Ethics, Values, and Cultures

Rooting Evaluation Guidelines in Relational Ethics: Lessons From Africa

Caitlin Blaser Mapitsa[1] and Tara Polzer Ngwato[1]

Abstract

As global discussions of evaluation standards become more contextually nuanced, culturally responsive conceptions of ethics have not been sufficiently discussed. In academic social research, ethical clearance processes have been designed to protect vulnerable people from harm related to participation in a research project. This article expands the ambit of ethical protection thinking and proposes a relational ethics approach for evaluation practitioners. This centers an analysis of power relations among and within all the different stakeholder groups in order to establish, in a context-specific manner, which stakeholders are vulnerable and in need of protection. The approach also contextualizes the nature of "the public good," as part of an ethical consideration of interest trade-offs during evaluations. The discussion is informed by our experiences in African contexts and speaks to the "Made in Africa" research agenda but is also relevant to other global contexts where alternatives to "developed country" ontological assumptions about the roles of researchers and participations and the nature of vulnerability are being reconsidered.

Keywords

evaluation, ethics, evaluation guidelines, Africa, relational ethics

Introduction

Global debates on professional standards in evaluation practice are becoming more aware and inclusive of contextual assumptions and conditions. This is illustrated by the fact that regional evaluation guidelines are being codified and revised in ways that go beyond cosmetic adaptations of developed-country guidelines and rather start from local conceptions of core evaluation-related concepts such as the nature of the person, accountability, value, relevance, and impact, among others (Deane & Harre, 2016). One of the regions with an active debate on "indigenizing" or contextualizing evaluation theory, methods, and personnel is Africa, where an expanding and professionalizing evaluation sector is looking to define itself beyond its roots in the global aid

[1] Centre for Learning on Evaluation and Results, Wits University, Johannesburg, South Africa

Corresponding Author:
Caitlin Blaser Mapitsa, Centre for Learning on Evaluation and Results, Wits University, PO Box 85166, Emmarentia, Johannesburg 2195, South Africa.
Email: caitlinmapitsa@gmail.com

industry and the still-dominant unequal power dynamics of international donor/local beneficiary relationships (Chilisa & Malunga, 2012). This article argues that current debates around power, decolonization, and indigenization in Africa, and elsewhere, are not yet sufficiently clear on the contextual nature of ethics as a core aspect of evaluation practice. Furthermore, we argue that regional evaluation association guidelines are a relevant and important place to codify contextually nuanced conceptions of ethics.

The African Evaluation Guidelines (AEG) have sought to move beyond "developed country" assumptions about methods, program design, and development outcomes and are, at time of writing, being further revised with these issues in mind. As part of these discussions, we argue that the guidelines should provide evaluators with the questions and tools to identify and manage the complex ethical dilemmas often generated by the specific contextual realities of evaluations in Africa. We believe that similar dilemmas arise in many other evaluation contexts and that African experiences may contribute to a global debate.

We propose an approach to ethics in evaluation, which centers on an analysis of the specific, contextual power relationships between and among all the stakeholder groups in an evaluation. For this, we build on an approach to ethics known as "relational ethics," which "is a contemporary approach to ethics that situates ethical action explicitly in relationship. If ethics is about how we should live, then [relational ethics] is essentially about how we should live *together*" (Austin, 2007, p. 378). This approach applies to evaluations in all thematic disciplines and sectors as any technical program exists within a context of power and contestation.

We argue that there are three main implications of this approach. The first relates to who and what knowledge is valued. We need evaluators who understand the local context and local stakeholder relationships. This kind of knowledge is at least as important as technical knowledge about evaluations theory or technical knowledge about the intervention sector being evaluated. This means that ethical decision-making in evaluations, and effective evaluative judgment, is a function of contextual knowledge. Donors and other evaluation commissioners, in contrast, often prioritize "neutrality" (i.e., distance from the relationships in the evaluation context) and technical competency in their commissioning of evaluators, which results in individuals without context knowledge being appointed (Ngwabi & Wildeschut, in press).

The second implication is that evaluators need to be trained in and have tools for recognizing and analyzing multiple layers of power and complex webs of relationships between stakeholders. Such analysis is a mainstay of critical social science theory (Freeman & Vasconcelos, 2010, pp. 7–19) but is rarely included in evaluation courses. We argue that such training and tools are at least as important as the conventional evaluation tools for analyzing intervention logic, such as log frames, theories of change, and outcome mapping, since without them, the ethical implications of an intervention design, implementation, or outcome cannot be assessed, and there is therefore no normative compass for judging "success" or "failure" (and for whom).

Finally, an approach which takes power dynamics seriously implies that ethical decision-making in evaluations must be about identifying interest trade-offs and dealing with competing and conflicting interests. In contrast, the expectation expressed in existing guidelines is that it is possible to keep all stakeholder groups equally happy and that the only ethical principle is to "protect the vulnerable." Evaluators need to be equipped with skills and tools to negotiate often tense relationships for participation, ownership, and use. Our case studies echo many other evaluation studies in showing that it is not always easy to identify which interests serve the "public good" and whose vulnerabilities need protecting. For example, we discuss contexts where individuals or groups who have abused power within the context of the intervention being evaluated may also be the most vulnerable in terms of their station relative to the power dynamics of the broader society.

The implications of these three points mean different ways of working for professional evaluators, the evaluation capacity development sector, and commissioning agencies. It forces us to

consider who is hired as an evaluator, which skills sets are valued and built, and the contextual, thematic, and disciplinary background of the individuals on the evaluation team. It also impacts on which tools are developed and applied and how guidelines are formulated. Finally, all of these things inform both how engagement takes place with ethical issues in the evaluation process and how power is conceptualized and discussed in evaluation results reports (Abma, 2006).

