

## Quick and Dirty: Bureaucracy-Driven Evaluation

I entered the media assistance evaluation ‘field’ through documents. The crafting of and the following up of documents is a central part of the development practitioner’s material work (Escobar 1995:146). The bureaucratic logic is underpinned by an assumption that if the documents could be perfected, poverty would be eradicated (Jassey 2004). This chapter pays attention to the “moments of document-making” (Riles 2006:18) in media assistance evaluation. By focusing on the evaluation document as a material, technocratic artifact, it reveals the histories of documents, and the ways in which documents are made in anticipation of the future careers of documents (Brenneis 2006; Riles 2006:18). Tracing the document-making moments and the anticipated careers of evaluation documents illuminates how the evaluation document is shaped by and instrumentalized within the development bureaucracy. As such, this chapter connects with important critiques of development, examining how the bureaucratic concerns with efficiencies, systems, procedures and policies, entrench top-down, supply-driven aid.

What was striking about the almost 50 media assistance evaluation reports for this research was the similarity. My analysis of publicly available media assistance evaluation reports over a 10-year period (2002–2012) found that a typical media assistance evaluation report can be summarized in the following ways: It is undertaken at the mid-point or end of the funding cycle, probably by a commissioned consultant who is usually paid for about a three-week period to review the project documents, carry out stakeholder interviews or focus groups and observe the

running of the operation, perhaps with some minor additional methods. The types of stakeholders included in interviews (or other similar, qualitative methods) were the donors, the implementing agency staff, partner staff, and trainees or other participants (see Noske-Turner 2015 for the full analysis). This pervasive ‘template’ was familiar to the evaluators interviewed. It was the “classic model” (Renneberg 2013, personal communication, 26 February) and “the known approach” (Susman-Peña 2013, personal communication, 24 July). It is reasonable to extrapolate, therefore, that the template is generalizable beyond the sample of evaluation reports analyzed for this study. In general, this template, or classic model of evaluation of media assistance, did not enable high-quality evaluation reports or provide evidence of on-going social or governance changes. Evaluators themselves are frustrated with the usual “five-day visit with just a bunch of key informant interviews and document review” which, one evaluator stated, means “you can write a report ... but you can’t really give a good evaluation” (Abbott 2013, personal communication, 26 July).

This chapter involves a deeper examination of this template, finding that methods used are chosen either due to explicit direction in the Terms of Reference (ToR), itself a procedure-driven document, or because consultants follow the same process due to the absence of time and resources to do anything else. Reports are required to be concise, around 30 pages, with easily extractable information to enable subsequent reporting to parliament. The evaluation document works as a bookend, bracketing the project with the Project Design Document (PDD) written at the beginning. Together these documents form the primary material artifacts, or the “documentary reality” (Smith 1974) of development managing. These methods, procedures, and forms are not deliberately chosen in order to achieve the best possible evaluation, but rather these document-making moments are shaped by the bureaucracy to, as one research participant put it, “feed the angry beast of donors” (Renneberg 2013, personal communication, 26 February).

This chapter extends existing analysis of the systems and of evaluation with particular reference to media assistance. By tracing the formation of evaluation documents into a material form, I seek to make visible the relationships between ‘the template’ of evaluation reports and the systems of bureaucracy that make this the default. Despite the push for participatory and learning-based approaches, particularly strong in communication for development and social change (see Lennie and Tacchi 2013), the core purpose of evaluations in international development

continues to be bound, first and foremost, as a management tool for accountability processes (Chouinard 2013). The need for accountability is rarely questioned, much less any questioning of how accountability is pursued and the costs associated with its attainment (Chouinard 2013:239). As this chapter will show, the systems put in place to manage evaluations, which are intended to ensure development effectiveness, actually impede effective practices in evaluation and have a series of damaging consequences. The analysis repeatedly indicated that factors beyond questions of the most suitable methodology and methods were influencing the choices being made about evaluation, making it imperative to engage in the critiques of aid and development in a macro sense. The demand by the bureaucratic systems of donors for documents, in particular forms and at particular times, drives evaluation decisions. It is the bureaucratic system that shapes the methods used, the sources valued, the issues included and excluded, and what is actually being evaluated. The choices made about evaluation are largely passive and procedural. But while the systems that shape decisions about evaluation appear to be inevitable and immovable, since ultimately donors control these systems, to close this chapter, I point to some emerging examples, which, in small ways, resist the default evaluation from within the system.

### HISTORIES: TOWARDS ‘PROCEDURALIZATION’

The media-missionaries mind-set meant that the increased expectations in relation to evaluation and evidence across the development sector during the 1990s was slower to reach media assistance. Evaluators commented that 10 years ago, there may have been some limited efforts to collect data, but without any commitment and planning, the data was almost useless. For example, consultant and scholar Maureen Taylor said, “About 10 years ago, nobody was doing anything except counting, and they were doing a poor job of counting” (2013, personal communication, 28 August). Evaluators perceived a gradual groundswell of attention to evaluation by media assistance organizations, and sensed increasing “demand” from donors to produce more and better-quality evidence from monitoring and evaluation (Warnock 2013, personal communication, 9 April). But while a growing post-media-missionary mind-set from the late 1990s to the early 2000s brought an increased interest in evaluation, there were competing imperatives. There were increased efforts to consistently achieve better evaluations and more transparency, leading

to the development of sets of procedures to guide ‘best practice,’ but, at the same time, there were increased pressures on delivery and less time to engage with the evaluation process. To explore this phenomenon further, I draw upon Anderson, Brown, and Jean’s concept of “proceduralization” (2012:65–82). These authors point out that procedures are a genuine response to an endeavor to be more efficient through streamlining, simplifying and standardizing repeated tasks so that they are undertaken in ways that are more consistent, transparent, and reflective of best practice. However, while procedures sometimes achieve this, the authors found that both donors and recipients saw downsides to the increasing number of procedures in aid and development—in particular, that procedures can be counter-productive and disconnected from the original purposes and objectives sought. They therefore use the term ‘proceduralization’ (a corollary to ‘bureaucratization’) to refer to “the codification of approaches that are meant to accomplish positive outcomes into mechanical checklists and templates that not only fail to achieve their intent, but actually lead to even worse outcomes” (2012:67).

