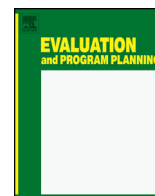




ELSEVIER

Contents lists available at ScienceDirect

# Evaluation and Program Planning

journal homepage: [www.elsevier.com/locate/evalprogplan](http://www.elsevier.com/locate/evalprogplan)

## Conclusion: Agency in the face of complexity and the future of assumption-aware evaluation practice

Nathan Morrow<sup>a</sup>, Apollo M. Nkwake<sup>b,\*</sup><sup>a</sup> Payson Program for Global Development, Tulane University, 5208 Magazine St, #188, New Orleans, LA 70115, United States<sup>b</sup> African Women in Agricultural Research and Development (AWARD), Hosted by World Agroforestry Centre, United Nations Avenue, Gigiri, P.O. Box 30677-00100, Nairobi, Kenya

## ARTICLE INFO

## Article history:

Received 12 May 2016

Accepted 23 May 2016

Available online 7 June 2016

## Keywords:

Future of assumption-aware evaluation

Complexity

Agency

## ABSTRACT

This final chapter in the volume pulls together common themes from the diverse set of articles by a group of eight authors in this issue, and presents some reflections on the next steps for improving the ways in which evaluators work with assumptions. Collectively, the authors provide a broad overview of existing and emerging approaches to the articulation and use of assumptions in evaluation theory and practice. The authors reiterate the rationale and key terminology as a common basis for working with assumption in program design and evaluation. They highlight some useful concepts and categorizations to promote more rigorous treatment of assumptions in evaluation. A three-tier framework for fostering agency for assumption-aware evaluation practice is proposed—agency for themselves (evaluators); agency for others (stakeholders); and agency for standards and principles.

© 2016 Elsevier Ltd. All rights reserved.

### 1. Introduction

The articles in this special issue outline the philosophical basis, mechanisms, and evidence from recent cases to illustrate how working with assumptions improves evaluations and program design particularly in complex environments. Nkwake and Morrow lay out the philosophical basis for working with assumptions and make an argument for the development of a typology of the assumptions that permeate our practice of Evaluation. Mertens in the second article links the practice of working with assumptions in evaluation to the transformative power of rooting evaluative practice in social justice and human rights. Chen makes an argument for better approaches to working with assumptions on a pragmatic basis – they improve both program theory and evaluation theory—leading to more effective programs and better evaluation designs.

Archibald et al. present evidence of how working with assumptions specifically enables adaptive management, adapting program theory to context and resulting in greater agency and empowerment of front line staff. The final article, by Morrow and Nkwake, attempts to put these concepts into some historical

perspective by underlining the changing role of the evaluator and development of assumption-aware tools as our profession seeks to improve evaluation and design for increasingly complex evaluands in ever dynamic and more complex contexts. The particular focus on the development of tools is intended to both provide insight to the development of assumption-aware practice but also provide readers with some options for integrating more assumption-aware design and evaluation into their own work.

#### 1.1. Diverse view points on assumptions found in this volume

Nkwake and Morrow open this volume arguing the case for a more systematic approach to working with assumptions in program evaluation and design. The practice of surfacing and examining assumptions has deep historical and philosophical roots. Although evaluators face a plethora of unexamined assumptions in their practice and work with stakeholders, the research and has taken place in a piece-meal fashion with the concept of evaluations being picked up by a variety of authors. Nkwake and Morrow believe that better approaches and tools for working with assumptions is fundamental to improving evaluation and program design in complex contexts and with complex evaluands for a variety of reasons including unexpected outcomes, ubiquitous feedbacks, fuzzy boundaries and linkages at a variety of scales. Articulating and testing assumptions is one way for making incremental progress and learn through reflection. Ultimately

\* Corresponding author.

E-mail addresses: [nmorrow@tulane.edu](mailto:nmorrow@tulane.edu) (N. Morrow), [nkwake@gmail.com](mailto:nkwake@gmail.com) (A.M. Nkwake).

working with assumptions enables the agency of evaluators and stakeholders to move forward despite uncertainty. The first step to evolving an assumption-aware practice is development of a common typology for the most common types of assumptions found in evaluative practice. The authors propose the following as the basis for an assumption aware typology for evaluands:

1.1.1. Normative assumptions

These are the considerations, right from before an intervention is devised, that there exists and problem (or opportunity) that deserves a response—that there is a discrepancy between the reality and what is ideal.

