Chapter 2

How to assure the Quality of Grey Literature: the Case of Evaluation Reports

Markus Weber, Federal Office of Public Health, Switzerland

2.1 Introduction: Grey Literature needs quality control

The production of grey literature has grown considerably compared with the more traditional type of academic literature. This is due to several developments in modern society and many of them are covered by the two key words «knowledge society» and «internet». One consequence of this is that today readers have more difficulty in judging the quality and relevance of what they read. This is especially true for information professionals in all sectors of government, academics, business and industry and stays a major challenge for developments in our information society.

At first this seems to be mainly a problem on the demand side of grey literature (the reader), but is of course also a problem perceived and being tackled by the supply side: How can one be assured of the quality of grey literature in a similar way to the quality hallmark of “white literature” i.e. peer reviewing in scientific journals?

This chapter is about a quality assurance system¹ for a specific category of grey literature, evaluation.² It was developed by the Competence Centre for Evaluation (CCE)³ of the Swiss Federal Office of Public Health (FOPH) as a six

-
- 1 «Quality assurance», «quality control» and «quality management» are used as synonyms in this text.
 - 2 «Evaluation is the process of determining the value (contribution to societal well-being), quality, and/or justification of the object in question. Its judgment is based on the use of (mostly) social science research methods and procedures for the systematic collection and analysis of data, not necessarily routinely available, regarding various aspects of a public measure. The judgment criteria most commonly applied include RELEVANCE, EFFECTIVENESS, and EFFICIENCY, and occasionally, SUSTAINABILITY.» (FOPH Swiss Federal Office of Public Health 2005)
 - 3 The CCE is responsible for commissioning and managing the FOPH's evaluations of public health measures - mostly of health promotion and prevention programmes and projects³. It is an internal service that has to assure the studies' scientific quality, ethical conduct and trustworthiness, on the one side, and, on the other, their usefulness. Most studies are mandated to external university research institutes or private evaluation consultancies.

step system for assuring the quality of the commissioning and management process as well as the evaluation products, especially the written reports. The “effects” of such work have been well recognised; «Managed by the CCE» is increasingly perceived in Switzerland as a quality label for evaluation studies. For example, The Swiss FOPH and its CCE are mentioned in several international and national studies as a good and successful example of how to handle evaluation in public administration (e.g. Fornerod 2001; Jacob and Varone 2002; Widmer et al. 2001).

The CCE’s quality assurance system is described in detail in section 2.2. Whilst much of its experience is concerned with public health evaluations, the system itself could probably also be applied to other areas concerned with producing knowledge for grey literature. The implications of such a transfer are discussed in 2.3, and some conclusions are presented in section 2.4.

2.2 An example of a quality assurance system for commissioned evaluation studies

2.2.1 Overview and objectives of the system

By introducing and using a quality assurance system the CCE aims at achieving two main objectives; firstly, that the studies are conducted according to sound evaluation standards, including the scientific quality of the applied methods and methodology and secondly that the products are useful and practicable, i.e. the studies need to address questions and draw conclusions that are relevant to the needs of a wide and varied audience, and come up with a set of recommendations that can be implemented.

2.2.2 General description

In most cases, the final product of an evaluation takes the form of a written report; quality control is most often therefore focused on this end product. However, evaluation is a process as well as a product and thus there are many steps along the way that need to be controlled for quality. It would not be very sensible to just come in at the end of a study and judge the quality of a report; rather it has to be steered from the beginning. The CCE has standardised processes, guidelines, models and checklists that are used to guide the process from A to Z, i.e. from the first request for a study to actual commissioning, accompanying the study throughout, assessing the report (meta-evaluation) and discussing and supporting a work plan for the utilisation/implementation of the study results.

Figure 1 shows the 6 main steps of the evaluation process from a commissioner’s point of view. The many sub tasks that have to be considered within each

step are described on the following pages. Many of these are supported by CCE checklists, models, etc.