The procedures of evaluation featured heavily in independent consultant Robyn Renneberg’s accounts of doing evaluations for AusAID. The procedures she outlines include negotiating the ToR written by AusAID which, when an agreement is reached, is followed by the contract and submission of an evaluation plan, which is then taken to a joint forum where the program management group are asked to comment on the methodology and commit to it (2013, personal communication, 26 February). Although intended as a consultative process (described as such in the AusAID *Monitoring and Evaluation Standards* (ADG Quality Performance and Results Branch 2013)), the procedures to create the ToR served to limit the flexibility. ToRs are written either using formal templates or from existing samples and adapted by AusAID for use in particular projects (AusAIDnotes 2013, personal communication, 17 June), and although officially there is “consultation” and opportunities for “comment,” donors largely remain in control. The consultation process is primarily the donors’ “way of getting buy-in from their stakeholders” (Renneberg 2013, personal communication, 26 February), rather than being used to build an evaluation plan that responds to the needs of other stakeholders. The procedures continue after the evaluation itself. Renneberg describes a series of meetings, draft reports and summaries after the “mission” facilitated by AusAID, before the final report is eventually signed off (2013, personal communication, 26 February).

The ToRs are core mechanisms used to formalize and systematize the production of the evaluation document, having the effect of specifying the required form of the evaluation. In many instances, the ToR will outline suggested methodologies, and in some instances, traces of the ‘template’ style of evaluation (a review of the key documents as specified, followed by a one- or two-week visit to consult with stakeholders) are visible even in the ToR. Very often, the contents and length of the document are also stipulated. The ToR controls what is included, by outlining the purposes and objectives, and, crucially, what is excluded, as indicated by statements such as “it was not in this reviewer’s ToR to look at the provision of this kind of information from governments to public” (Myers 2011).

The ToR is part of a web of documents, each of which is highly inter-referential, and each of which shapes the evaluation. There is a particularly important and defining relationship between the PDD (or other similar documents created at the beginning of a project), the ToR, and the resulting evaluation document. For example, Renneberg explains how, by linking to the existing documents, particular evaluation purposes and questions are defined:

In the context in which I work, which is usually for a donor, there’s a very clear purpose and that is to actually assess the progress – or the completion – the effectiveness against a design. So you don’t go in with some broad general idea, ‘I’m going to see how this is all going’, you go in with a design document and a series of documents that have changed the design over time. (2013, personal communication, 26 February)

Therefore, although there are concentrated systems shaping the document-making moments towards the end of the project cycle, the final evaluation document is anticipated from the projects’ conception and, throughout the life of the project, is planned and scheduled in the PDD (or similar). In this way, the PDD, and the evaluation document, form brackets around the project and, reinforced by the ToR, this relationship has the effect of concentrating attention on what was proposed, rather than on what actually happened. The creation of the PDD and the ToR are therefore key document-making moments in the creation of evaluation documents.

Corresponding with increased proceduralization are increased pressures on delivery. While the expectations, systems and resources for

evaluation have increased, standardized procedures are increasingly used in place of spending time working through the specifics of evaluations in ways that were possible in the past. C4D and media assistance consultant Birgitte Jallof describes the changes in time pressures as this relates to the design of evaluations:

The donors are very pressed [for time]. There was a time where you spent time with the program officer, discussing approaches and so on. In the [19]90s I worked [with] fantastic people ... sitting in DANIDA or Sida for days, and we designed the tools, tested and compared and so on. That doesn't happen anymore. [Donor staff] are so pressed, they don't have time. (Jallof 2013, personal communication, 6 March)

Research participants from media assistance organizations and donors similarly discussed the “pressures of delivery” and of “getting things done” (Testa 2013, personal communication, 18 September),<sup>1</sup> and a lack of space and time to reflect on broader issues or alternative approaches to evaluation (AusAID01 2013, personal communication, 17 June; CCAP06 2013, personal communication, 23 May).<sup>2</sup> Increases in procedures can impact on the quality and usefulness of the resulting evaluations, where evaluations become just “another box that they need to check off in their extremely full schedules” (Susman-Peña 2013, personal communication, 24 July).

Proceduralization is also evident in the timing of the production of evaluation reports, which is neither arbitrary nor a considered decision based on when impacts might manifest. There are clear patterns in the timing of when media assistance evaluation reports are undertaken, which corresponds closely with the ‘packaging’ of assistance into one-, two-, three- or five- year cycles (Anderson et al. 2012:35). Discussing these issues, a staff member of a donor agency said, “So mid-term evaluation—it was set up in the contract to be done, so that’s why—I mean,

---

<sup>1</sup>Statements made are in the context of BBC Media Action’s approaches in 2013. BBC Media Action advised that some of its views have evolved since the time of the interview, and perspectives shared by Testa do not necessarily represent the current views of the organization.

<sup>2</sup>With the exception of consultants, who were able to choose whether to be anonymous or not, interviews are anonymized and coded by the organization e.g., AusAID# refers to interviewees employed by AusAID, CCAP# refers to interviewees who were employed by CCAP, etc.

I'm new—so that's what has to happen"<sup>3</sup> (AusAID01 2013, personal communication, 17 June). The implementation of evaluation procedures is triggered by systematized time periods rather than by active decisions about when an evaluation is needed. For media assistance and other forms of communication and social change where changes often take place over long periods of time, these timeframes, produced to satisfy the demand for measurable results, are for those involved—not least the evaluators—seen as “totally unrealistic” (Renneberg 2013, personal communication, 26 February). Renneberg explains how the donor's imperative to report to parliament (or similar) can lead to vague statements such as “while the evidence is not clear, it is reasonable to assume that at this point in time the impact in this area is blah-blah-blah” being inserted into documents “knowing full well that [the impact] probably isn't.” (Renneberg 2013, personal communication, 26 February).