This article substantiates and broadens these arguments by counterposing existing guidelines on professional and ethical evaluation practice in Africa (drawing mostly on the AEG adopted by the African Evaluation Association [AfrEA] in 2002 and updated in 2004 and 2006/2007) with examples from evaluations conducted or closely accompanied by the authors in the past 5 years. While guidelines are by definition general, and cannot and should not prescribe specific action, we argue that more conscious and focused research, commentary, and region-wide debate on the process of applying the guidelines in a variety of situations are necessary to build up a better understanding of their relevance and appropriateness. This process is necessarily a work in progress, as are the guidelines themselves.

We begin by defining our approach to ethics in evaluations and by summarizing current debates on ethical evaluation practice in the African context. We then discuss three important considerations in evaluation practice that pose interrelated but distinct sets of ethical challenges: positionality, participation, and relationships. Within each of these considerations, we summarize how existing guidelines advise evaluators to act and contrast this with examples of real ethical dilemmas we have experienced as evaluators. We conclude with reflections on implications for guidelines, evaluation practice, and voluntary organizations of professional evaluators (VOPEs).

Setting the Frame

Evaluation Systems in Africa

With the Sustainable Development Goals mandating country-led evaluation systems, many countries in Africa are taking steps to build or strengthen the monitoring and evaluation function within the public service. Many scholars have highlighted the need for transformation in the sector (Basheka & Byamugisha, 2015; Cloete, 2014), which is still strongly linked to the international development sector, and its postcolonial power dynamics. Professional associations in Africa's evaluation sector often have nascent institutional capacity, highlighting the challenges in organizing and supporting local evaluation practice (Mouton, 2010).

This development has led to dynamic discussions in the evaluation sector around the availability of local skills and capacity, the donor-driven nature of evaluation practice, and what this produces in terms of the relationships between the commissioning agency, the program implementers, the evaluation team, and the beneficiaries (Cloete et al, 2014). Due to the region's colonial history, the varied landscape of tertiary education for evaluators, public sector capacity, access to information, and political incentives have all shaped the region's political economy in ways that fundamentally impact on the way that evaluation needs to be understood and practiced (Cloete, 2016).

While many of the ethical dilemmas mentioned in this article are therefore common to evaluations everywhere, and many are common to developing contexts around the world, there are aspects of the relational power dynamics which are exacerbated in African contexts.

Relational Ethics

Our main argument, worked through in the rest of the article, is that ethical practice requires recentering the negotiated relationships between the "parties interested," based on an understanding of the power relations between them, within the specific context of the program being evaluated. This means considering the commissioning agencies and the evaluators as interested parties with

their own power dynamics, contextual roles, interests, and vulnerabilities rather than considering them to be neutral arbiters by virtue of their claims to a professional role or position.

In taking a relational approach to ethics, we build on critical theory which recognizes that the very idea that ethics are power neutral is not power neutral. The concept of ethics as an abstract, universal set of norms—"an autonomous authority over and beyond the parties interested"—is itself an idea that "hegemonic sectors of Western culture have fashioned of themselves for themselves" (Pels, 1999, p. 10) and that has been used to both impose and obscure specific interests, especially "professional" interests, in the colonial and postcolonial contexts. We extend this view by challenging the idea that an evaluator ought to be "neutral" or that "neutrality" is ever achievable, in identity or worldview. Just as "insider" evaluators may hold biases based on their personal experiences and understandings of the evaluation context, "outsider" evaluators also bring with them assumptions and biases which shape their evaluative judgment.

In the evaluation sector, the claim of "professional ethics" as an "abstract universal set of norms" finds expression when implementing agencies and beneficiaries have no say in how commissioning agencies and evaluators define the parameters of ethical practice. This is because commissioning agencies and evaluators are too often viewed as the benchmark for neutrality, organized in an established hierarchy of power, with donors at the top and beneficiaries at the bottom. Where evaluators hail from countries, education systems, and/or racial identity groups associated with the commissioning agencies, they are often treated as "spokespersons" for commissioning agencies rather than as independent mediators between all stakeholders (Johnson, 2015), compounding the reification of ethical immutability. One of the effects is to limit the space for implementing agencies and beneficiaries to engage honestly with the evaluator about the nature of their relationship with the commissioning agency, and how this relationship is impacting on the outcomes and impacts of the intervention (Bornstein, 2006). It also implies that the value of the evaluation itself, its process, timing, findings dissemination strategy, and so on, by de facto determined by the commissioning agency. We argue that a relational ethics lens would allow the evaluator to engage with the full range of power dynamics in the evaluation context, including the positionality of the commissioning agency itself, which is too rarely questioned.

Within this consciousness of context and power, we still define an ethical position as a judgment through which evaluation stakeholders make decisions in the interest of the public good while protecting the interests of vulnerable groups. However, sometimes there is a tension between varied interests, normative social good, and the protection of the vulnerable, and it may not always be easy to determine who is vulnerable. While guidelines cannot prescribe specific actions in different circumstances, they can provide some parameters and robust discussion around the boundaries and potential trade-offs defining an ethical position. Just as some guidelines already mention the need for transparency about an evaluator's theoretical and ideological position, we argue that there should also be transparency about the evaluator's analysis of power relations within and among evaluation stakeholders. There is no expectation that there is one objective interpretation of power relations, so transparency allows the evaluator's particular interpretation to be either contested or factored into the interpretation of the overall evaluation findings.

Guidelines

When compared with traditional biomedical ethics clearance procedures and the ethical criteria of "good clinical practice," the global evaluation community has already made advances in the discussion of ethics by going beyond individual consent and "doing no harm," to considerations for commissioning evaluation research, reporting on results, and the needs and rights of final evaluation users. These advances are reflected in a range of national and regional guidelines for evaluators that

set out ethical practice principles. Among the most commonly cited globally are the guidelines of the Australian Evaluation Association and the American Evaluation Association.

In parallel to the development of these guidelines within the (largely northern) "core" of the global evaluation sector, there has been a trend toward formulating regional guidelines, in line with the formation of regional associations as the profession expands. This is particularly important in Africa, given that the expansion of the evaluation profession is both decades newer than other regions and that the donor-driven expansion is at least partly due to political priorities outside of Africa.