1.1.2. Diagnostic assumptions

Diagnostic assumptions are stated as stakeholders’ perceptions of the major and minor causes of the core problems. Since the intervention to address a problem is based on the causes of that problem, diagnostic assumptions are crucial to a normative theory and need to be examined from design, implementation and evaluation perspectives

1.1.3. Prescriptive assumptions

Prescriptive assumptions have to do with the intervention or strategy devised to resolve the problem or to reach a stated program goal, which represents stakeholders’ beliefs of what could be the most appropriate approach for addressing the problem or responding to an opportunity.

1.1.4. Causal assumptions

Causal assumptions explain how initial and intermediate changes resulting from program implementation will bring about longer term changes. The difference between prescriptive and causal assumptions is that while prescriptive assumptions are related to strategies (and alternatives) devised to address a problem, causal assumptions relate to how the immediate results of a strategy program or intervention (outputs) are expected to lead to long-term desired changes (outcomes and impacts).

1.1.5. External or contextual assumptions

Considered to be factors in the external environment of a program beyond stakeholders’ control that are preconditions for achieving expected outcomes.

At the foundation of evaluation practice is the determination of value and how to go about assessing it. Mertens’ Transformative Evaluation begins with articulating the underlying assumption of all evaluands as action for the promotion of social justice and human rights (Mertens, 2008). By starting with a clear articulation of the primary purpose of program or intervention, there is a larger framework to guide design and methodological choices. With this articulation of the axiological basis of Transformative Evaluation, a complete evaluation approach can be built on a foundation that clearly favors cultural relevance and responsiveness, promotion of underrepresented and marginalized voices, and the promotion of social change. Understanding the power dynamics and the assumptions behind actions and relationships are therefore the key to obtaining the desired programmatic or policy impacts. In this volume, Mertens extends these ideas of Transformative Evaluation to the different types of assumptions that are made in design and evaluation processes. She then provides examples of the utility in assumption awareness when working with wicked problems that involve the interaction of multiple systems, contradictory perspective and complexity of interactions that obscure causal relations such as climate change, health behaviours and sexual violence. The transformative perspective is then articulated with respect to axiological, ontological, epistemological and methodological assumptions (see Fig. 1).

Transformative Axiological Assumption holds that evaluation should be performed in the service of social justice; this is important because discrimination and oppression are systemic; failure to address these problems sustains an oppressive status quo.
Transformative Ontological Assumption holds that different versions of reality exist and these come from different social positionalities; there are consequences associated with accepting one version of reality over another; evaluators have a responsibility to make visible those versions of reality that sustain oppression and those that support the path to social justice.
Transformative Epistemological Assumption holds that differences in power impact the ability to accurately identify problems and solutions; evaluators need to establish trusting relationships with the full range of stakeholder groups in order to obtain an accurate picture of the phenomenon under study.
Transformative Methodological Assumption holds that dialogic moments are critical to understanding phenomenon from different perspectives; evaluators can use mixed methods to be responsive to diverse stakeholder groups and to capture the complexity of the phenomenon under study in ways that contribute to social transformation.

Fig. 1. Different levels of transformative assumptions (Mertens this volume).

Mertens argues that interrogating assumptions from a social justice perspective leads to more culturally relevant and therefore more appropriate and effective interventions. In particular, transformative causal assumptions take into account contextual assumptions related to cultural complexity and human rights. A clear outcome of transformative evaluation is a focus on the role of the evaluator, the participation of a diversity of participants, and a focus on process. “Evaluators reflect and make explicit their axiological, ontological and epistemological assumptions, they are better able to choose the methodologies, for use in their inquiries” (Mertens this volume).

Choice of methodology is perhaps the most concrete and visible aspect of the transformative approach and implies many of the aspects seen in assumption-aware tools including iterative participation and surfacing underrepresented or unarticulated theories of change.

