Service for National Agricultural Research (ISNAR) Briefing Paper*, 50.

King, Jean A. (2007). "Developing Evaluation Capacity Through Process Use" in: *New Directions for Evaluation*, no. 116 (winter 2007), pp. 45-59.

Mackay, Keith (2002). "The World bank's ECB experience", in *New Directions for Evaluation*, 93, 2002, 81-99.

McDonald, Bron, Patricia Rogers, and Bruce Kefford (2003). "Teaching people to fish? Building the evaluation capability of public sector organizations", in *Evaluation*, 9 (1), pp. 9-29

Taut, Sandy (2007). "Defining evaluation capacity building. Utility considerations" in *The American Journal of Evaluation*, 28 (1), pp. 120

Varone, Frédéric, Steve Jacob, and Lieven De Winter (2005). "Polity, politics and policy evaluation in Belgium", in *Evaluation*, 11, pp. 253-273

Reports

Guerrero, Pablo (1999). *Evaluation capacity development. Comparative insights from Colombia, China, and Indonesia*. ECD Working Paper Series, 5. World Bank Operations Evaluation Department.

Lahey, Robert (2004). *Evaluation in the government of Canada*. Report prepared for the International Evaluation Benchmarking Seminar and Workshop. Brasilia: Brazil, 23-24 November 2004, 51p.

United States General Accounting Office (2003). *Program Evaluation. An evaluation culture and collaborative partnerships help build agency capacity. Report to Congressional Committees*. May 2003, 27p.

Online resources

European Commission (2008) The Guide, in EVALSED, The resource for the evaluation of socio-economic development, http://ec.europa.eu/regional_policy/sources/docgener/evaluation/evalsed/guide/index_en.htm (consulted on 7 April 2008).

Williams, B. (2001) Building Evaluation Capability. Unpublished paper. Available from: <http://users.actrix.co.nz/bobwill/>

Working papers

Mackay, Keith (1999). Evaluation capacity development. A diagnostic guide and action framework. *ECD Working Paper Series*, 6. World Bank Operations Evaluation Department. 29p.

Annex 1: List of specific indicators for evaluation capacity and culture, per source and classified into 17 categories of indicators (See table 1).

Boyle et al. (1999)	
Human capital (skills, knowledge, experience)	SK
Financial, material resources	FR
Sound data systems	DCM
Familiarity with social sciences	SK
Absence of corruption (political and economic)	POL
<u>Key issues that must be dealt with by governments wishing to institutionalize evaluations</u>	
Anchoring the evaluation regime	LE
Anchoring evaluation capacity within organizations	ORG
Evaluation coverage	CPS
Linking evaluation with other public sector functions and institutions	CPM
Using evaluation in decision-making	USE
Professionalizing the evaluation function	SK
Fostering demand	POL
Mackay (1999)	
<u>Barriers (and hence points of attention) for building evaluation systems in developing countries:</u>	
Genuine demand	DEM
Supply: evaluation, accounting, or auditing skills	SK
Information infrastructure: high quality financial and other performance information; accounting/auditing standards and systems	DCM
Ownership	OWN
Culture of fact-based accountability	AVE
Absence of corruption	POL
Evaluation feedback mechanisms into decision making processes	DIF
Critical mass	HR
Comprehensive approach; link with performance measurement and performance management	CPM
Presence of supportive culture or set of values- or at least the possibility to develop such a culture!!!	AVE
Guerrero (1999)	
<u>Demand</u>	
Leadership and vision	POL
Awareness	AVE
Incentives: laws and regulation	LE
Institutional set-up	ORG
Enabling environment (internal/external pressures)	DEM
<u>Supply</u>	
Staffing	HR
skills	SK
Financial resources	FR
Methods	SK
Professional evaluation standards	QLI
Information	DCM
Evaluation timing	CPM
<u>Organisation of the evaluation function and actionable steps</u>	