The AEG, adopted by the AfrEA in 2002 were an important step forward both in providing African leadership to global evaluation debates on the evaluation profession and in providing guidelines that are contextually relevant. These guidelines specifically aim to "present a framework for conducting evaluations that are sensitive to the African context" while being "in line with international practice, standards and norms for evaluations" (AfrEA, 2002). The revision to the guidelines in 2006/2007 and the additional revision currently underway (as of early 2019) take these same aims forward.

The coverage of ethics in the 2002/2007 AfrEA guidelines mentions political feasibility and the need for contextual understanding. This is positive, but we argue that the AEG do not go far enough in fundamentally centering, rather than peripherally accommodating, an understanding of local value systems and power dynamics. We believe that guidelines that do this could be one step in allowing the evaluation profession in the region to develop in a locally appropriate way. Since 2002, very little has been published about these guidelines, with the exception of Patel (2013), who formally presented updated guidelines in the *African Evaluation Journal* as they came out, with a description of the revisions that were made. We are therefore trying to bring substantive practitioner research into the current revision of these guidelines and specifically draw attention to research on what "Made in Africa" curriculum and process design means in practice.

In the rest of this article, we illustrate how a relational ethics approach to evaluations can change how evaluators think about positionality and relationships. We consider how this approach can be entrenched in evaluation guidelines, and we reflect on several real-world case studies that present some of the ethical and relational dilemmas faced by all stakeholders in an evaluation process.

Positionality

Every evaluation by its very nature includes a wide range of stakeholders (Bryson, Patton, and Bowman, 2011). One of the strengths of evaluations is their triangulation of different points of view. In giving ethical direction to evaluators, it is important not to essentialize the perspective of any of these groups. Ethical guidelines need to be part of an array of tools that allow an evaluator to make nuanced assessments of the power dynamics of all stakeholders, including herself. There should be a recognition that sometimes interests will be competing and that—going back to our initial definition of an ethical standpoint—public goods and vulnerable populations will not always be apparent (Lewis and Naidoo, 2004). Furthermore, no stakeholder in an evaluation is either all-knowing or immune to the vulnerabilities and risks faced by other evaluation participants.

Current guidelines around ethical engagement with stakeholders focus heavily on inclusion, that is, on ensuring that a wide range of stakeholder views are reflected in the evaluation process and results. While this is important, what is missing is any guidance on the relationships between the stakeholders who are included and therefore how inclusion of one stakeholder may impact on the inclusion of another.

Some examples of unproblemmatized guidelines on stakeholder inclusion include the following excerpts from the AEG (2002, emphasis added):

- Principle of Utility 2, Stakeholder identification: Persons and organizations involved in or affected by the evaluation (*with special attention to community participants and vulnerable groups*) should be identified and included in the evaluation process in a participatory manner, so that their needs can be addressed and so that the evaluation findings are utilizable and owned by stakeholders, <u>to the extent this is useful, feasible, and allowed</u>.
- Principle of Utility 5, Information scope and selection: Data and information collected should be broadly selected to address pertinent questions and be responsive to the needs and interests of stakeholders, *with special attention to vulnerable groups*.
- Principle of Utility 9, Report dissemination: Significant interim findings and evaluation reports should be disseminated to stakeholders, <u>to the extent that this is useful, feasible, and allowed</u>. Comments and feedback of stakeholders on interim findings should be taken into consideration prior to the production of the final report.
- Principle of respect and ethics (General): These principles safeguard the respect of legal and ethical rules as well as the well-being of stakeholders involved in the evaluation or affected by its findings.
- Principle of respect and Ethics 6, Disclosure of conclusions: The findings of the evaluation should be owned by stakeholders and the limits of the methodologies in use should be precise. Recommendations resulting from this ownership will be designed with stakeholders. <u>Privacy should be maintained during the whole process to avoid any attempt to intimidate executing agencies or evaluators</u>.

We see two related problems with these guidelines. The first is that they come with assumptions about what stakeholder groups look like, what power they hold, and how they behave, which may not be accurate. The passive voice of the guidelines leaves the actor unstated, which is itself a grammatical elision of self-awareness for the actors, assumed to be the evaluator, as well as the commissioning agency. These two sets of actors are instructed to protect the ethical rights of other stakeholder groups, focusing on "community members" and beneficiaries as "vulnerable," without consideration of transparently addressing their own internal ethical dilemmas.

The second challenge is that these guidelines are difficult to operationalize by an evaluator in a context where there are conflicting interests, and where trade-offs between these interests need to be thought through. The caveats written into the guidelines (underlined above) suggest an awareness of these challenges, but the guidelines do not provide the evaluator with tools to assess these power relations and competing interests.

To illustrate our points, we provide several examples of cases where commissioning agencies' and evaluators' positionalities have not been "neutral" or ethically unproblematic. Our brief case studies show how transparency about their ethical choices would have been beneficial to the process and outcomes of the evaluation but also how sometimes conflicts of interests are not easily resolvable even when there is transparency.

Positionality of Commissioners

Commissioners of evaluations play two important roles by nature of their position. First, they have a contractual relationship with evaluators, which makes the evaluation possible but also creates an unequal power relation. Second, they play an important role in the use of the evaluation findings, which means both they, the evaluators, and the beneficiaries need to negotiate interests in program design following the evaluation results. The area of ownership and use is where the evaluator is most commonly put in a dilemma of mediating conflicting interests between the commissioners and beneficiaries, and where commissioning organizations also have obvious

ethical considerations. The evaluator has a dual imperative to serve the needs of the commissioning organization (based on the contractual relationship and dependent power relation), and still serve the "public interest," or the needs of beneficiaries.