Finally, resourcing patterns are a highly influential moment of document-making. In contrast to the ToR, which represents a deliberate act of document-making, the effects from resourcing patterns on evaluation are largely unintentional. Media assistance evaluators frequently used terms such as “quick and dirty” to describe the evaluations resulting from the one or two weeks generally allocated for in-country research. Constraints on time and budget limited the methodologies that could realistically be used. Although there may be the appearance of flexibility to adapt the methodologies listed in the ToR and “add any other ideas of who you'd like to meet or how you'd like to do this,” the caveat is always that it must be within the allocated time and budget (Myers 2013, personal communication, 20 March). Jallof, one of the consultants known in this field for her use of participatory approaches to evaluation, similarly explained how time and budget directly impacted on her evaluation designs:

Quite often ... I don't have time to carry out a whole Most Significant Change process, because it takes about a month and it is more expensive than the ordinary quick and dirty three-week thing. (Jallof 2013, personal communication, 6 March)

---

<sup>3</sup>While the reliance on procedures is understandable when staff are newly appointed in their roles, at the same time, evaluators indicated that high turnover of staff often impacted on evaluation processes (Myers 2013, personal communication, 20 March; Renneberg 2013, personal communication, 26 February).

A further concern raised among evaluators was that resources for evaluation were only made available at the end of the project, or, at best, mid-way and at the end. Once again, due to the lack of alternative options, methodologies were limited to what is possible, rather than what might be most useful. Evaluators lamented that although “You wish that you could do it by the book,” and that “You could have been there at the beginning, middle, and end” (Myers 2013, personal communication, 20 March), entrenched budgetary procedures means that evaluation resources are held until the end of the project. This has dire impacts on the quality of evaluation. Consultant Susan Abbott advocated for more ‘frontloading’ of evaluation efforts, meaning a commitment to collecting baseline data and doing good monitoring throughout the duration of the project. Without this, she says, “you really can’t evaluate something. You can write a report, which is what happens, and you can collect success stories, and you can review lots of stuff, but you can’t really give a good evaluation” (Abbott 2013, personal communication, 26 July).

Procedures do, of course, have a role to play in evaluation and in development more generally. Procedures clarify the basic expectations and can help to implement an organization’s understandings of best practice consistently. However, if fulfilling procedures becomes a proxy for active engagement to adapt best practice principles to the specifics of the context, and if procedures become inflexible requirements rather than guides, procedures can reduce the effectiveness of evaluations rather than improve them. Imposing mandated evaluation processes and procedures can lead to a “compliance mentality,” or “mechanical” implementation, both of which reduce the utility of evaluations (Patton 2008:108). There is no inherent problem with the methods (document reviews and stakeholder interviews) per se; these are standard and common qualitative methods that can be implemented rigorously or superficially. The real problem is that the media assistance evaluation ‘template’ is not dominant by deliberate design, but rather as an outcome of the bureaucratic system. The recent history outlined by research participants points to an overall increase in proceduralization, which sees greater value placed on evaluation, while simultaneously meaning more templates, less time, less engagement, less creativity, and less flexibility.

## ANTICIPATED CAREER OF DOCUMENTS

The moments of evaluation document-making as outlined here are shaped by the anticipation of “future moments when documents will be received, circulated, instrumentalized and taken apart” (Riles 2006:18). The ‘utilization’ of an evaluation has become a foremost criteria for evaluations according to the 1994 *Joint Committee on Standards for Evaluation* (Patton 2008:26–29), indicating that the uses and utility of evaluation is recognized as both a significant challenge and a priority for the profession. This section outlines how the document functions within the development system, which values certain kinds of data for checking off boxes, and thereby shapes the document itself.

### *“Donors Love Numbers”*

One unequivocal observation among in-house evaluators, evaluation consultants, and donors, was the centrality of evaluation reports as a function of the donor’s reporting responsibilities. In order to fulfill this function, the donors desire easily extractable information, ideally in quantitative form. While the AusAID staff interviewed in Cambodia placed significant value on and interest in qualitative data, even going so far as to say that understanding impact requires qualitative data, quantitative data are necessary for their own parliamentary reporting obligations. Their annual reporting process requires filling in another document, called the Quality at Implementation (QAI) report, which involves applying scores to specific aspects of the project, such as implementation, and monitoring and evaluation. These are monitored and audited by senior staff from AusAID (AusAID notes 2013, personal communication, 17 June). The American equivalent is the F-Process Indicators (Foreign Assistance Indicators), which go to the U.S. Congress. U.S.-based consultant Susan Abbott explains how the demand for the F-Process Indicators does not necessarily produce useful indicators for media assistance agencies themselves:

You still have to keep in mind that this goes to the US Congress ... and you can imagine [that] if you’re a member of Congress given such a composite sketch, along with all the other thousands of things that they have to read, it makes some sense in some weird way. But for our little

community, it's not that helpful and it doesn't really help make the case for what we do. (Abbott 2013, personal communication, 26 July)

The consequence of the anticipation of the QAI report or F-Process Indicators is that evaluators, both in-house and consultant, are highly aware that they need to provide donors with quantified data within the evaluation documents they provide. That “donors like numbers” (Testa 2013, personal communication, 18 September) was a repeated observation. A staff member from ABCID expresses a sentiment representative of those of many others:

I think that [the donors] also are subject to requiring those statistics for their reporting. So we also have a clear understanding that they also need to be able to demonstrate [results] and quant[itative data] is important in that realm as well. (ABCID02 2013, personal communication, 18 November)

Research participants understood the pressure to produce quantitative outcomes “even if that’s not the best measure of impact” (Testa 2013, personal communication, 18 September). Consultants reported a conscious effort to “find something that you can quantify,” knowing that some “quantification does soothe donors” (Warnock 2013, personal communication, 9 April). Furthermore, failure to provide quantitative data in evaluation documents has, in Kitty Warnock’s experience, led to criticisms from the donor (DFID) through their internal auditing processes (2013, personal communication, 9 April). Donors love numbers, therefore, not only for reassurance, but also in anticipation of their own reporting responsibilities within the system.