For evaluation to be successfully implemented, it needs to be located in several places within the governance structure to meet the demands of various markets and stakeholders	ORG
<u>General</u>	
In depth diagnosis of a country's institutional framework → country-adapted strategy	CTX
Matching evaluation capacity interventions with public sector management reforms	CPM
Mackay (2002)	
Demand as main prerequisite	DEM
Long-term commitment of government	POL
Tailor ECB according to country circumstances	CTX
Achieving an ongoing, enduring and sustainable state of affairs where evaluation is utilized	USE
ECB needs guided process (clear guidance and support)	POL
Emphasis on utilization	USE
<u>Operations Evaluation Department: Checklist for which it regards as good-quality country level-ECB</u>	
Based on a formal country diagnosis and clear action plan	CTX
Form part of a public sector management reform program	CPM
Develop and implement a customized training program for ECB	SK
Establish linkages with financial management and accountability programs	CPM
Develop linkages with statistical system improvements	DCM
Establish linkages with research initiatives	DCM
<u>Evaluation Capacity Building Results Chain: Performance indicators</u>	
<u>Outcomes:</u>	
M&E findings are used in budget decision making, in sectoral strategy making, and in line management decisions	USE
M&E findings are used by media, in parliamentary debates, and in NGO submissions to government	USE
Government structures and processes have been realigned to commission M&E findings and to feed them into budget processes and into ministries' planning and management processes	CPM
<u>Outputs:</u>	
Greater quantity of monitoring information and evaluation findings	DCM
Better quality of monitoring information and evaluation findings	QLI
Formal M&E framework is established by government	LE
Number of officials who undertake M&E training	HR
Number of officials working on M&E	HR
Number of evaluations or reviews conducted	CPS
<u>Activities:</u>	
M&E training and trainer-training offered	SK
In-country seminars provided to build awareness and strengthen demand for M&E	AVE
Stufflebeam (2002)	
Take into account the context-dependent nature of evaluation systems	CTX
Identify, support and address internal and external driving forces for evaluation	DEM
Locate the evaluation unit as a staff operation at a high level of the organization	ORG
Promote and support stakeholders' buy-in, participation, and support from all levels	OWN
Adopt and apply the evaluation field's Standards and Guiding principles	QLI
Define and apply clear, functional evaluation policies and contracts	QLI
Define and pursue clear, appropriate evaluation purposes	SK
Engage and support a capable, credible evaluation team	HR
Supply the evaluation effort with sufficient funds, facilities, equipment, services, software and technical	FR

support	
Adopt and apply appropriate evaluation models	SK
Employ a range of qualitative and quantitative evaluation methods	SK
Regularly conduct evaluations	CPS
Establish and maintain functional databases	QLI
Employ effective communication channels and mechanisms	DIF
Provide evaluators and stakeholders with ongoing evaluation education	SK
Establish and maintain a quick response mechanism to address emergency evaluation needs	CPM
Periodically secure internal and external meta-evaluations	QLI
Maintain and employ an evaluation system review and improvement process	QLI
Furubo et al. (2002)	
Evaluation takes place in many policy domains	CPS
There is a supply of evaluators, specializing in different disciplines who have mastered different evaluation methods and who conduct evaluations	SK
National discourse concerning evaluation in which more general discussions are adjusted to the specific national environment	CTX
Profession with its own societies or frequent attendance at meetings of international societies and at least some discussion concerning the norms or ethics of the profession	NET
Institutional arrangements in the government for conducting evaluations and disseminating their results to decision makers.	ORG
Institutional arrangements are present in Parliament for conducting evaluations and disseminating them to decision-makers	ORG
An element of pluralism exists, that is, within each policy domain there are different people or agencies commissioning and performing evaluations	CPS
Evaluation activities within the Supreme Audit Institution	ORG
Some public sector evaluations must have program or policy outcomes as their object (in addition to focus on outputs or technical production)	CPS
United States General Accounting Office (2003)	
The ability to systematically collect data on program results	DCM
The ability to systematically analyze data on program results	SK
The ability to systematically use data on program results	USE
Evaluation culture: Regular self-assessments to inform program improvement- Commitment to self-examination and learning through experimentation.	AVE
Data quality: credibility, reliability and consistency	QLI
Analytic expertise: Knowledge of research methods and relevant subject matter	SK
Collaborative partnerships: the sharing of resources and expertise among stakeholders	NET
Dabelstein (2003)	
Lack of demand means low impact of evaluation institutions on policy and management decisions	DEM
Policy advocacy and senior management demand	POL
Legal foundation or a firm statutory organizational regulation	LE
Evaluation unit's independence from line management	ORG
McDonald et al. (2003)	
Evaluation capability only to be understood as self-evaluation by programme staff and managers; but refers also to the effective use of external evaluations and ongoing monitoring. Hence: not only conduct evaluations but also commission, manage and use them.	USE
Equipment to successfully fish	SK
Effective distribution system	DIF
People who want to eat fish	USE