Case Study 1

An evaluation was commissioned with the goal of shutting down an organization which the commissioning agency considered ineffective. This only became apparent to the evaluation team midway through the evaluation, although it was known to the implementing agencies from the beginning. As a result, the implementation agency managers and staff were hostile to the evaluation team. By the time the evaluation team understood the context, there was little space to negotiate the scope of the evaluation with the commissioning organization. Furthermore, the implementing agency had other programs, not funded by the commissioning agency, the beneficiaries of which were negatively affected by the outcomes of the evaluation.

Case Study 2

The commissioning agency of an evaluation was legally obliged to act on reports of criminal wrongdoing, which were exposed in the evaluation report. However, the commissioners were unwilling to engage with the costs and reputational complexities of the legal processes that recognizing the report would have required. As a result, the evaluation report was buried by the commissioning organizing, and never acted on.

Case Study 3

In several evaluations, there was an attempt by the evaluation team to suggest research pieces on the basis of evaluation results that are of significant academic merit and of interest to the general public. However, due to contracting around intellectual property, the commissioning agencies did not permit the learnings to be shared and used more widely.

The situations above demonstrate the well-known fact that commissioners have a powerful position in any evaluation, which they can use in ways that have ethical implications for the evaluators, implementing agencies, and beneficiaries. Evaluation guidelines state that evaluators should ensure the inclusion and consideration of the interests of implementing agencies and beneficiaries, as well as the general public in sharing the results of evaluations. However, there are few mechanisms through which these ethical provisions can be enforced in practice when powerful commissioners see their own interests in conflict with these ethical provisions.

Taking a contextually considered relational ethics approach to this power dynamic, we identify national or regional VOPEs as potentially important mechanisms to enforce the guidelines which demand ethical behavior and information sharing from commissioning agencies, just as they are supposed to oversee evaluator behavior. Taking this role would require VOPEs in Africa to move beyond many of their histories as government-housed or -initiated agencies (so not independent of governmental commissioners) or loose associations of individual evaluators who compete against each other for commissioner contracts. According to the same principle of organization and collective action that has allowed unequal power relations to be reduced in other employment sectors, information sharing among evaluators, in an organized manner, about commissioner ethics, and feedback to commissions by an association rather than a vulnerable individual evaluator, would make ethical provisions in guidelines more meaningful.

Another way in which commissioning agencies have ethically problematic impacts on evaluations is by dictating the composition of the evaluation team. Team decisions can be intentional or determined by administrative contingencies, as in Case Studies 4–6.

In essence, in the examples above, the quality of ethical insight and contextual understanding of the evaluations were compromised in order to meet the administrative ease and compliance needs of the commissioning organizations. If the evaluations were judged simply on their usability by the commissioning agencies, to fulfill their narrowly defined aims (i.e., ease of cultural/linguistic fit with commissioning agency culture or compliance), they could be considered successful. However, if the goal is around the ethical imperative of improving program design in order to generate a

Case Study 4

An international donor commissioned an evaluation of a programme it had funded in an African country. While the evaluator selection process was supposedly open and competitive, the requirements for the team lead included speaking the commissioning agency language and having a formal qualification in technical evaluation processes, with no stated requirements for local context experience.

Case Study 5

An evaluation was commissioned by an organization that was required to conduct it for compliance reasons. There were tight time lines, and existing procurement policies meant that it was impossible to make revisions in the scope of the project. To solve this problem, a consortium of evaluation organizations with standing contracts was cobbled together. While the compliance requirement was met, none of the evaluators could speak authoritatively to the implementation or outcomes of the work.

Case Study 6

A government department mandated to oversee monitoring and evaluation for all government-funded and implemented programmes consistently complained of a lack of evaluation capacity and diversity in the country for them to commission. They then commissioned a study on capacity in the sector which found that there were in fact many more qualified evaluators, including indigenous and local language–speaking evaluators, but that government procurement processes excluded all but the largest and most established evaluation companies, almost of all whom were foreign-based or dominated by evaluators with non-African heritage.

positive impact for past and future beneficiaries (i.e., "public interest"), the evaluations are less likely to fulfill any such claims of success.

Positionality of Evaluators

A consideration of positionality in relation to the evaluator requires a consideration of the power-arbitrating role of the evaluator. If we conceive of a participatory process not as devoid of power and politics, but as bringing all the power plays of an intervention into the evaluation process (Cooke, 2001), then the role of the evaluator is not one of a neutral bystander and documenter, but rather of an active arbitrator between stakeholder interests and perspectives. Such a role brings with it certain implications. The evaluator does not have an a priori neutrality but must position herself consciously within a certain power dynamic. Firstly, it requires specific skills of the evaluator. The AfrEA Guidelines usefully state that the credibility of an evaluator depends on her "cultural sensitivity [and] appropriate communication skills" (Utility Principle 3: Credibility of the Evaluator and Utility Principle 4: Credibility of the Evaluation Team). We would specify that "cultural sensitivity" needs to include self-knowledge (understanding of one's own cultural assumptions and their contextuality). We would further add the additional skill sets of identifying power relations and mediation skills.

Many emerging evaluators, or even established professionals working in different contexts or thematic areas, may not be equipped with the skills of considering the varying sorts of power they hold, power the participants hold, and power the evaluation commissioner holds. As discussed further below in the section on relationships and in the conclusions on training, there is extensive existing literature within the social sciences and development studies on how to analyze power. In addition to identifying power relations, evaluators need to be able to mediate when there are competing interests and manage the relationships around such conflicts. These power analysis and mediation skills, in our view, should be central components of Monitoring and Evaluation (M&E) training and recognized as core skills in evaluation guidelines.

The second implication is that it is necessary to recognize that the evaluator comes to this arbitration role with an existing worldview and set of values. Again, the existing guidelines are useful in stating that "the rationale, perspectives and methodology used [by the evaluator] to

interpret the findings should be carefully described so that the bases for value judgments are clear" (Utility Principle 6: Values Identification), although this is formulated in relation to the evaluation findings and not in terms of the process of mediating competing and conflicting participant interests per se.