Furthermore, surfacing assumptions helps avoid the misunderstanding of the nature of problems from the stakeholders’ perspective—so they can use their assets and actions to participate in the solution. Purposeful inclusion, local oversight and mixed methods are the recommendations derived from cases presented from India, Mali, Kyrgyzstan—examples of problems that had more complex roots in context of power relations than perhaps initially considered by program designers.

Finally, examining assumptions increases potential for social change. Focus on composition of the team, governance and context are the often unstated but crucially important aspects of successful program design and evaluation. Mertens (this volume) then incites evaluators to base evaluation questions on the central axiological assumption of the proposition of social justice and human rights and to focus inquire on resilience and capabilities of the stakeholders. Evaluators should use mixed methods with a focus on participation and broad qualitative understanding of contextual aspects and a focus on stakeholder experience and world-view. Applying the transformative approach to working with assumptions ultimately promotes empowerment, dignity and agency of all stakeholders involved.

In the third article, assumptions about best practices and gold standards approaches to complexity in program evaluation theory are discussed in Chen's 'Interfacing Theories of Program with Theories of Evaluation for Advancing Evaluation Practice: Reductionism, Systems Thinking, and Pragmatic Synthesis'. Reductionism, where simple cause and effect relationships between interventions and outcomes are measured through experimental evaluation designs, are first examined in situations that Chen refers to as 'moderate complexity'. Moderate complexity are typified by development and public health interventions where actions are based on some formal theory and/or experimental results but at the same time applied in real world contexts. Often put forward as the gold standard for evaluation validity, experimental evaluation designs seen today, according to Chen, proceed from work by Tyler in the 1940s and 50s on the relation of activities to program objectives and goals that have been married with a singular focus on the importance of internal validity and controls following Campbell evaluation (Tyler, 1942; Tyler, 1950; Campbell & Stanley, 1963). An important set of assumptions underlie these approaches. First it is assumed that the goals and objectives of the program are clear, shared, appropriate and relevant. These types of assumptions can be broadly grouped under the categories of prescriptive and diagnostic assumptions put forward by Nkwake (2013). Perhaps most problematic in complex contexts is the predominant assumption that the functional relationship between an intervention and intended outcome observed under controls are maintained in real world settings—captured in the concept of a category of causal assumptions (ibid.)

Systems approaches that focus on relationships and capture feedbacks are often put forward as a way to better characterize the function of a program in a complex setting and capture both intended and unintended consequences (Patton, 2011; Peters, 2014). Perhaps most useful to understanding how programs actually function in real world settings, the systems concept of bounded rationality captures information about people acting based on differing purposes or rationalities as a primary determinate of social system behavior (Midgley, 2003). The potential benefits of increased understanding from these systems approaches seem to be offset by a lack of development in analytical methods and the tendency to provide too much information and produce results are difficult to communicate to stakeholders. The underlying assumption then seems to be that both complexity and systems thinking are 'cutting-edge' in evaluation and should therefore be considered together (Chen this volume). This could be considered as a paradigmatic assumption.

Pragmatic synthesis is then proposed by Chen as another way to develop evaluation program theory and assumes that interventions have dual purpose. Interventions are intended to both be effective through appropriate action and at the same time to manage context. Borrowing a criteria from a pragmatist perspective, pragmatic synthesis asks if the program evaluation theory is useful by in view of a real world challenges and expectations of a manager or client involved in an actual development or global health type program. In Chen's pragmatic synthesis, assumptions about the program theory and the evaluation theory are continually challenged assessed and the program is adjusted in the following steps:

- seeking the existence and desirability of joint effects as complementary to ruling out threats to internal validity
- capturing unintended positive effects and other real world benefits from simple tools like a checklist as an alternative to information overload from some holistic approaches
- expand credible evidence on effectiveness and viability
- focus transferability as a targeted approach to evaluating external validity

- valuing stakeholder theories of causality and added value of these theories to formal theory
- bottom-up dissemination of evaluation learning opposed to top-down approaches

These steps interactive and participatory approaches to articulating and challenging the assumptions that may limit the usefulness of evaluations in specific ways and are particularly problematic in evaluations of complex evaluands or in complex contexts. Chen essentially looks at explicating assumptions in program theory as being useful in two ways. Prescriptive assumptions about how to do the intervention inform an action model. Descriptive assumptions about causality inform change model and the intended causal linkages. Chen is basically saying that explicating assumptions both helps a program be effective and work in context. Pragmatism is about interventions that work on these two levels.