Figure 1: Commissioning an evaluation in 6 steps

The process starts with the CCE's analysis of an evaluation request that it receives from the specialist internal service needing the evaluation (e.g. the HIV/AIDS prevention unit). If the request is considered justified and necessary the CCE then develops the evaluation specification (step 2). After calling for offers, the CCE, together with its internal partner, selects an external evaluation team (step 3). The study is then commissioned and the CCE regularly meets with the external team, reviews the tools developed for data collection and analysis and manages the contract (step 4). At the end (and sometimes midterm) the study's findings are received, considered and a plan of action is drawn up to put them to effective use (step 5). The final part of the process includes following up on the "action plan" about one year later to see what was done and what was achieved.

This system has been successfully used for several years (FOPH Swiss Federal Office of Public Health 1997; Läubli Loud 2004) and, more recently, adjusted to take into account the Swiss Evaluation Society's (SEVAL) quality Evaluation Standards (Widmer et al. 2000)⁴. As SEVAL's standards underpin several aspects of the CCE's quality control system, they and their role are explained in the next section (2.2.3), before going on to present the details of the CCE's 6 steps (2.2.4).

2.2.3 The SEVAL Standards

The SEVAL developed its set of good practice standards for guiding the conduct of evaluations in Switzerland. They refer to the processes involved in seeking and collecting data for making judgements and producing the written report. They describe what an ideal evaluation should be like in an ideal situation. They also promote the need for self-reflection and professional discussion between commissioners, evaluators and any other stakeholders so as to build a common ground for

4 Approved in 2001 by the Swiss Evaluation Society SEVAL, www.seval.ch [site visited 04.08.2009]

the execution of a study. As such it is hoped that the risk of a study being instrumentalized or manipulated is reduced.⁵

The 27 standards are grouped into four quality dimensions: Utility, Feasibility, Propriety and Accuracy (cf. Figure 2). The objective of each dimension is as follows (Widmer 2005):

- The **8 Utility standards (U)** guarantee that an evaluation is oriented to the information needs of the intended users of the evaluation
- The **3 Feasibility standards (F)** ensure that an evaluation is conducted in a realistic, well-considered, diplomatic and cost-conscious manner
- The **6 Propriety standards (P)** ensure that an evaluation is carried out in a legal and ethical manner and that the welfare of the stakeholders is given due attention
- The **10 Accuracy standards (A)** ensure that an evaluation produces and disseminates valid and usable information

Figure 2: The 27 standards ordered by the 4 quality dimensions (Widmer 2005).

The SEVAL Standards define the expectations of an evaluation but do not specify either the methodology or the methods to be used. Overall, they share the same concerns and objectives as those defined by the CCE: sound scientific quality and ethical conduct (especially through the accuracy and propriety standards) and production of practical knowledge (utility and feasibility standards). The standards are categorised according to the quality dimensions. But they are not all equally relevant to every evaluation (e.g. subject to which methodology is applied) and certainly not to every phase of an evaluation (from initial planning to utilisation).

5 This common ground is usually referred to as an «evaluation culture». A more detailed description on the background of an «evaluation culture» and on the current status quo of its development in Switzerland can be found in Läubli Loud (2004)

Those who use the SEVAL Standards need to relate them to their specific evaluation methods, needs and situation. The head of the CCE team was involved in the development of the SEVAL Evaluation Standards (Widmer et al., 2000) and has systematically advocated their use in the commissioning process ever since.⁶

2.2.4 Step by step through the CCE's system

The step by step description of the CCE's quality assurance system provides a good overview of how each part of the system works. Even though all of the 6 steps consist of several sub-steps (cf. 3 to fig. 8), only those of key importance to quality assurance are described in further detail.

Step 1: Analysing the evaluation request (pre-evaluability)

At this point, a CCE staff member is asked to study the request and determine its main aim or purpose, the key evaluation “needers and users”, and whether it is worthwhile and feasible, e.g. can the information be obtained in some other way such as through performance review, or audit and/or can the expected information be delivered “in time” enough to be useful i.e. to help decision making? Background knowledge has to be gathered and processed to help clarify the main purpose of the requested study and the intentions and expectations of different purposes of different stakeholders.