Finally, taking on an arbitration role is complicated by the fact that the evaluator is likely to share a certain set of vulnerabilities, in addition to power positions and values that may be the same or may be different to the various vulnerabilities of other evaluation stakeholders. The evaluator may either share or not share a whole range of characteristics with the evaluation participants, including gender, racial identity, educational background, ethnicity, religion, language, political or ideological affiliation, and so on. These overlapping characteristics will shape not only the material gathered by the evaluator (Bledsoe & Donaldson, 2015) but also the risks that could be faced by the evaluator and the participants in engaging with the evaluation program. Understanding and unpacking these are important for appropriate analysis of the data as well as an appropriate consideration of the nature and subjects of the ethical risks in the evaluation process.

The above considerations regarding the role of an evaluator in mediating competing interests in participative processes have implications for how we see the roles of insider and outsider evaluators. As noted in the Introduction and Context sections, many of the evaluators active in Africa are "outsiders," often not of African origin, not living full time in Africa or not connected in their everyday lives with the contexts they are evaluating. A recent survey conducted by the Centre for Learning on Evaluation and Results (CLEAR) and the University of Stellenbosch (unpublished) has highlighted the complexity in identifying cultural knowledge on the basis of various characteristics, with many evaluators, including those of African heritage, having cosmopolitan identities, countries of birth different from their country of residence, and multiple nationalities.

One of the narratives of a conventional approach to ethics is that the evaluator should hold a position of professional neutrality (assumed to be associated with outsiders) rather than being "biased" by being involved in the local context (assumed to be associated with insiders). This assumes that all evaluators start with the same basic information about the ethics of a situation and then differentiates insiders and outsiders based on the quality (or "neutrality") of their judgment about that information. If, however, we recognize this narrative as a component of a colonizing normative supremacy, and that in fact all evaluators bring individual and group biases and vulnerabilities with them into any stakeholder engagement, we must question both the assumption of information parity and the assumption of what "good" ethical judgment entails.

On the question of information, "insider" evaluators may be more likely (although of course not homogeneously or unbiasedly capable) to recognize and understand the subtle contextualized power relations between participants, while "outsider" evaluators are more likely to miss important dynamics entirely and be more easily misled by context-savvy participants. On the question of ethical judgment, if this requires identifying multiple and crosscutting forms of "public interest" and "vulnerability," "insiders" may be more equipped (if they are sufficiently self-reflective about their own biases and not just as distanced from local relationships based on urban, class, or educational backgrounds), while "outsiders" may unwittingly place some participants in danger because of not understanding the nature of relationships.

While this section has spoken largely to the individual professional evaluator, the same lessons are also of relevance to other evaluation stakeholders. As noted above, commissioners of evaluations should be aware of the importance of identity and contextual knowledge in the evaluators that they hire. Evaluation capacity development organizations and programs should be cognizant of the importance of contextual knowledge and offer training that prioritizes these skills.

Relationships

Building on our arguments relating to positionality, a relational ethics approach requires a rethinking of various other aspects of the relationships in an evaluation system overall and among the stakeholders in any particular evaluation. In this section, we discuss how the concepts of participation, conflict of interest, and vulnerability should be approached more contextually in discussions of ethics in evaluations and in the AEG.

Participation

In addition to ethics guidelines about the *range* of stakeholders who should be considered in an evaluation, and concomitant assumptions about the *power relationships* between those stakeholders (as discussed in the previous section), ethical guidelines also set out *the manner in which evaluators should engage* with stakeholders. The buzzword is participation.

Once again, we argue that by insufficiently considering the power dynamics, diversity, inequalities, and other contextual characteristics of different stakeholders in African contexts, the existing ethical guidelines for the inclusion and participation of stakeholders not only fail to prevent or resolve ethical dilemmas but can generate them (Cooke, 2001). While guidelines cannot prescribe specific actions in every context, we argue the existing AfrEA Guidelines should go further to help evaluation stakeholders understand the process of ethical adjudication between the competing priorities of participation, protecting the vulnerable and producing a high quality product.

The issue is not whether to encourage participation or not in different circumstances but rather how to be conscious of the process of participation and its implications. As a starting point, some practical steps for evaluators could include asking stakeholders about any concerns they may have with participation, giving them options about the form their participation may take, and writing up a reflection on participation process and results as part of the methods of an evaluation. We recognize that there are trade-offs evaluators may face between following good practice in ethics and other logistical considerations around cost and time lines, but these logistical considerations have been more extensively explored in evaluation literature, and evaluators may be better equipped to deal with these trade-offs (Bamberger, Rugh, & Mabry, 2011).

There is ample scholarship available in the fields of development and social science research methods around participatory research and the many considerations in both process and results that come with it. Particularly in the contexts of the aid industry, development programing, and humanitarian aid, the critical literature on participation is rich and accessible (Molenaers & Renard, 2009).

However, due to the rapid evolution of evaluation as a profession, not all evaluators come from an academic or professional background which covers these tenets of research and development work/humanitarian ethics. Research on participation and appropriate methods from various technical sectors can now be found in the development studies literature (governance, water, and sanitation). As the professionalization of evaluation is driving curriculum design and training practice, this material needs to be included in evaluation qualifications and professional development courses.

Among several provisions in the AEG regarding participation, those included under the theme of feasibility most strongly place a participative approach at the center of any evaluation. Feasibility itself is defined in relation to participation—"The feasibility principle is designed to ensure that evaluation is useful, participative, realistic and efficient" (AfrEA Guidelines, p. 5)—suggesting that evaluations which are not "participative" are not feasible, meaning they cannot be done.

