The politics of evaluation: The case of cultural organizations

Francesco Chiaravalloti
Utrecht University/The Netherlands
Faculty of Humanities/Department of Media and Culture Studies

Martin Piber (corresponding author)
Innsbruck University/Austria
Department for Organization and Learning
Universitätsstraße 15
A-6020 Innsbruck
phone: +43 512 507 7569
mail: martin.piber@uibk.ac.at

Biographies

Francesco Chiaravalloti is researcher and lecturer at the University of Utrecht/The Netherlands. Previous Affiliations include the University of Groningen/The Netherlands and the Fraunhofer IPK in Berlin/Germany.

Martin Piber is professor at the Department of Organization and Learning and head of the research centre of Organization Studies at the University of Innsbruck/Austria. His research and publications focus on the management and organization of cultural organizations, the relevance of culture for society, aesthetics, and business ethics. In 2009/2010 he was visiting professor at the University of Cape Town in South Africa.

Abstract

Cultural policy always means to decide how to spend money. This is a problem, as there are usually more useful options to use financial means than money available. Hence a sound evaluation of the proposed results and the predicted outcome as well as to some extent also a comparison of different projects is required. We raise the question, if and how the interests of various stakeholders can be considered in the theoretical framework of evaluation in the field of arts and culture. This entails the question, what kind of political processes are hidden behind processes of evaluation in the cultural sector and how the relation between the purpose of evaluation, the evaluators, and the other stakeholders can be understood and explicitly addressed. Hence, this requires a focus on the inherent (political) values and interests as well as their contexts.

Key words: evaluation, politics, policies, stakeholder, cultural organizations.

Word count: 3803

First draft – please don’t quote without permission of the authors.
I. Evaluation – theory and practice

The literature on evaluation developed in the public sector in the 19th century – namely in the areas of education, healthcare, and justice (literature review required). At the beginning of the 20th century actions against unemployment, social insurances and other public policy programs came to the fore. Especially in the 60ties and 70ties of the 20th century, programs in the area of education have been evaluated. Even more: It was a necessary condition to conduct an evaluation in order to be able to invest public money from a certain size onwards. Already at that time politicians had a vested interest in supporting the programs launched with evaluative action.

In the 70ties we saw a strong professionalization of evaluation and its actors. First professional bodies as the Joint Committee on Standards for Educational Evaluation were founded. The idea of evaluative practices has been integrated into manifold conceptualizations of New Public Management and meanwhile evaluation also entered the sphere of private institutions. With management ideas like Total Quality Management, Benchmarking, EFQM, Auditing, and Controlling evaluation is spreading into the key fields of managerial action. As it gains additional legitimacy with that move, the relevance for public spending is even increasing. Evaluation understood as “… an applied inquiry process for collecting and synthesizing evidence that culminates in conclusions about the state of affairs, value, merit, worth, significance, or quality of a program, product, person, policy, proposal, or plan” (Fournier: 2005: 139) develops to become a key social practice in order to take decisions what is important and what is not in different contexts.

On the one hand conceptualizations of evaluation are consciously used in order to make organizational or societal decisions. It addresses values, merits, worth, significance, or qualities of various targeted items (Fournier 2005). On the other hand evaluation proceeds towards being a general contemporary practice inherent in various social and economic fields of society. In this vein, evaluative practices have also gained an essential importance for cultural programs, organizations and issues. In general evaluations serve different purposes: legitimacy, cognition, development, and control. In this paper we focus especially on the connection of these functions with interests and influence from various stakeholders.

Evaluative practices can be distinguished in two different streams: First, the evaluator compares the target object with an abstract state of the object representing a desirable situation. This abstract state can be outlined with the help of targets, experience, values, or theoretical considerations about good practices. Second, the evaluator can compare the target object to existing benchmarks. That can be similar objects, organizations, processes, or actions. Insofar, one key component of evaluative practice is the explicit and implicit process of comparison. In this vein also cultural enterprises and projects are compared with past actions, programs, experience from the evaluators’ practice, targets, expectations of different stakeholders, or with professional standards. However, it is definitely a complex question, to find adequate benchmarks and in some cases a lack of benchmarks will lead to an abstract evaluation.