Figure 3: Step 1, Analysing the evaluation request (pre-evaluability)

Step 2: Drawing up the Evaluation specification

This is a very important step since it makes a considerable contribution to determining the final quality of the study. It sets out the key questions that need to be addressed, the scope and focus of the study, who needs what information and how

6 A “slimmed down” version of the SEVAL standards was later produced specifically for the use of commissioners within the Swiss federal administration (Widmer 2005).

and by whom it is intended to use the findings, and most importantly, the time frame for receiving the results.

Figure 4: Step 2, Drawing up the evaluation specification

Step 3: Selecting an evaluation team

Once the evaluation specification is finalised the study is put to tender and the CCE, together with the internal specialist service needing the evaluation, examine the offers and select a suitable evaluation team.

Figure 5: Steps 3, Selecting an evaluation team

Step 4: Managing the contract

In this step the CCE staff member responsible for managing the project keeps in regular contact with the external evaluators for monitoring progress, reviewing the tools, helping with gaining access to data and organising regular feedback sessions. Often intermediate results can already be very useful to the internal commissioners of the study (end user). Such regular contacts are therefore essential for identifying and bringing forward useful information “along the way”.

Figure 6: Step 4, Managing the contract

Step 5: Considering the findings

In step 5 the evaluation's findings, conclusions and recommendations are presented in both written and oral form to the commissioners and other stakeholders. However, the quality of the report is first checked by the CCE and this "meta-evaluation"⁷ phase is the most important for assuring the overall quality of the work. As the end users have to take responsibility for interpreting the findings (what do these mean to their work?) and developing a dissemination and action plan (what has to be done consequently?), they depend on the CCE's quality control of the overall evaluation

Figure 7: Step 5, Considering the findings

A set of checklists supports the CCE in its examination and evaluation of the final report. Step 5 is therefore the second most important and time consuming task of the overall system.

Step 6: Following up on the utilisation of evaluation findings

This step is less relevant for the CCE's control of an individual study, but more for the control and accountability of its overall products and services. It is of paramount importance for legitimizing evaluation in an institution such as the

⁷ A meta-evaluation is defined by the FOPH as the scientific and ethical quality control of an evaluation (FOPH Swiss Federal Office of Public Health 2005).

FOPH. Evidence on the usefulness of evaluation per se and of CCE's work in general has to be demonstrated – the CCE therefore seeks out evidence on how the knowledge and lessons highlighted through the evaluation studies have helped in making further decisions, fundamental changes and/or slight modifications to the public health strategies and measures studied. Towards this end, it compares and contrasts the intended “action plan” with what was finally implemented and why.

Figure 8: Step 6, Following up on the utilisation of evaluation findings

2.2.5 Short discussion of the quality assurance system

The 6 step system's procedures and tools help the CCE to achieve its two main objectives (assuring its evaluation studies are of sound scientific quality and produce useful and usable knowledge which can be put to practical use). Scientific quality and professional ethical conduct is assessed through a strict review of the final product (meta evaluation of the final evaluation report) - the last of the quality assurance procedures. Steps 1, 2 and 4 are the most staff resource-consuming; however, given the guidelines and checklists produced to support each step along the way, the 6 steps of the system can now be accomplished in much less time than was possible before.

2.3 What for other study types and constellations?

In this last part of the chapter we have described the system used to help the CCE assure its partners of quality evaluations. The procedures are based on some general principles of quality assurance and therefore should be readily transferable to other “grey” literature areas. Below, we suggest some possible ways of transferring “good practice” principles to other areas albeit adapted to the needs of other contexts.⁸

⁸ Perrin (2006) argues that in most cases speaking of «best practice» should be avoided. Before transferring «some successful way of doing something somewhere» to your own

For evaluators: adapt slightly

For evaluators themselves, much of the CCE's system for overseeing the commissioning process cannot be applied directly. However, the SEVAL evaluation standards address evaluators and commissioners alike as they set out standards to be addressed both during the process and when producing written reports. In step 5, the evaluation team should therefore conduct a meta-evaluation of their own work.

Other kinds of studies than evaluations: Use other, relevant standards for quality assessment or ...

In many other areas standards of good practice exist and can act as a starting point for developing a quality assurance system e.g. in clinical research one could take aspects from the CCE's 6-step-model (or any other defined process) and combine them with the quality standards of "good laboratory practice GLP" of the OECD⁹.