Chen proposes that the benefits of following a pragmatic synthesis approach include:

- With the availability of pragmatic synthesis, evaluators have a better perspective to follow for assessing these program and better serving stakeholders evaluation interest and needs.
- without the information overload or overwhelming complexity—
- Proposes a set of tools/approaches: adjuvants, joint effects, viability evaluation, and the bottom-up approach provide more manageable and applicable options for addressing systems issues.
- Testing and cumulating stakeholders' knowledge and technologies.
- New ideas and insights for developing cutting-edge evaluation concepts and approaches.

In summary, Chen rejects Universalist claims of gold standards and best practices such as reductionism and systems thinking on the pragmatist grounds that they are just not that useful for the development of program evaluation theory in moderately complex contexts where most development and global health interventions take place. A contingency approach that acknowledges that different situation require different perspectives and approaches is preferred. Although there may be other tools, approaches to surfacing, testing and adjusting assumptions seem to be at the heart of the proposed steps to making program evaluation theory and their associated policy and programs more useful.

In the fourth article of this volume, Archibald et al. use case studies based on evaluative thinking workshops from Ethiopia and Zambia to illustrate the links between examine assumptions, developing Theories of Change, and adaptive management. Evaluative Thinking builds upon the basis of Critical Thinking, the systematic review of evidence to support beliefs and/or action, to the question of the value of that action (Trochim et al., 2012; Preskill & Torres, 1999; Buckley et al., 2015). The authors identify the four elements of ET at the basis of their ToC workshops as, "(1) identifying assumptions, (2) posing thoughtful questions, (3) pursuing deeper understanding through reflection and perspective taking, and (4) informing decisions in preparation for action." Further developing the link between examining assumptions and the pragmatist tradition discussed in the article by Chen, examining assumptions is undertaken to clarify and come to agreement on a common understanding as the basis for promoting action. And then assessing the value of that assumption by the resulting desirable outcomes of that action.

The workshops that are the subject of the case studies are intended to improve project management by Catholic Relief Services staff in Ethiopia and Zambia as part of the larger Gate's Foundation Emergency Capacity Building program. The projects

themselves are funding by USAID that has been promoting the use of Theories of Change to enhance management particularly of complex programs in complex environments and more specifically in its Technical References for Food For Peace Development Food Assistance Projects, “identify and explicitly present all assumptions underlying the ToC” (USAID, 2015b, p. 15). The workshops, according to the authors, are intended to foster continuous reflection and learning to enable the emergence of adapted practices and increase the relevance and sustainability of CRS programs. To do this, they follow a process of surfacing assumptions, reviewing evidence for alternative explanations, and then building consensus on a course of action or common understanding.

The authors identify three types of assumptions, following Nkwake’s classification, as important to their workshops causal, prescriptive and paradigmatic assumptions are useful in surfacing during the development of a ToC. Authors suggest that ToC models are seen as offering practitioners a way to engage more transparently with complex change based on experience. They also identify some risks associated with this process are related to the subjectivity and a degree of arbitrariness surrounding the identification and selection of assumptions. A second round of workshops after initial implementation of the projects found that causal assumptions in the program logic that had been shown to not hold true such as willingness to take on new approaches or include men in women’s savings groups. The reaction of the participants was to take another look at causal assumptions and prompted staff to collect more data, set up small experiments to elaborate the program logic.

Perhaps the greatest revelation in the Evaluative Thinking workshops, was the engagement and interest of a wide range of frontline and back office program staff. “What we found instead was that these front-line community educators were hungry to learn ways to conceptualize the theory behind their work, to share their frustrations about the real-world barriers to project success that they experience in their daily practice, and to express their desire to be more involved in program planning.”