II. Four generations of evaluation

To understand the development of evaluative practices, we build on the four generations of evaluation of Guba & Lincoln (1989) which step-by-step includes political arguments in the interaction of evaluators and the evaluated (object). For the evaluation of museums analyzed from the frame of Guba & Lincoln, see Gstraunthaler and Piber (2012).

The 1st generation of evaluation is described as the “measurement paradigm”. In this paradigm the role of the evaluator is exclusively technical. It is expected that s/he knows the full panoply of available instruments, in order to measure any variable within the spectrum of the targets. His ‘superior knowledge’ about the targets to be achieved and the measures to be used enables a seemingly clear reproduction of ‘the performance’ on different scales.
The 2nd generation was triggered by serious deficiencies of the 1st generation, as the measurement process of the 1st generation takes place without reference to any previously agreed targets. Performance measurement and evaluation in the 2nd generation compares the results with the objectives or the mission statement of the organization. This refers to the abstract process mentioned above. The role of the evaluator is thus to describe patterns of strength and weakness with reference to clear-cut targets. Measurement is no longer treated as the equivalent of evaluation but is redefined as one of several tools that might be used in its service (Guba/Lincoln 1989: 28).

Breaking down a mission or a set of strategic targets into manageable figures is sometimes far more difficult for cultural organizations than in the profit context. In contrast to companies which pursue the maximization of their earnings, cultural organizations face multifaceted targets. As an illustration to this problem we refer to the mission statement of a Museum of Applied Arts: “The museum must continue to develop as a place of awareness, free of external influences, by advancing the discourse between the interplay of experience and perception. Since it navigates along the borders that separate art and awareness from innumerable forms of fashionable consumption, the museum’s articulation of qualitative assessment makes it a central forum for resistance against the widespread loss of meaning pandemic in contemporary popular culture.” Herewith it becomes obvious, that any evaluation or measurement of the targets mentioned in this statement cannot be done without various judgments. Therefore it is no surprise that in the 3rd generation of evaluation judgement and expertise enter the arena. A judgment requires a real decision, which can presuppose the encouragement of one and the dis-encouragement of another value system. In contrast, a simple choice would be a mere application of clear-cut and predefined rules within one single value system. There is no room for special expertise and interpretation. Judgement requires both: expertise and the decision between different value systems. Expertise is a quantum of knowledge that is recognized as suitable to give a judgement (Madsen & Polesie 1981). In some cases power issues complement the expertise. With the assignment of judgement to an individual, evaluation leaves the ground of objectivity and gets more and more subjective and – in consequence, political. As any evaluation incorporates a political dimension, a conscious handling and a critical reflection on judgements become imperative. Starting from here, judgement is considered as an integral part of evaluation (Guba/Lincoln 1989: 30).

In another turn Guba & Lincoln develop the 4th generation towards a “responsive constructivist evaluation.” It takes into account that any notion of performance is constructed by the participating actors. Therefore a thorough understanding and an evaluation of performance has to be executed in a dialogue of the organization and its relevant evaluators. In addition, any evaluation has to take into account that it inevitably has political components. This is the reason why evaluators have to act knowingly and responsibly towards this insight (Abma 2006).

III. Evaluation in the field of culture

Until now evaluation theory didn’t directly refer to the field of culture. When applied on the field of culture, several key points have to be problematized. First, the research paradigm plays a key role for the definite gestalt and scope of evaluation. Positivist-empirical research with a twist towards objectivity and generalization and constructivist approaches with their paradigm of competing realities heavily influence the relevance of the evaluative processes itself (Chiaravalloti & Piber 2011). Second the field of culture is always embedded in a network of various stakeholders with different interests (Brooks et al. 2002). Therefore, evaluation has always to be understood in the context of the interests – especially in the context of the interests of the contracting body of the evaluation.

If we take the performing arts as an example, we have the artists, the audiences, the management, the support staff, the art community, and the general public as entities with a considerable interest in the organization. Additionally the interests of these groups might not only be different but also conflicting. Hence, organizing and managing means as well handling these different interests. Sometimes even evaluators themselves are counted as organizational stakeholders (Markiewicz 2005).
Hence, we have to raise the question, how the interests of various stakeholders can be considered in the theoretical framework of evaluation in the field of arts and culture. This entails the question, what kind of political processes are hidden behind processes of evaluation in the cultural sector and how the relation between the purpose of evaluation, the evaluators, and the other stakeholders can be understood and explicitly addressed. Hence, this requires a focus on the inherent (political) values and interests as well as their contexts.