Feasibility Principle 2: Political viability most clearly illustrates the challenge of prioritizing the principle of participation per se over a consideration of the nature and power positions of those participating. The guideline states that

> Evaluation should be planned and conducted in a participative manner in order to achieve *total* involvement of *all* stakeholders. It should be prepared and conducted on the basis of *scientific principles of neutrality* and strictness to *avoid disputes* conducive to a negative impact on processes and findings as well as on implementation and recommendations. (emphasis added)

The *Utility Principle 2: Stakeholder Identification* which is quoted above also explicitly mentions participation. We have already discussed why we believe that a relational ethics approach changes the focus from *neutrality* to a focus on self-aware and transparent positionality and shifts us from *avoiding disputes* to an intentional and transparent identification of conflicting interests and their active mediation.

Conflict of Interest

Similarly, "conflict of interest" is commonly considered a basic professional principle for ethical best practice. Rather than engaging in an analysis of actual, contextually defined interests, however, "conflict of interest" is generally defined in terms of the presence of a preexisting relationship between the evaluator and other stakeholders in the evaluation—especially the implementing agency. The supposedly professional opposite of "conflict of interest" is therefore a "neutral" outsider status. The problem with this position has been discussed above.

In addition, in countries with small populations, which may only have one national university and a strongly interconnected educated population, it may be difficult to develop "conflict of interest" policies that are useful and implementable if "conflict of interest" is defined through preexisting relationships. In certain professional fields in Africa, with monitoring and evaluation being no exception, it is likely that all practicing professionals have multiple relationships with the limited number of organizations in the field. The case study below demonstrates one of the challenges of trying to avoid the pitfalls of these overlapping relationships.

Case Study 7

A commissioning organization had an elaborate conflict of interest policy as a means of safeguarding the ethical relationship between the commissioners and the evaluators. This policy disallowed using an evaluator twice as part of the same country program. As a result, evaluation methods were chosen which were less effective, or evaluations were even abandoned because the scarcity of certain evaluation skills meant that the organization did not want to "burn" various evaluators' skills on any but the most critical projects.

Even when evaluators are not connected to commissioners or implementing agencies through preexisting family, friendship, or collegial ties, long-standing relationships within the sector are almost inevitable within small professional circles. The individual evaluator, evaluation team, or evaluation organization comes with a certain identity. The livelihood of the evaluator or evaluation organization hinges, to a certain degree, on their ongoing relationship with evaluation commissioners, although this depends on the specific professional location of the evaluator. This could simultaneously be an incentive to maintain high professional standards, or limit criticism of the program in question, depending on the context and role players (Thomas, 2010). Program implementers could have a range of relationships to the evaluation commissioner, the evaluation organization, and the program beneficiaries. Particularly in niche technical areas, or countries with a relatively small educated elite, all of these functions can be drawn from a very limited pool of people who wear different hats in different situations.

Case Study 8

An evaluation organization found evidence of wrongdoing by the implementing agency that was substantial enough that under normal circumstances, the evaluation would have ended, and the agency reported. However, this evaluating organization was also collaborating with the implementing agency on another project that delivered services to vulnerable communities, and they did not want to jeopardize this relationship because it would have harmed the beneficiaries of the second project.

Vulnerability

From a relational ethics perspective, vulnerability is not a function of an abstract positionality but of a specific set of power imbalances within a specific relationship. This means that commissioning agencies, evaluators, implementing agencies, and other supposedly powerful stakeholders can also be vulnerable in some ways and in some situation.

Case Study 9

In an evaluation, different levels of anonymity were granted to participants on the basis of presumed levels of vulnerability. The evaluation team assumed unemployed community members would be relatively more vulnerable than local government counselors. However, it turned out that counselors were caught in a complicated web of violent threats which impacted on their engagement with the evaluation. In fact, great resources were invested in protecting the anonymity of the wrong category of participant, while other stakeholders were unknowingly put at risk.

Case Study 10

An evaluation of a local government initiative involved interviews with a community association that was widely recognized for vigilantism. Its head office was in an isolated area in a forest and was heavily guarded. The head of the association would only meet at his office, in the evening, with a single member of the evaluation team, putting that evaluator in a very vulnerable position.

Case Study 10 calls into question the traditional assumption that researchers are always in a position of power in relation to the stakeholders being evaluated or the final beneficiaries. In evaluations, there are many perspectives by which the participant is more powerful than the evaluator. The participant, by definition, has contextual knowledge of the program and the other participants, which can be a form of power (Bledsoe & Donaldson, 2015). Participants are likely to know the other role players better than the evaluator and may have long histories and ongoing relationships with the other stakeholders. This contextual and relational knowledge is a form of power that the participants could use to compromise the safety and well-being of the evaluation team as well as the quality of the evaluation findings. Furthermore, the feedback and engagement of participants shape the findings of an evaluation in a way that could be significant for the career of an evaluator.

Implications and Discussion

Homogenizing

The emergence of democratic evaluation approaches (Patton, 2008; Stufflebeam & Coryn, 2014) has come with a hope that participants in evaluations may take part in some sort of constellation of activities around the use of evaluation, through codetermining the purpose of the program, by being part of the accountability processes in program implementation, and by allowing the evaluation to promote ownership in the program. Evaluation theory regularly recognizes that participatory approaches require significant investments in "timing, resources, and focus" and that they require strong facilitation skills (Guijt, 2014, p. 18). What is less often discussed is how participatory approaches may risk glossing over the differences in interest among stakeholder groups, and how agendas and power within a program and its evaluation may be contested (Maeckelbergh, 2016).

In the social sciences, and especially in critical studies of development projects, it has long been recognized that participatory approaches risk homogenizing "the community" and that they can reinforce some of the exclusionary power dynamics within communities (Chambers, 1997; Fowler, MacLeod, Mandiola, & INTRAC, 2016). Other stakeholder groups, including people within and associated with the implementing agency, can also be homogenized in the same manner. Such homogenization can directly threaten some members of a group, as Case Study 11 describes below.