... use GLISC guidelines as an instrument

A very useful tool is the «Guidelines for the production of scientific and technical reports: how to write and distribute grey literature», also called «Nancy style» (GLISC 2007). These guidelines are mainly focused on writing (and distributing) accurate, clear and easily accessible scientific reports in different fields, but they also «include ethical principles related to the process of evaluating, improving, and making available reports, and the relationship between GL producers and authors» (GLISC 2007, p. 1). «The Guidelines state the ethical principles in the conduct and reporting of research and provide recommendations relating to specific elements of editing and writing» (GLISC 2007, p. 2).

Minimal procedure: Clarify everything along process

The main element of the CCE's system is of course the same as for any good research: clear, transparent and well documented procedures for guiding the procedures and conduct of the work.

2.4 Conclusions

The quality assurance system described above was specifically developed by the CCE to oversee the evaluation process and product. The system is based on the fundamental premise that the "end" is the result of the "means" used to get there. Thus the quality of the evaluation report as an example of "grey literature" is as good as the processes, tools and conduct applied throughout the study. It therefore makes a significant contribution towards helping readers judge the quality of what they read. Could such a system be generalised? Given the wide variety of grey

situation it has to be adapted to your actual context. This necessary adaptation is better expressed by using the term of «good practice».

9 Organisation for economic co-operation and development OECD, http://www.oecd.org/department/0,3355,en_2649_34381_1_1_1_1_1,00.html [site visited 04.08.09]

literature it is of course not feasible and nor the aim to develop a universal system for all producers of grey literature. But a basic set of steps for guiding the production of quality output in the field of grey literature could go part way towards improving its overall quality.

References

- FOPH Swiss Federal Office of Public Health. Guidelines for Health Programme & Project Evaluation Planning. 1997. Berne, FOPH. URL: [http://www.health-evaluation.admin.ch \(>tools >guidelines\)](http://www.health-evaluation.admin.ch (>tools >guidelines) [site visited 04.08.09]) [site visited 04.08.09]
- FOPH Swiss Federal Office of Public Health. Glossary of Evaluation Terms. 2005. Berne, FOPH. URL: [http://www.health-evaluation.admin.ch \(>tools >glossary\)](http://www.health-evaluation.admin.ch (>tools >glossary) [site visited 04.08.09]) [site visited 04.08.09]
- Formerod, Serge. A quoi et à qui servent les évaluations ? 2001. IDHEAP.
- GLISC. Guidelines for the production of scientific and technical reports: how to write and distribute grey literature. Version 1.1. 2007. GLISC Grey Literature International Steering Committee.
- Jacob, S. and Varone, F. L'institutionnalisation de l'évaluation des politiques publiques : expériences étrangères et scénarios pour la Belgique. 2002. Louvain-la-Neuve, Université catholique de Louvain.
- Läubli Loud, M. 2004, Setting Standards and Providing Guidelines. The Means Toward What End?, *Evaluation* 10, 237-245.
- Perrin, B., 2006, How Evaluation Can Help Make Knowledge Management Real., in: R.C.Rist and N.Stame, eds., *From Studies to Streams: Managing Evaluative Systems*. (New Brunswick, NJ).
- Widmer, Thomas. Leitfaden für Wirksamkeitsüberprüfungen beim Bund. Instrument zur Qualitätssicherung gestützt auf die Evaluationsstandards der Schweizerischen Evaluationsgesellschaft (SEVAL-Standards). Erarbeitet im Auftrag des Bundesamts für Justiz. 2005. Berne, Bundesamt für Justiz.
- Widmer, Thomas, Landert, Charles, and Bachmann, Nicole. Evaluation Standards of the Swiss Evaluation Society (SEVAL Standards). 2000. Berne, SEVAL. URL: [http://www.seval.ch >standards](http://www.seval.ch >standards [site visited 04.08.09]) [site visited 04.08.09]
- Widmer, Thomas, Rüegg, E., and Neuenschwander, Peter. Stand und Aussichten der Evaluation beim Bund. 2001. Zürich, University of Zurich.