### *Anticipating Decision-Making*

Coupled with the need for data that satisfies reporting requirements is the assumption that evaluation documents perform an important role in future funding decision-making. From the donors’ perspective, evidence is the basis of future funding decisions, as a staff member of AusAID in Cambodia explains, “With evidence, we can make a case for a continuation of funding. And actually, we should be continuing to fund this sort of stuff. I would like us to. But I have to make arguments, and I have to use evidence” (AusAID01 2013, personal communication, 17 June).

In the highly competitive funding environment, NGOs look to evaluations as a mode of securing future funding. Research participants from BBC Media Action referred to evidence collected through monitoring and evaluation processes as contributing to making a good “business case” (Testa 2013, personal communication, 18 September), or as vital in “business development” (C4DNGO01 2013, personal communication, 19 June), indicative of the relationship between evidence and future funding.

However, although evaluation documents are expected to contain evidence, and in turn enable informed decision-making, in practice, evaluation documents do not routinely lead to expected decisions. The Media Map Project found contradictions between the donors’ stated evaluation goals and the actual use of the evidence in decision-making, stating that, “Ultimately, we found little evidence that M&E was changing the landscape of funding decisions, other than the now ubiquitous requirement to provide some sort of M&E component to project proposals” (Alcorn et al. 2011). This situation is reflected in the broader evaluation field, with Patton’s *Utilization Focused Evaluation* (2008) being a direct response to concerns over the lack of engagements with evaluation findings. In other words, the increased emphasis on evaluations has only increased the number and complexity of the procedures in evaluation, not the use of them, despite, or perhaps more accurately, because of the fact that so many of the procedures are shaped by the needs of the bureaucracy rather than by deliberate choices about the best practice or needs for each particular situation.

## DEPENDENCE ON INDEPENDENCE AND THE AURA OF TRANSPARENCY

Commissioning an external consultant to produce ‘independent’ reports to fulfill quality assurance processes is part of the language of evaluation policies for many donors and agencies. For example, the United Nations Evaluation Group’s norms and standards stipulate the need for independence and impartiality in conducting evaluations (United Nations Evaluation Group 2016). There is a pervasive link between independence and the perception of credibility, and this was evident in the sample of evaluation documents analyzed in this research. However, this practice is in need of serious critique. The problems raised here are founded on

two concerns. The first is that the commissioning of consultants is used as a relatively controlled and repeatable technique for “operationalizing accountability” (Brenneis 2006:44) to create an “aura of transparency” (Riles 2006:19). The second is the flaws inherent in basing the credibility of evaluations on a notion of a detached neutrality of external evaluators (Chouinard 2013). This critique of objectivity is more than a simple problem of clashing epistemological positions; it relates to the ways consultant evaluators are commissioned to respond to predetermined evaluation designs, and how the systems undermine the independence that consultant evaluators are able to exercise. This occurs to the point where a dependence on independence comes at the cost of thorough evaluation.

Documents are a technology that make things auditable and construct legitimacy (Brenneis 2006; Power 1996). Furthermore, techniques and technologies of accountability and audit are often produced to meet rhetorical, rather than instrumental, outcomes (Aho 1985). The use of independent consultants in evaluation can be seen as an example of this, where the technique serves the rhetorical function of justifying “an activity about which there existed ... considerable suspicion” (Aho 1985:22), or at least a bureaucratic demand for “quality assurance.” The term ‘quality assurance’ itself is significant, originating in engineering discourse in relation to specific parameters such as product defect rates, but which has become a generalized, abstract management concept, with the effect of commodifying quality assurance services (Power 1996:300). Of course, this is not to say that such techniques are illogical, but they do need to be understood as more than the neutral procedures of calculations they are purported to be. One common mode of achieving auditability is the use of experts (Power 1996). In the evaluation of media assistance, consultants are not only seen as bringing expertise and objectivity, but also functioning as a repeatable technique that creates the aura of transparency. In this way, the single most important criterion for a credible document is that it was produced by a consultant; the quality of the evidence itself, which as shown is compromised by the inflexibility and routine resource allocation practices, is of secondary concern.

While evaluation consultants are professionals who are commissioned to provide truthful accounts of a project, there is a need to critically engage with the assumptions and limitations of the idea that consultant evaluators can provide absolute objectivity. This analysis follows a similar line of enquiry to that which has been explored by several authors (e.g., Chouinard 2013:244; Fetterman 2001:94–96), including by Cracknell, who has cautioned that consultant evaluators:

... approached their task with not only their own preconceived ideas, but also with values acquired (probably without even realizing it) from the donor agency, which has its own criteria of success, often embodied in the Terms of Reference (however obliquely). (2000:336)<sup>4</sup>

While there are some important benefits of having an independent evaluator in terms of the openness with which statements of a sensitive nature can be made, the technical expertise they may bring and the weight that the findings will have, it is important to problematize the objectivity of independent evaluations.