Entire fishing system that is sustainable	ORG
Hence: not just developing skills of individuals, but of the whole organization	
<u>Lessons in building evaluation capability</u>	
Develop a common evaluation framework	LE
Build knowledge about what works in evaluation in your context:	SK
Knowledge building through partnerships between government and research institutions	NET
Good evaluation practice depends on context	CTX
Systematically and visibly evaluate each stage (evaluation of the evaluation capability process)	QLI
Lahey (2004)	
<u>Developing the capacity to 'do' evaluations</u>	
Recognition of the role played by a champion/advocate for the evaluation function, to ensure the necessary resourcing and maintenance of momentum	ORG
Financial resources	FR
Trained/experienced personnel.	SK
Vehicles to train:	
Evaluator Internship Program	SK
Networking events with departmental evaluators	NET
Formal conferences	NET
Professional development series	SK
Links with other professional organizations	NET
Educating the users of evaluation	SK
<u>Developing capacity 'to use' evaluation.</u>	
Central agency driven demand for evaluation	ORG
Steering the evaluation function in Departments	ORG
Credibility building for evaluators	AVE
Developing a comprehensive evaluation plan	ORG
Evaluation in full public view (transparency)	QLI
<u>Success factors needed for effective evaluation ["Evaluation infrastructure checklist"]</u>	
<u>Cultural</u>	
Institutions are prepared to divulge information	DIF
Managers trust that assessments will be objective	AVE
Agencies are willing to be reviewed	AVE
Managers have the courage to make changes and implement recommendations	AVE
Evaluation function is prepared to evaluate itself	AVE
Relevant accountabilities have been clarified	ORG
<u>Operational</u>	
Technical, professional resources are available	SK
Financial resources are available	FR
Time is sufficient	CPM
Evaluation policies and standards are in place	QLI
Need for objectivity can be met	QLI
Authority exists to oversee evaluations	QLI
Authority exists to act on findings	USE

Baizerman, Compton and Stockdill (2005)	
Evaluation capacity building is a context-dependent intentional action system of guided processes and practices [for]	CTX
And its appropriate uses are ordinary and ongoing practises within and/or between one or more organisations/programs/sites	USE
ECB is oriented toward the organization and its structure	ORG
Social and cultural practices [and]	AVE
Personnel	HR
<u>Program level:</u>	
Designate coordinators and staff	HR
Dedicate resources	FR
Create logic models with linked evaluation plans and information systems	DCM
Create technical assistance directories	ORG
<u>Organisational level:</u>	
Designate organisational leader or champion	ORG
Develop evaluation consultation corps	ORG
Train staff	SK
Varone et al. (2006)	
<u>Existence of formal organizations</u>	
Evaluation body within the executive	ORG
Evaluation body within the Parliament	ORG
Evaluation activities within the Supreme Audit Institution	ORG
<u>Existence of an epistemic community</u>	
National society of evaluators	NET
Scientific review on evaluation	QLI
Quality standards	QLI
Bustelo (2006)	
<u>Professional consolidation of a discipline</u>	
Development of higher academic programmes for training professionals	SK
Publication of specialised periodicals	NET
Existence of professional associations	NET
Existence of a job market	NET
Development of ethical codes and normative guidelines or standards: not only indicator but also creative element that encourages a greater maturity and the institutional / professional consolidation of a discipline.	QLI
<u>Reasons for absence evaluation culture (in Spain)</u>	
Few public references to concrete practices	DIF
Professionals do not know what evaluation is and how it can be carried out	SK
No training; no possibility to become socialized to a way of thinking about judging the value of policies if there is little public awareness of such an undertaking	SK
Institutions do not have practices that resemble evaluation or these practices are not recognized as such	AVE
Lack of facilitating mechanisms to define evaluation, identify guidelines, such as standards	QLI
<u>Information gathering:</u>	
Good dose of authority and persuasion	POL
Good information systems	DCM
Adequate coordinating structures	ORG