In evaluating the effectiveness of the workshops, participants were asked what they could do now after the workshop that was new— and surprising number of the comments focused on working with assumptions:

“Teasing out assumptions along pathway model; Seek alternative explanation of assumptions before I take them as true and valid assumptions; Question certain assumptions through meetings and trainings including during my field visits; Critically thinking about assumptions; Consciously identify assumptions; analyze assumptions for alternative explanations; Have colleagues to mine my assumptions on the model and help me use evaluative thinking to re-evaluate my model; How to spot implicit and explicit assumptions in the design of a project/activity; Test assumptions; Am able to identify assumption in my project pathway model; I am able to accept critics because the peer review challenged me that it’s not all that I see right that is right; other people may see something different; I knew about the six hats exercise but feel I can use it more effectively with a group of people to surface “unseen” interpretations or understandings related to a specific issue; Be more critical of the assumptions that I held about my work— a reawakening of reason.” (Archibald et al. this volume)

Participants in the ET workshops identified working with assumptions as critical learning and readably applicable to their work with frontline staff. Identifying assumptions, testing alternatives and revisiting causal assumptions seem to be critical to adaptive management. Working with assumptions in Evaluative thinking seemed particularly relevant when facilitators asked

participants “to identify possible assumptions that they might be leading to the poor state of health in their community. This aspect of our work hints strongly at the prospect of ET serving as a conduit for frontline staff and beneficiary involvement in evaluation and program planning processes”.

A methodological note on working with the assumptions that underlie methodological choices in evaluation design for Complex Social Programs is the subject of the 5th (or perhaps move to 4th) article in this volume. Chatterji (this volume) uses a variety of social programs from sectors as diverse as improving educational attainment and supporting basic needs of refugee populations to illustrate the complex characteristics of many contemporary evaluands. When choosing approaches to designing an impact evaluation for a CSP, these complex characteristics confound the basic assumptions of more experimental designs and specifically violate some key assumptions of Randomized Control Trials (ibid). Chatterji then provides insight on how to recognize is a social program is indeed a CSP and suggests potential approaches to undertaking impact evaluations as an alternative to RCT. Particularly important is iterative reflection and the use of both qualitative and quantitative methods in an appropriate sequence the author describes as Extended-term Mixed Method approach.

Chatterji (this volume) continues with the premise that a plethora of assumptions underlie any social intervention and that it is necessary to use a number of approaches to open up the black box of program theory. Comprehensive programs with multiple levels and multiple interventions introduce more complex and an even more diverse set of underlying assumptions. The author counsels a pragmatic approach in which appropriate tools are chosen based on the phase of the design to implementation to evaluation continuum (Chatterji, 2005, 2009). Through the use of multiple sources of information and multiple approaches and tools, alternative theories of causation may begin to be ruled out. As the program is implemented in context, reflection and analysis can begin to prove, disprove and eventually strengthen the program theory as causal assumptions are systematically identified and tested.

In the sixth article, Morrow and Nkwake, argue that assumption-aware tools for more appropriate and effective program theory development have not been systematically compared or organized into useful descriptive categories based on intended use. The paper reviews twelve alternative approaches with associated tools that work with assumptions to strengthen the causal linkages and theory for program design and evaluation. Evaluators interested in the potential benefits of understanding how programs work with different stakeholders in different contexts have often characterized evaluations that focus only on measuring achievement of predetermined objectives as ‘black box’ evaluations (Dyehouse et al., 2009; Dyehouse, Bennett, Harbor, Childress, & Dark, 2009). These benefits of theory-based approaches are numerous and appealing for a number of different applications such as learning, greater accountability, implications for scalability as well as contribution to innovation and adaptation. Evaluators interested in unpacking the mechanisms seen to be producing outcomes through explicating mechanisms and how they work in context have developed a number of methods and tools widely referred to as theory-based approaches likened to ‘white boxes’ or ‘clear boxes’ (Dyehouse et al., 2009). The Logical Framework Approach placed assumptions definitively external to the program and outside the boundary of program management. Nkwake’s (2013) diagnostic, prescriptive, and causal assumptions highlight the value of working with assumptions internal to the program as well as assumptions that help define the nature and extent of program boundaries or programing ‘box’, what should be contained in the ‘box’, and how best make linkages within and across boundaries. Ten of the reviewed assumption-aware tools

are mentioned in previous assumption focused articles by Nkwake (2013) and Leeuw (2003). The Casual Lop Diagram and Strategy Assessment Approach were added to the review from subsequent discussion with fellow evaluators on the list serve M&E News (2013).