Case Study 11

In one evaluation site, evaluators identified key stakeholders, with the intention of organizing focus groups with groups of common stakeholders. When experiencing extreme logistical challenges in organizing a focus group with local councilors, it eventually became apparent that due to infighting within the governing political party, councilors were being intimidated and one had been murdered, allegedly organized by another councilor. In this particular case, homogenizing a stakeholder group posed a problem for both process and content of the evaluation results.

The implication is not that participative approaches are unethical per se, but that "empowerment must be re-imagined as an open-end and ongoing process of engagement with political struggles" (Williams, 2004, p. 557) rather than being seen as a depoliticized process. One of the obligations of the evaluator is to disaggregate the interests within groups, understand the points of contention, and ensure that the process of participation allows for a balanced analysis of power dynamics that may exist among stakeholders. The existing guideline directives do not help evaluation practitioners navigate the nuanced and complicated process of deciding whose voices to accommodate and include, and indeed when there might be situations in which some voices should be excluded because they pose a danger to other voices. While a guideline may be unable to balance sensitivity to context, and sufficient broadness to be useful to a wide audience, it would be possible for guidelines to include processes for engaging with participation in a way that is inclusive, disaggregated, and analyzed in a way that is contextually appropriate.

Participation and Quality of Evaluation Outcome

An unreflective "quantitative" approach to a participative process that includes *all* stakeholder groups (i.e., *Feasibility Principle 2: Political viability*) may clash with the ethical principles relating to quality outputs of an evaluation. Case Studies 12 and 13below highlight that everyone has different motives for participating in the evaluation process and that the opportunities and risks brought by this may not be known to the evaluation team. These motives may be legitimately competing, both with the interests of other stakeholders and also with the eventual quality of the evaluation report. This can often play out in the availability and accessibility of data or the willingness of certain stakeholders to participate in the process.

Case Study 12

In an evaluation of a municipal project, the program manager of the implementing agency being evaluated lied to the evaluation team and said that a request had been sent for all managers to participate in the evaluation. In fact, no such request was sent, and to the contrary, rumors were circulated that the evaluation team was there to spy on the municipality. As a result, most managers pretended they were not available or refused to participate in the evaluation in other ways.

Case Study 13

Evaluation F was started with a data collection strategy approved by the commissioning agency, only to find out that a previous program manager had run off with all the program data, making the proposal impossible to operationalize.

Part of the question of power in an evaluation is the extent to which an evaluator has the power to compel stakeholders to participate or whether it is a matter of stakeholder choice. It is a central tenet of research ethics that participation in any research endeavor must be voluntary. However, stakeholders, especially people connected to implementing agencies, beneficiaries and people in any form of dependency relationship with either the commissioning agency or the implementing agency, may not feel that they are able to refuse to participate. This is especially the case if an evaluation is known or assumed to be the condition for continued funding. Even as individuals or groups may formally participate in an evaluation by attending meetings, granting interviews, or providing data, they may still deny the evaluator their full participation. This can take the form of selectively providing information, sending compliant or junior individuals to represent a group, or predisposing other stakeholders to distrust the evaluation team.

While most evaluation methods' discussions focus on how an evaluator can ensure completeness of information, the relational ethics approach we are proposing requires the evaluator to also think through the motivations different stakeholders have to participate in an evaluation, with a recognition that the evaluator may not have sufficient knowledge of the program, stakeholders, and relationships to assess these motivations in advance. The opposing side of this norm shift is that stakeholders may also have an interest in nonparticipation in the evaluation. The evaluator also has the responsibility to understand, document, and—significantly—respect and reflect motivations for nonparticipation, which may include participant self-preservation in the face of evaluation processes or outcomes that they think would make them vulnerable. The evaluator then has to balance the impacts of such nonparticipation against any prescriptions on scope or reach that may come from the evaluation commissioner, as well as impacts on the "completeness," "balance," and therefore quality of evaluation results.

In summary, the shifting norms in favor of participatory approaches to evaluation are a welcome trend. However, this unearths ethical issues that have not been sufficiently unpacked. For participation to play a positive role in evaluation guidelines norms and standards, it must come with a nuanced understanding of the interface between power and participation. An evaluator should be viewed not as the all-powerful researcher of the past, but as an individual, like those who are participating in the evaluation, who is both part of an enmeshed web of vulnerabilities, knowledge, and relationships, and trying to untangle this same web in the process.

Conclusion

It is evident that evaluations are embedded in a context of complex power relations that include, but are not always understood by, the evaluator. These power dynamics are embedded in both the program being evaluated and also the evaluation process itself. For an evaluator to be able to navigate this landscape ethically, it is important that guidelines and training acknowledge and reflect this context. It is also important that other evaluation stakeholders, and the institutions supporting the sector such as the VOPE, share a common analysis of how to engage with their respective roles, as well as these complex power dynamics. As it stands, these power dynamics are insufficiently understood, and the relational component of them often hidden to evaluation stakeholders.

As the evaluation sector becomes entrenched in Africa, and the debates and discussions on the merits and pitfalls of professionalization move forward, evaluators may have access to greater contextual knowledge to inform their work. This makes it critical both that the guidelines reflect this knowledge, but also that each stakeholder has sufficient guidelines around their diverse roles in the evaluation process. Evaluators, commissioners of evaluation, evaluation capacity development organizations, and program managers all need to engage with the guidelines for different reasons. Their ethical considerations are different, and while guidelines cannot prescribe actions for everyone in every situation, it can help provide parameters and processes for consideration from each

perspective. Furthermore, these guidelines cannot be documents that are isolated from institutional mechanisms that make them meaningful. They need to become a widely recognized and enforceable norm for a sufficiently wide range of actors. Well-organized and independent VOPEs are critical for ensuring power relations among different evaluation stakeholders can be mediated and that there can be some sort of recourse by any role player if they are not adhered to.