Volkov & King (2007)	
<u>Organizational context</u>	
<u>Cultivate a positive, ECB-friendly internal organizational context:</u>	
Make sure that key leaders of the organization support and share responsibility for ECB	POL
Locate existing and enlist new evaluation champion(s) in the organization	ORG
Determine and work to increase the organization's interest in evaluation information	AVE
Determine and work to increase the organization's demand for evaluation information	DEM
Provide opportunities for sufficient input in decision-making, ensuring that people in the organization are able to use data and make decisions	CPM
Organize opportunities for socializing around evaluation activities during the workday	NET
<u>Understand and take advantage of the external environment and its influence on the organization</u>	
Identify external mandates/accountability requirements and expectations, and integrate them into the ECB efforts.	ORG
Determine if and to what extent the external environment is supportive of change	AVE
<u>ECB Structures</u>	
<u>Develop and implement a purposeful long-term ECB plan for the organization</u>	
Establish a capable ECB oversight group	SK
Generate an appropriate conception of evaluation for organizational policies and procedures	ORG
Create a strategy for conducting and using evaluations in the organization that applies existing evaluation frameworks, guidelines, and professional standards	QLI
Integrate evaluation processes purposefully into organizational policies and procedures	CPM
Make sure that a detailed written ECB plan exists, is distributed throughout the organization, and is used to assess progress	ORG
Evaluate the capacity building activities routinely	QLI
<u>Build and reinforce infrastructure to support specific components of the evaluation process and communication systems</u>	
Create organizational structures that will facilitate evaluation activities	ORG
Assign responsibility for facilitating the ongoing development and evaluation of evaluation processes	ORG
Build individuals' readiness to implement evaluation activities	AVE
Build individuals' skills to implement evaluation activities	SK
Develop and use an internal reporting/monitoring/tracking system	DIF
Develop an effective communication and reporting capability	DIF
Introduce and maintain purposeful socialization into the organization's evaluation process	OWN
Establish clear expectations for people's evaluation roles and provide sufficient time during the work day for evaluation activities	ORG
Offer tangible incentives for participation in the evaluation process	OWN
Provide or make available formal training, professional development, and coaching in evaluation	SK
Promote and facilitate people's learning evaluation	SK
Model a willingness to be evaluated by insuring that evaluations and the ECB processes itself are routinely and visibly evaluated	AVE
Build and expand peer learning structures	NET
Purposeful trust building and interdependent roles in the evaluation process	AVE
Incorporate a feedback mechanism in the decision-making process and an effective communication system	DIF
Create ongoing learning activities	SK
Provide opportunities for individual and group reflection	NET
<u>Resources: make evaluation resources available and use them</u>	
Provide and expand access to evaluation resources	ORG

Use evaluation personnel effectively	HR
Provide access to research bases that contain 'best practices'	DIF
Ensure the availability of sufficient information on how to access existing evaluation resources	DCM
<u>Secure sources of support for program evaluation in the organization</u>	
Assure long-term fiscal support from the board or administration	FR
Provide basic resources	FR
Allow adequate time and opportunities to collaborate on evaluation activities	NET
If needed, develop revenue-generating strategies to support program evaluation	FR
King (2007)	
<u>CONTEXT</u>	
External mandates and accountability requirements	CTX
Evidence of external support for change	POL
Management support	POL
Broad-based interest in and demand for evaluation information	DEM
Commitment to use information	USE
Willingness to generate new information	POL
<u>INPUT</u>	
Purposeful ECB plan	ORG
Identifiable evaluation champions (incl. managers)	ORG
Peer learning structures specific to evaluation	NET
Infrastructure to support the evaluation process	ORG
Access to evaluation resources	FR
Sufficient input into decision making	CPM
<u>PROCESS</u>	
Systematic use of the organization's ECB plan	ORG
A viable evaluation advisory group	ORG
Incentives for participation in evaluation activities	OWN
Purposeful socialisation into the evaluation process	AVE
Active facilitation of and reflection on evaluation processes	ORG
Documentation of existing evaluation resources	DCM
Ongoing and high quality communication about evaluation	DIF
European Commission (2008)	
<u>Individual level: necessary skills and competencies</u>	
Skilled people (appropriate recruitment and training)	SK
Training resources, training courses, diplomas (open to both practitioners as well as commissioners)	SK
<u>Organizational level: management arrangements and structures</u>	
Evaluation is routinely undertaken at each stage of policy and programming	CPS
Evaluation findings are integrated into decision-making	CPM
Evaluation is considered as important input for improving performance and management for results	AVE
Regular flow of evaluations is commissioned, covering the broad spectrum of policies and programmes	CPS
Follow-up procedures to ensure use of evaluations	USE
Procedures to accumulate evaluation findings	DIF
<u>Inter-organizational level: Bridges public and private bodies through networks, procedures and partnerships</u>	
Co-ordination through a network of evaluation units and functions	NET

Requirements that evaluations take place embodied in legislation, articulated policy and regulatory activity	LE
Well-defined evaluation market	HR
Culture of evaluation that values professional standards, independence, learning from experience and evidence based policy	AVE
<u>Societal level: Embeds evaluative thinking in civil society (including professional organizations) as well as in the public sector</u>	
Open and systematic dialogue between policy makers and evaluation specialists	NET
Evaluation community of practice	NET
Evaluation associations	NET
Awareness of evaluation activity and outputs	AVE
Dissemination of reports and findings	DIF
<u>Developing evaluation capacity</u>	
Architecture: locating and structuring evaluation functions and their coordination	ORG
Strengthening evaluation demand: ensuring there is an effective and well managed demand for evaluations	DEM
<u>Strengthening evaluation supply:</u>	
Professional evaluation community	NET
Independent consultants	HR
<u>Institutionalising evaluation:</u>	
Extending evaluation more widely in the public sector	CPS
Greater integration of evaluation processes into policy making	CPM
Utilization into policy making and programme management	USE
Human resources	HR
Financial resources	FR
High level political commitment	POL