This means that cultural organizations have to pursue a broad variety of objectives in order to fulfill the expectations of their stakeholders. Definitely the targets and interests of these stakeholders are not well aligned in many cases. In addition they might change continuously and in unforeseeable, individual patterns. Furthermore single patterns are under continuous discussion.

In this contested field different stakeholders try to pursue their interests with evaluative practices. Often they attempt to improve their power-position by using legitimating arguments and comparisons. Insofar it is crucial to understand in what name an evaluative process is undertaken. It might be possible, that the genuine interests of some stakeholders are hidden behind the notion of evaluation and that the organizers of the evaluative action drive the process with the main target of pushing interests.

In this context two important issues have to be addressed: First we consider the role of metric representations (Selwood 2001). They play a key-role in attracting legitimization of outside constituencies. Second we address the role of experts in several approaches of evaluation. The full impact of evaluative action will only be achieved, when we can relate to well-respected experts in the field of action. In some contexts, the representing power of metrics as well as the affirmative impact of experts are consciously used in order to back the results of evaluative practices.

**IV. The politics of evaluation**

Evaluative action is always entangled in a field of stakeholders. Furthermore any evaluation has an instructing party. Hence, even seemingly objective methods and measures might camouflage the interests of various stakeholders. In this case it is their selection and the weighting of measures, which might influence the results of the evaluation. If someone acts s/he has a reason for that. With this seminal reason the basis for political action is already given.

It has to be questioned if the reason to act corresponds to societal standards, if the reason to act is ethical. We are facing not only the question, if it is enough to act according to humanitarian and democratic values. It is also the question, if it is possible to judge at all, which humanitarian and democratic values to choose. This piece of work does not stand out in order to judge (non-)appropriate actions. We rather want to problematize certain acts in the area of evaluative practices. “It is now widely argued and widely acknowledged that evaluation has to do intrinsically with values and value judgements“ Cordray and Lipsey (1986).

In the literature the topic is approached from different angles: The arguments of Datta (2011: 279) can be summed up in the following table:

<table>
<thead>
<tr>
<th>Point</th>
<th>Counterpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>justice requires stakeholder participation and decision making</td>
<td>justice requires best allocation of program resources and the strongest possible designs</td>
</tr>
<tr>
<td>Feasible, sound evaluation technically requires stakeholder participation and decision making</td>
<td>Sound evaluation requires awareness of but avoids capitulation to biases, agendas, self-interests</td>
</tr>
<tr>
<td>utilization requires attention to all stakeholders and negotiated designs</td>
<td>without fullest assurance of independence and integrity, utilization will be minimal</td>
</tr>
</tbody>
</table>
Politics is a necessary condition and the realistic evaluator can expect to contribute to illumination, to negotiate but not control. Politics can blow up any and the evaluator needs to be observant about it.

Evaluators are saturated with their own values and thus inherently subjective. Evaluation can only seek to reveal different realities. Objectivity is an illusion. Evaluators can and should disqualify themselves if there are conflicts of interests in values or anything else. Objectivity and impartiality are possible and essential.

Designs that are more process-oriented, qualitative, grounded where appropriate in indigenous values and epistemology, and capacity building for evaluands best fit social justice and utilization concerns.

<table>
<thead>
<tr>
<th>Table 1: The politics of evaluation observed from different angles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source: Modified from Datta (2011: 279)</td>
</tr>
</tbody>
</table>

The main question is, if we can or should decide with or without consulting the main stakeholders of the evaluated objects. If the answer to this question is at least partly yes, the evaluator has to take into account the power and the influence of the respective stakeholders. As a consequence, especially the following two questions arises: How the interests of the least powerful stakeholders are to be addressed? How can we handle the influence of powerful stakeholders in the evaluation process? On the basis of these basic considerations theorists found different approaches to deal with the problem.