First, several evaluators described situations where donors and others commissioning evaluation reports had a sense of ownership over the evaluation report, which challenged the consultants' roles as independent evaluators. As a commissioned piece of work, in the end, the evaluation document belongs to the client. Australian consultant Robyn Renneberg describes evaluation work as a job where her "primary client" is AusAID; "So if AusAID is contracting me, in the end that's who I'm serving" (2013, personal communication, 26 February). While Renneberg is at pains to state that she is "quite fearless about giving them feedback about where they've done things badly" (Renneberg 2013, personal communication, 26 February), she and other evaluators interviewed noted that at times this relationship had implications in relation to the content of the reports. This tension was particularly pronounced when independent evaluations were commissioned by the implementing organization (the NGO). For some, the commissioning process unambiguously meant that the document was "their property" and that the commissioners could do whatever they want with it (EvaluationConsultant04 2013, personal communication, 13 March).<sup>5</sup> A researcher who had been contracted by ABCID for research on the CCAP project similarly felt that "we are not that independent," expressing an awareness that "because [we] take the money from the client, [we] have to fit with the client's interests all the time"<sup>6</sup> (Consultant05

---

<sup>4</sup>A similar critique of the role and positioning of consultants is available as provided by Carothers (1999:287).

<sup>5</sup>Anonymized at the request of the research participant.

<sup>6</sup>This quote required significant editing for clarity and readability.

2013, personal communication, 15 June). In another (unrelated) case, when changes were requested by a commissioning NGO,<sup>7</sup> the evaluator felt that this was “not very ethical” (Myers 2013, personal communication, 20 March). Consultant Mary Myers describes her experience of being asked to make changes to a section of her report, which stated that aspects of the project were not going well:

They came back and said, ‘Well what are we going to tell our board with this evaluation that you’ve done?’ I didn’t say it in so many words but I said ‘that’s your problem, I’ve said what I’ve said, you can tell the board what you like, but I would prefer if you didn’t cut stuff out of my writing’. But the trouble is I was paid by those people so in a way I suppose they felt that I should do what they wanted me to do, i.e., [in] a final version I should cut out certain words or nuance them ... So I had a long phone call with the director ... we went through word by word and [the director] said ‘can you just nuance this word a bit or that word’. I said ‘well OK if you want, but it’s not very ethical, I mean, I’ve done what I’ve done.’ (Myers 2013, personal communication, 20 March)

Similar ethical dilemmas associated with the commissioning process are a known challenge in evaluation (Fitzpatrick et al. 2004:419–421; Patton 2008:25). The situation described above illustrates the complexity of commissioning independent reports, which are expected to be simultaneously owned by the commissioner and independent of them. This was a genuine dilemma that evaluators actively engaged in. Their responses depended on the context and the agreements made up front, as well as the level of agency an evaluator perceived they had. For example, Myers, a consultant with more than 15 years of experience, negotiated all the requests for nuances, only agreeing to change the wording where she felt it was ethically appropriate. Locally engaged researchers may not feel that they are in a position to negotiate with the commissioner to undertake or write the evaluation differently. In all cases, however, the client-consultant relationship is far from straightforward, and to refer to these situations as entailing “objectivity” would be to over-simplify the ways in which evaluation documents are produced.

Second, evaluation reports are written with an awareness of the potential audiences, which has some subtle influences on the content of the

---

<sup>7</sup>The name of the NGO is withheld at the request of the research participant.

report. In interviews, evaluators explained that when they know that the reports will be public documents, they write with a self-consciousness of the possible consequences of the documents. Evaluators described “being very careful ... I weigh every word. I’m very conscious that I may cause harm [to communities]” (Jallov 2013, personal communication, 6 March). Another described writing a “warts-and-all” review when such documents were internal, compared to in the new “transparent environment” where there is a need to be “more circumspect” (Renneberg 2013, personal communication, 26 February). Once again it is important to see these decisions as carefully considered and ethically based, with awareness that any criticisms may be taken out of context and may have other negative repercussions, especially for local communities.

Finally, the choice of consultant is often a deliberate decision based on the expectations about the kinds of approaches they are known to use. In this way, the commissioners are subtly shaping the report. While critiques of the role of consultants as mediators (in Latourian terms) in development networks have pointed to consultants’ interests in maintaining these relationships to secure future work (Hayes and Westrup 2014), this is just one side of the equation. In interviews, consultant evaluators were conscious that they were commissioned based on the commissioner’s knowledge of their reputation and past work. As a representative example, Kitty Warnock describes her sense of independence, while at the same time showing a self-awareness of the expectations underpinning the interaction: “I would say I had complete freedom. Obviously, because they knew me, they knew what I was likely to do. So we weren’t strangers” (Warnock 2013, personal communication, 9 April). Other evaluators were also conscious of their reputations for using certain approaches, such as participatory approaches (Jallov 2013, personal communication, 6 March), or conversely, their reputation for not using these approaches (Myers 2013, personal communication, 20 March). These factors shape who regularly commissions them and with what kinds of expectations.

For these reasons, while a consultant might be “independent,” there are limitations to the notion that this means that objectivity and neutrality are possible aspirations and expectations of this process. There are, of course, many advantages to commissioning a consultant, but objectivity, a much more complex expectation, is compromised by the commissioning process, the ownership and agencies of those

involved, and the sensitivities and ethical dilemmas faced. For these reasons, aside from satisfying the rhetorical functions of the bureaucracies' aura of transparency, such systems do not lead to the best evaluation practices and outcomes for any of the actors concerned. There are much better ways to make use of the consultants' methodological and comparative expertise, including using them as guides or "evaluation coaches" (Fetterman 2001; Hanssen et al. 2008) in the evaluation planning stages, or as 'scrutineers' of research by project teams. Both these uses would be more likely to lead to richer sets of evidence on which to make funding decisions, but this would mean relinquishing the dependence on independence as a bureaucratic operationalization of legitimacy.

### COSTS OF THE SYSTEM

Here's a really common sentence [in a final report]. 'The project collected an enormous number of output indicators showing how many people were trained and blah blah blah, however, because of limited M&E efforts, there are few outcomes and no impact measurements.' And then you're stuck with key informant interviews. (Taylor 2013, personal communication, 28 August)

By now, some of the costs of 'evaluation as usual' where key document making moments are shaped by bureaucratic imperatives and procedures will be apparent. However, it is worth taking stock of these in order to build a case for the need for alternative practices in the post-media-missionary environment.