Contextual knowledge based on a process of relational power analysis must come to inform evaluation guidelines in Africa. At the moment, the guidelines do not go far enough to provide evaluators and other role players with the boundaries of a process that will help guide in the selection of a principle, method, or approach within which to make decisions in an evaluation. Additionally, there is no consensus about what falls in and outside of ethical evaluation practice, leaving all stakeholders without recourse in situations where ethical violations are taking place. Until these tools have been debated and developed, and discussed through their application in context, the evaluation sector will remain vulnerable to the very same power dynamics that are currently challenging program effectiveness.

Evaluation has the potential to strengthen development in the region, acknowledge and balance varied interests, and help program design to be effective in context. It has a unique role in driving transformative change in program design and implementation. However, for power dynamics to be transformed, a critical first step is to acknowledge the centrality of power-based relationships at all stages in the evaluation process. The sector will be able to live up to its potential most effectively if it embraces the contestation and power dynamics directly and supports all stakeholders in an evaluation process to first understand, and then act with integrity in a complex world.

Declaration of Conflicting Interests

Funding

References

Abma, T. A. (2006). The practice and politics of responsive evaluation. *American Journal of Evaluation*, *27*, 31–43.

African Evaluation Association. (2002). The African evaluation guidelines: 2002. Retrieved January, 11, 2005, from https://www.sciencedirect.com/science/article/abs/pii/S0149718902000587.

Austin, W. J. (2007). Relational ethics. In L. M. Given (Ed.) (2008). *The Sage encyclopedia of qualitative research methods.* Thousand Oaks, CA: Sage. doi:10.4135/9781412963909

Bamberger, M., Rugh, J., & Mabry, L. (2011). *Real world evaluation: Working under budget, time, data, and political constraints*. Thousand Oaks, CA: Sage.

Basheka, B. C., & Byamugisha, A. (2015). The state of monitoring and evaluation (M&E) as a discipline in Africa. *African Journal of Public Affairs*, *8*, 75–95.

Bledsoe, K., & Donaldson, S. I. (2015). Culturally responsive theory-driven evaluation. In Stafford Hood, Rodney Hopson, & Henry Frierson (Eds.), *Continuing the journey to reposition culture and cultural context in evaluation theory and practice* (pp. 3–27). Charlotte, North Carolina: Information Age Publishing.

Bornstein, L. (2006). Systems of accountability, webs of deceit? Monitoring and evaluation in South African NGOs. *Development*, *49*, 52–61.

Bryson, J. M., Patton, M. Q., & Bowman, R. A. (2011). Working with evaluation stakeholders: A rationale, step-wise approach and toolkit. *Evaluation and Program Planning*, *34*, 1–12.

Chambers, R. (1997). *Whose reality counts* (Vol. 25). London, England: Intermediate Technology.

Chilisa, B., & Malunga, C. (2012, November). Made in Africa evaluation: Uncovering African roots in evaluation theory and practice. In *Bellagio Centre. Report of the African thought leaders forum on evaluation for development: Expanding thought leadership in Africa* (pp. 14–17).

CLEAR-AA. (n.d.). A survey of evaluators in the Afred database. Unpublished

Cloete, F. (2016). Developing an Africa-rooted programme evaluation approach. *African Journal of Public Affairs*, *9*, 55–70.

Cloete, F., Rabie, B., & De Coning, C. (Eds.). (2014). *Evaluation management in South Africa and Africa*. Stellenbosch, South Africa: AFRICAN SUN MeDIA.

Cooke, B. (2001). *Participation: The new tyranny?* London, England: Zed Books.

Deane, K. L., & Harré, N. (2016). Developing a thoughtful approach to evaluation: Values-driven guidelines for novice evaluators. *Evaluation Matters–He Take Tō Te Aromatawai*, *2*, 53–78.

Fowler MacLeod Mandiola, INTRAC. (2016). How aid really works the problem with participation. Retrieved from https://www.intrac.org/aid-really-works-problem-participation/ [Special issue]

Freeman, M., & Vasconcelos, E. F. S. (2010). Critical social theory: Core tenets, inherent issues. *New directions for evaluation*, *127*, 7–19.

Guijt, I. (2014). *Participatory approaches: Methodological briefs—Impact evaluation no. 5* (No. inn pub 750). Retrieved from, https://ideas.repec.org/p/ucf/metbri/innpub750.html.

Johnson, C. M. (2015). *Power and participation: Relationships among evaluator identities, evaluation models, and stakeholder involvement* (Doctoral dissertation). Boston, MA: Boston College.

Lewis, S. G., & Naidoo, J. (2004). Whose theory of participation? School governance policy and practice in South Africa. *Current Issues in Comparative Education*, *6*, 100–112.

Maeckelbergh, M. (2016). Whose ethics? *Impulse to act: A new anthropology of resistance and social justice* (p. 211). Bloomington: Indiana University Press

Molenaers, N., & Renard, R. (2009). 11 the trouble with participation: assessing the new aid paradigm. *Doing good or doing better* (p. 255).

Mouton, C. (2010). *The history of programme evaluation in South Africa* (Doctoral dissertation). Stellenbosch, South Africa: University of Stellenbosch.

Ngwabi and Wildeschut. (in press). Scandinavian donors in Africa: A reflection on evaluation reporting standards. *Evaluation Landscape in Africa*. Stellenbosch, South Africa: SunMedia.

Patel, M. (2013). African evaluation guidelines. *African Evaluation Journal*, *1*, 5.

Patton, M. Q. (2008). *Utilization-focused evaluation*. Thousand Oaks, CA: Sage.

Pels, P. (1999). Professions of duplexity: A prehistory of ethical codes in anthropology 1. *Current Anthropology*, *40*, 101–136.

Stufflebeam, D. L., & Coryn, C. L. (2014). *Evaluation theory, models, and applications* (Vol. 50). Hoboken, NJ: John Wiley.

Thomas, V. (2010). Evaluation systems, ethics, and development evaluation. *American Journal of Evaluation*, *31*, 540–548.

Williams, G. (2004). Evaluating participatory development: Tyranny, power and (re)politicisation. *Third World Quarterly*, *25*, 557–578.