Participation and enhanced facilitation is a primary method for both surfacing assumptions and for comparing alternative causal paths, linkages, and assumptions. Slightly more than half of the reviewed tools explicitly require participation. Participation is usually further defined by large and diverse number of stakeholders including non-traditional stakeholder groups. Nearly all of these approaches cite a weakness as these processes can be time consuming. In the end, any additional time requirements need to be balanced against the potential benefits and authors tend to propose that the extra effort is warranted when working with complex evaluands and contexts.

### 1.2. Discussion current and future state of assumption-aware practice; what is the potential to empower the agency of evaluators, stakeholders and evaluation practice?

Perhaps most striking when taking stock of assumption-aware evaluation practice is the sheer ubiquity of assumptions underlying the design and evaluation process. Assumptions play key roles at different stages of program design, implementation, and evaluation. There are assumptions in the program theory and also in the methodologies and approaches to valuing the program. The authors in this volume and those that have been cited from the literature discuss the problems encountered while working with assumptions from a diversity of viewpoints and for application in a diversity of evaluation situations. This is likely one of the key factors contributed to a proliferation of tools that address felt professional requirements of evaluators to do more to surface and articulate assumptions at each of the stages of design and evaluation.

To determine if these tools and approaches have, on the whole, been proven to improve practice and have potential to improve future practice, one may ask if they help evaluators be more efficient and effective at their work in the face of complexity. Perhaps more importantly if the evaluation process has had an impact. We propose agency as a valid expression of evaluation process impact particularly for complex evaluands; are the stakeholders more empowered with deeper knowledge of their theories including articulation of previously unarticulated assumptions; Are evaluators better able to work in dynamic and chaotic contexts and with complex evaluands in ways that benefit a broad and inclusive group of stakeholders?

Narayan (2005) defines Agency as the expansion of freedom of choice and action to shape one's life-. According to Sen (2001, 2002), a person's agency is one's ability to act on behalf of what he or she values and has reason to value. Agency is defined with respect to the goals at hand, as well as the freedoms and capabilities to pursue those goals (Alkire, 2006).

Meyer and Jepperson (2000: 117) analytically distinguish three types of agency:

- Agency for itself
- Agency for others
- Agency for standards and principals

*Agency for itself* (or for 'themselves') takes into account the capabilities and goals of individuals and collectivities such as professions, specialist organizations and national states acting on their own behalf. *Agency for others* considers actions of individuals, groups, organizations and states on behalf of others-such as

individuals, families, groups, organizations or other states. *Agency for standards and principles* acts on behalf of a principle, such as human rights, transparency, good governance, social responsibility, and science-driven rationalization (Meyer and Jepperson 2000; Ahonen, 2015).

As demonstrated in this volume, examining evaluand and evaluation assumptions is essential for successful interventions and helpful evaluations. We therefore propose three levels at which agency for assumption-aware evaluation practice can be fostered:

1. Assumption-aware evaluation as agency for themselves: Evaluators are committed to and have capacity for examining their own assumptions about evaluands, stakeholders' expectations, and appropriateness of methods.
2. Assumption-aware evaluation as agency for others: Evaluators are committed to and have the capability to help/facilitate program stakeholders to examine their assumptions about evaluands and evaluations (their use, credibility and boundaries for rigor and resourcing).
3. Assumption-aware evaluation as agency for standards and principles: That institutions (development organizations, donors, programs, evaluation associations, etc. put in place a policies and standards that encourage the explication of assumptions.

### 1.3. Assumption-aware evaluators and agency for themselves

In order to become more aware of the assumptions we make, evaluators need to recognize the power of the unconscious mind. Social psychologists have seen a rapid accumulation of evidence both for the limitations of the conscious mind and the power of the unconscious mind. The limitations of the conscious mind are highlighted in Wegner's (2002) analysis of the role of consciousness in human thinking and action. He dramatically demonstrates "the illusion of conscious will" in which human beings not only claim responsibility but also intention for actions over which they had exactly no control. In a variety of tasks and contexts, humans tend to attribute their own behavior to premeditated intention, rather than to unconscious processes. Conscious will is consistently given more credit than is due, despite robust evidence about its limitations. The mere conscious desire not to be biased does not eliminate implicit bias, contrary to conscious intention.