Vestman & Conner (2006) see three different approaches on the relationship on politics and evaluation: The value-neutral, the value-sensitive, and the value-critical position. The value-neutral approach draws a firm line between politics and evaluation. The evaluator does not entangle with any stakeholder position. The value-sensitive approach constructs evaluators, who might raise the arguments, but who don’t take position for one or the other stakeholder. Finally, the value-critical evaluator is aware of different political positions. (S)he is conscious about necessarily having own values and decision criteria. Hence this approach opens up the possibility for a connection of politics and evaluation, which can’t be avoided. In this sense, evaluation by its very nature includes political statements (Weiss 1987).

Datta (2011) distinguishes two main groups of approaches: On the one hand the populist family considerably takes into account the values and interests of stakeholders: “That is, evaluators’ values can lead to insisting on a participatory, empowerment, values clarification, or social justice stance as their gold standard. The evaluator negotiates, facilitates, proselytizes for this stance […]” Datta (2001: 281.) On the other hand the public interest family advocates for independent evaluations: “These evaluators nonetheless believed in, valued, and honored the need for trustworthy, politically unbiased, independent information for decisions” Datta (2001: 282).

Taking into account the arguments of Guba & Lincoln (1989) in the third and fourth generation of evaluation, sound evaluative action cannot be done ‘from nowhere’. On the one hand evaluators are entangled in their personal view of the world and developed their values and understandings, which might have supported their external perception as renowned experts in the field. On the other hand, it could be advantageous to understand the interests and values of the stakeholders involved in an evaluative project, in order to be able to better judge situations, capacities and possible future developments. Hence, a value-neutral and partly even a value-sensitive position according to Vestman and Conner seems to be neither a realistic assumption nor helpful to thoroughly understand the intricacies of the project. Consequently, we will draw the conclusion rather from the value-critical approaches.
V. Conclusions

On the basis of the discussion presented above we can draw general and specific conclusions for the cultural sector: On a general level the approaches presented showed, that the interests of the stakeholders are a key element of evaluative projects. Previous research showed, that evaluation is consciously or at least unconsciously done from a defined perspective of values and understandings. Evaluation from nowhere is in danger to be lost in arguments between conflicting paradigms. Thus, the interests of the participating parties are on the agenda – if they are consciously addressed by the evaluators or not. Herewith, we argue that the consideration of these interests – might enhance and is a key question of ethical behavior in evaluative projects. However, this requires enhanced analyzing capacities, as usually powerful interests will have more means to be presented in favorable colors. On the other side it doesn’t mean that the values and interests of the disenfranchised are always to be considered.

Second, the continuous engagement with stakeholders has significantly improved the methodology of evaluation. We have learnt how to cope with vested interests and political arguments. We learnt that addressing politics in many cases improves the analytical quality of evaluative action and helps to understand subtle relationships among stakeholders. Hence, it is advisable to include a dialogue with the relevant stakeholders in evaluation methodology (Palumbo 1987).

Consequently it is far better to be aware of the sometimes hidden political agenda than to intentionally look the other way, as soon as political dimensions of evaluative projects occur. Even in this case the interests will openly or less openly find their way in the pen of the judgement.

Especially for the cultural field we can draw two preliminary conclusions: If we address cultural policy decisions we argue for a sound analysis of the stakeholder landscape in order to be able to evaluate qualities and trade-offs of policy and managerial decisions. This enables a more thorough picture of the present accounts and future developments. Sometimes even conflicting interests could have their cultural value in itself.

On the level of cultural organizations we would see a rising potential to understand evaluative action not only as a necessary intervention from the outside or as a mere external accountability tool but rather as an opportunity for internal management practice and strategy development. This understanding can spur discussion within the organization and trigger desirable learning effects between different stakeholders within and outside the organizational field.

For policy makers in the field of arts and culture further research in the field will show, how the conceptualization of the third and fourth generation of evaluation can support evaluative action of political agencies deciding upon fundings for cultural projects. On the one hand the funders have to take into account their or the policy makers’ goals and on the other hand a responsive evaluation would also consider and actively acknowledge the pluralistic interests of funding aspirants. Exactly in this case a stakeholder dialogue could lead to a deeper understanding of projects, targets, and their impact as well as their relevance from different angles.

VI. References