As has been implied, while the resulting document performs its purposes in the bureaucracy, the content of the evaluation report is often quite predictable. Documents that are typified with statements such as the one described above by Taylor are unsatisfactory to everyone involved. The "quick and dirty" process creates a report, but not an evaluation (Abbott 2013, personal communication, 26 July). From the donor perspective, there is a mismatch between the highly rigorous and often qualitative information they craved about the impact of media assistance, and the pressure to provide quantitative information simplified to the point of irrelevance for reporting to parliament. As has already been alluded to throughout this chapter, there are many negative consequences and missed opportunities resulting

from increased proceduralization and demands for evaluations to take a certain form, which cumulatively result in prioritizing the document over insightful evaluation. There are two other consequences worthy of note.

Firstly, the bureaucratic systems prevent even the possibility of participatory approaches being an option. Instead, proceduralization leads to very donor-centered evaluation. One staff member of ABCID explains this situation thus:

In terms of true participatory evaluation approach[es] I'm not sure whether or not the donor context actually facilitates that, because obviously at the end of the day, evaluation questions are driven by the program design rather than by the community, which is the core of participatory research. So that's the first challenge. (ABCID02 2013, personal communication, 18 November)

Evaluators saw a conflict between the product and the process and explained how, within the limitations of the system, they did their best to “actually listen” to enable a “mutual process” of “knowledge production” (Jallof 2013, personal communication, 6 March). Myers notes with irony that in producing the reports, only the very few “top people” would read what was written and that, even though they are valuable in some respects, the reports are largely inaccessible for the recipients of development:

The actual result is often just a 30-page report, which will also often be in a language they can't understand anyway. But I think they are valuable when they're done well for almost all stakeholders, apart from the direct beneficiaries on the ground (*irony intended: laughs*). (Myers 2013, personal communication, 20 March)

Participatory approaches are not only limited by the centrality of donors' needs in the making of the evaluation document (e.g., through the ToR), but are also undermined by the fact that resources for evaluation are only made available at the end; the very short time periods made available for evaluators; and the insistence on commissioning independent consultants to the role of objective expert. Once again, the importance of the document is a barrier to engaging local evaluators. Referring to an evaluation in Papua New Guinea, Renneberg said that the two local evaluation team members were “really good:”

But none of them can write to the standard required by donors. One of the things that AusAID hasn't got its head around yet is letting go of the professional standard of reporting. So in a way, the perfect evaluation wouldn't involve me at all, it would be done locally by locals and fed back to AusAID. But until AusAID is ready to let go and accept different standards, that's not going to happen. (Renneberg 2013, personal communication, 26 February)

It is therefore vital that any advocacy of participatory approaches to evaluation, so central to CfSC and thus increasingly relevant in the post-media-missionaries era, considers the systems and structures leading to passive decision-making, rather than simplifying the absence of participatory approaches in evaluation as an active decision made by donors to reject these types of approaches.

The second consequence of the proceduralization is the way it feeds insecurity. At the same time as there has been increased attention to evaluation, and more procedures and systems in place to produce evaluations, there is a sense of heightened insecurity around the future of media assistance. There is a significant and problematic contradiction between what stakeholders hope or expect evaluations will achieve, and what they actually produce, and this is particularly relevant to the donor context. The system produces a document that conforms to quality standards while simultaneously devoid of in-depth insights or useful evidence. Donors expressed a desire to see more than the kinds of "philosophical evidence" typical of the media-missionary mind-set—certainly more than basic reports that simply show that the project team has "been busy," and instead they want "hard data" around development impacts (AusAID01 2013, personal communication, 17 June). There is an expectation, therefore, that the evaluation will deliver high-quality insights based on evidence, which is in direct contradiction to the documents routinely produced by the systems and procedures in place.

There is a deep level and long-term insecurity in the media assistance. The related field of C4D has had many more public discussions and forums on the need for evidence, with the UN Inter-Agency Roundtables of C4D leading to a series of large research projects, including a partnership project that I have worked on with UNICEF C4D.<sup>8</sup>

---

<sup>8</sup> *Evaluating Communication for Development: Supporting Adaptive and Accountable Development*, ARC Linkage LP130100176.

Media Assistance has not had events or initiatives of this level and profile, and so the pressure to continually defend this as yet largely unproven, and complex development approach, permeates the field:

There's this feeling in media development that at any moment it could all be over. The donors will say ... 'we have lots of priorities; media is just not one of them'. Ever since I was at IREX everyone had this sense of 'will there ever be money again for media development?' ... And yet, lo and behold, here we are, at least for me, more than 13 years later and there's still media development and it's still going. (Abbott 2013, personal communication, 26 July)

A research participant from ABCID similarly described AusAID's position on media assistance as inconsistent and changeable, but on the whole "fairly ambivalent about media development" (ABCID04 2013, personal communication, 14 November). This environment of insecurity increases the risks involved in identifying project failings, which are higher in a field such as this than in other sectors where the evidence is more settled, as consultant Tara Susman-Peña explains:

I've kind of come to think of donor funded media, and I guess any sort of intervention as a bit of a trap, because if you admit that you failed, then what can the donors do but take away your money or not fund you again. But if you don't admit that you failed, how can you learn from anything? So it's ... difficult, [within] the financial structure of everything to really be able to learn from evaluations. (Susman-Peña 2013, personal communication, 24 July)

The repetition of practices that produce insufficient evidence, through broadly bureaucratically driven practices, therefore, stifles sector-wide knowledge generation, innovation, and improvement.

## EMERGING ALTERNATIVES

This section, which intended to make visible the structures and systems within which evaluation is called to perform certain functions, serves to enable all subsequent discussions of best practice principles to be informed by the context. It is also important to highlight the small, emerging examples of resistance to the system, within the existing structures. These examples set the scene for rethinking alternative ways to

plan and undertake the monitoring and evaluation of media assistance within the context of existing development systems.

Since 2011, when I began researching this topic, some significant shifts have been observable. Each of the three main media assistance organizations included in this study—ABCID, Internews and BBC Media Action—has restructured its organization to accommodate more in-house researchers, both in-country and at the head office, and has been arguing, with increasing success, for evaluation and research budgets of up to 15% of the overall projects' funds. The most compelling example is BBC Media Action, which has received a large, multi-country grant that includes significant research resources. According to Adrienne Testa, who at the time was the senior research manager for BBC Media Action, in 2013 the organization employed 40 research staff (2013, personal communication, 18 September).<sup>9</sup> These resources have allowed BBC Media Action to take ownership and agency over the research, rather than delegating these decisions to donors at the end. In the interview with Testa, it was clear that the major change the "Research and Learning" (R&L) team has implemented was to front-load their evaluation efforts. Testa emphasized how the preparatory work, including country information planning sheets, involves the project staff and research staff working together to "pin down the objectives," enabling more focused evaluation (2013, personal communication, 18 September).

This front-loading enables implementation-level learning. Although most research participants said that monitoring and evaluation should be about learning, the sense of insecurity, together with the production of evaluation documents following the completion of the project, limits opportunities to learn. Front-loading evaluation efforts subdue insecurities by enabling continual learning, so that the conversations frame failures as an issue of the past that has been solved, rather than as a final post-mortem.

Secondly, front-loading can transform the role of the donor's appointed consultant evaluator into more of a scrutineer of the quality of data, analysis, and integrity. In the case of BBC Media Action, the independent consultant uses the data and reports provided by the

---

<sup>9</sup>It should be noted here and in elsewhere that Testa's views represent BBC Media Action's practices in 2013, and do not necessarily reflect current views and practices.

BBC Media Action R&L team, rather than being the sole source of evaluation data. An in-house researcher at BBC Media Action says they are conscious of the need to “demonstrate to the donors that we are maintaining credibility and objectivity in the research,” noting that “the feedback from the evaluators was that we were too honest ... that we were being quite hard on ourselves and that we were very frank about our shortcomings” (Testa 2013, personal communication, 18 September). This is in keeping with other research on internal or participatory evaluation processes, which find that these types of evaluation processes tend to generate more critical judgements than those by external evaluators have, due to their deeper knowledge of the program and the personal stakes they have in improving effectiveness (Fetterman 2001). This is in direct contrast to the fears embedded in the dependence on independence approach. This further shows that when there is an environment largely stripped of the fear of being exposed as a failure, and which is “frank in that [it says] ‘look, this isn’t working’ or ‘we need to do things better’” (Testa 2013, personal communication, 18 September), organizations and donors alike can benefit from rigorous and useful evaluative evidence.

This type of approach was evident in one other case from the sample of evaluation reports analyzed in this study, which used Outcome Mapping. This evaluation undertook extensive monitoring throughout the life of the project. The consultant evaluator was engaged to “provide a degree of assurance” that the collected data is “reliable” (Graham 2009). In both cases, stakeholder interviews are still used, but the purpose is to check that the existing data corroborates the responses from key people. Unlike other media assistance organizations, this increased investment in in-house research means that BBC Media Action can speak with some “confidence” about impacts, to the degree of being able to compare differences across countries, and in relation to different types of program outputs (Testa 2013, personal communication, 18 September).

With direct reference to the BBC Media Action model, ABCID has been working to replicate the internal resource structures of research by setting up their ‘Insight and Impacts’ team, and by having a team of in-country researchers for each project, albeit on a much smaller scale. This has involved advocating for larger percentages of the project budget to be allocated to research, and advocating for the value of formative research against the usual concentration of research resources at the end of a project. The research participants from ABCID noted some

challenges in communicating the value of such a move to donors, but some of the benefits of these changes are emerging.

## EVALUATION VS. THE BUREAUCRACY: CONCLUSIONS

The remarkable consistency among the media assistance evaluation documents that were analyzed compelled an examination of the context in which decisions about methodologies were being made. These questions remained at the forefront as evaluators began to describe a separation between the process of evaluation, which facilitated learning and improvements, and the document, which only ticked a box. The same questions persisted when hearing of the ways AusAID staff craved evidence, especially qualitative evidence, but needed quantitative numbers and performance checking for their own reporting. It was clear that much of the evaluation process is geared towards producing a 30-page document as required by the system.

By drawing on the notion of documents as artifacts, I have sought to unpack the processes of production and the imposed constraints by the particular form, and used this to understand the influence of the donor bureaucracies in the creation of evaluation documents. Using these perspectives makes coherent the mechanisms triggering the document-making moments. For over a decade, well-meaning development planners have developed sets of procedures, policies, templates, and checklists intended to standardize best practices across the organization. While bringing greater awareness to the value and seriousness of evaluation for media assistance, this has simultaneously reduced the flexibility, the engagement, the specificity, and the level of agency that stakeholders have. The evaluation procedures reduce the moments of deliberate decision-making. Instead, since the quality assurance process requires a completion report, evaluation funds are held until the final weeks of a project, a consultant with no prior knowledge of the project is commissioned for one or two weeks in-field, and the consultant is explicitly directed to check the performance against the original plan. This compels a default to the 'template' style of evaluation; that is, a document review and stakeholder interviews.

Therefore, while the post-media-missionaries era implies an appetite for evidence beyond blind faith, the resulting evaluation document satisfies the bureaucracy but is largely irrelevant to those involved in the project. The root of the problem is that the bureaucratic systems construct

a document that complies with quality assurance, but that contributes little to achieving rigorous, insightful, useful, and relevant evidence and evaluation. Instead, the bureaucracy's principal demand is that the document be produced by a consultant, so that the document has the 'aura' of being independent and, in turn, that the system has the 'aura' of transparency. The resulting document is produced with minimal time, minimal data, and highly specified goals, with the only claim to integrity being bound up with problematic notions of objectivity.

As will be discussed in ensuing chapters, it is possible to achieve a document that satisfies bureaucratic demand, maintains accountability, and is based on a comprehensive, rigorous evaluation process. However, the mechanisms driving evaluation are deeply entrenched within the system. Simplistic calls for donors or others to just be more participatory are unlikely to succeed; awareness of these contexts is a critical foundation to developing more effective practices. To this end, I have highlighted some of the emerging ways that media assistance organizations resist the bureaucratic structures. In particular, BBC Media Action's model is an example of the benefits of reclaiming the role of in-house evaluation. Internal evaluation, both in the headquarters and in-country, enables ongoing research and analysis, and provides opportunities to use innovative methodologies. There can still be quality assurance processes; in the BBC Media Action case, a consultant was commissioned as a scrutineer over the research quality, finding that far from spinning a success story, the research teams were harsh critics of their own work. This role, as well as roles in methodology coaching, guidance, and capacity-building, would represent a much better use of the consultant's skills and the donor's resources, and, most importantly, would enable deeper and more useful evaluation processes. The remaining chapters in this book map out a better practice for media assistance evaluation, firmly situated within the bureaucratic contexts in which evaluations occur.

## REFERENCES

- ADG Quality Performance and Results Branch. (2013). AusAID monitoring and evaluation standards. <http://www.ausaid.gov.au/Publications/Documents/monitoring-evaluation-standards.pdf>. Accessed 13 Sept 2013.
- Aho, J.A. (1985). Rhetoric and the invention of double entry bookkeeping. *Rhetorica: A Journal of the History of Rhetoric*, 3(1), 21–43. doi:[10.1525/rh.1985.3.1.21](https://doi.org/10.1525/rh.1985.3.1.21).

- Alcorn, J., Chen, A., Gardner, E., & Matsumoto, H. (2011). Mapping donor decision making on media development. An overview of current monitoring and evaluation practice. <http://www.mediamapresource.org/wp-content/uploads/2011/04/DonorDecisionmaking.MediaMap.pdf>. Accessed 22 Jan 2013.
- Anderson, M. B., Brown, D., & Jean, I. (2012). *Time to listen: Hearing people on the receiving end of international aid*. Cambridge, MA: CDA Collaborative Learning Projects. [http://www.cdainc.com/cdawww/pdf/book/time\\_to\\_listen\\_pdf\\_Pdf1.pdf](http://www.cdainc.com/cdawww/pdf/book/time_to_listen_pdf_Pdf1.pdf).
- Brenneis, D. (2006). Reforming promise. In A. Riles (Ed.), *Documents: Artifacts of modern knowledge* (pp. 41–70). Michigan: University of Michigan Press.
- Carothers, T. (1999). *Aiding democracy abroad: The learning curve*. Washington, DC: Carnegie Endowment for International Peace.
- Chouinard, J. A. (2013). The case for participatory evaluation in an era of accountability. *American Journal of Evaluation*, 34(2), 237–253. doi:10.1177/1098214013478142.
- Escobar, A. (1995). *Encountering development: The making and unmaking of the third world* (2012 ed.). Princeton, NJ: Princeton University Press.
- Fetterman, D. M. (2001). Empowerment evaluation: The pursuit of quality. *Advances in Program Evaluation*, 7, 73–106. doi:10.1016/S1474-7863(01)80067-3.
- Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2004). *Program evaluation: Alternative approaches and practical guidelines* (3rd ed.). Boston: Allyn and Bacon.
- Hanssen, C. E., Lawrenz, F., & Dunet, D. O. (2008). Concurrent meta-evaluation: A critique. *American Journal of Evaluation*, 29(4), 572–582. doi:10.1177/1098214008320462.
- Hayes, N., & Westrup, C. (2014). Consultants as intermediaries and mediators in the construction of information and communication technologies for development. *Information Technologies and International Development*, 10(2), 19–32. <http://www.itidjournal.org/index.php/itid/article/view/1214>.
- Jassey, K. (2004). The Bureaucrat. In R. Hinton & Groves, L. (Eds.) *Inclusive Aid: Changing power and relationships in international development* (pp. 128–134). London and Sterling, VA: Earthscan.
- Lennie, J., & Tacchi, J. (2013). *Evaluating communication for development: A framework for social change*. Oxford: Earthscan, Routledge.
- Noske-Turner, J. (2015). 10 years of evaluation practice in media assistance: Who, when, why and how? *Nordicom Review*, 36, 41–56. [http://www.nordicom.gu.se/sites/default/files/kapitel-pdf/nordicom\\_review\\_36\\_2015\\_special\\_issue\\_pp\\_41-56.pdf](http://www.nordicom.gu.se/sites/default/files/kapitel-pdf/nordicom_review_36_2015_special_issue_pp_41-56.pdf).
- Patton, M. Q. (2008). *Utilization-focused evaluation*. Thousand Oaks: Sage.
- Power, M. (1996). Making things auditable. *Accounting, Organizations and Society*, 21(2–3), 289–315. doi:10.1016/0361-3682(95)00004-6.

- Riles, A. (2006). Introduction: In Response. In A. Riles (Ed.), *Documents: Artifacts of modern knowledge* (pp. 1–38). Michigan: University of Michigan Press.
- Smith, D. E. (1974). The social construction of documentary reality. *Sociological inquiry*, 44(4), 257–268. doi:[10.1111/j.1475-682X.1974.tb01159.x](https://doi.org/10.1111/j.1475-682X.1974.tb01159.x).
- United Nations Evaluation Group. (2016). *Norms and standards for evaluation*. <http://www.uneval.org/document/download/2601>. Accessed 13 Mar 2017.

Rethinking Media Development through Evaluation  
Beyond Freedom

Noske-Turner, J.

2017, XIII, 176 p. 5 illus., Hardcover

ISBN: 978-3-319-58567-3