Implicit social cognition is an umbrella term used to capture the idea that thoughts and feelings may operate outside the purview of conscious awareness, control, and intention; in contrast, explicit social cognition involves thought and deliberation (Banaji & Greenwald, 1994). In a growing, multi method body of research, automaticity (unexamined assumptions) has been found to play some role in virtually every cognitive process studied, and its inevitability has been cleverly termed the "unbearable automaticity of being" (Bargh & Chartrand, 1993). For example, research has indicated the influence of implicit gender stereotypes (Banaji & Hardin 1996) and implicit age stereotypes in human resource practices.

The more evaluators intentionally bring the conscious mind to bear, the more we become mindful of the potential for unconscious biases, being intentional about examining or unearthing implicit bias is similar Banaji's (2001) metaphor of the driver of a misaligned car deliberately counteracting its pull. I.e. evaluators can develop conscious strategies to counteract the pull of their unconscious biases. Evaluators need to constantly beware of these biases.

#### 1.4. Assumption-aware evaluation practice

Evaluators are committed to and have the capability to help/facilitate program stakeholders to examine their assumptions about evaluands and evaluations (their use, credibility and boundaries for rigor and resourcing). The major question here is that of tools and approaches available to facilitate the examination of stakeholder assumptions. These tools are discussed in chapter 6.

#### 1.5. Assumption-aware standards and principles

In her definition of agency within empowerment, Narayan (2007) argues that individuals' expressions of agency (agency for themselves and for others) must be supported by an enabling institutional and social environment. This is what is referred to as "opportunity structure". Agency for principles and standards requires institutions to establish policies that encourage the examination of assumptions. USAID's evaluation policy (2013) places clear emphasis on articulation of assumptions in program designs: "Compared to evaluations of projects with weak or vague causal maps and articulation of aims, we can expect to learn much more from evaluations of projects that are designed from the outset with clear development hypotheses, realistic expectations of the value and scale of results, and clear understanding of implementation risks."

Other institutions have established evaluation policies that do not directly emphasize the examination of assumptions but allude to the need to examine assumptions in terms of objectivity, etc. For example, The Global Environment Facility M&E policy articulates among the guiding principles that: "Evaluation practices will follow established standards, ensuring the credibility, impartiality, transparency, and usefulness of evaluation

projects." The European Bank for Reconstruction and Development Evaluation policy (2013) has for one of its principles: "Effective evaluation requires that the intended outcomes of institutional activities are expressed and monitored in a way that can be assessed on the basis of objective evidence."

United Nations Development program (UNDP)'s evaluation policy (2010) outlines the 'impartiality' principle that points to the need to clarify stakeholder expectations (and assumptions):

*"(e) Impartiality.* Removing bias and maximizing objectivity are critical for the credibility of the evaluation and its contribution to knowledge. Prerequisites for impartiality are: independence from management; objective design; valid measurement and analysis; and the rigorous use of appropriate benchmarks agreed upon beforehand by key stakeholders. In addition to being impartial, evaluation teams should include relevant expertise and be balanced in their gender and regional composition."

Opportunities for assumption-aware evaluation practice are visible in guiding principles for evaluators. E.g. American Evaluation Association's guiding principle for evaluators (A3) – "Commitment to systematic, data-based inquiry" emphasizes that evaluators should provide accurate and detailed information on methods and approaches used, including limitations, from conceptualization to use of findings, to allow others to understand, interpret, and critique evaluators' work. A list of examined assumptions that inform methods choice would be a great addition to this list of detail.

Another principle (Commitment to honesty and integrity) requires evaluators to be explicit about the interests and values of evaluators, clients, and other stakeholders concerning the conduct and outcomes of an evaluation. This in some ways may involve

#### **Box 1.** Assumptions in discussing evaluation purpose

Examining these assumptions gets to the following questions (Nkwake, 2015):

- Has the evaluation's purpose (including questions) been appropriately derived?
- Did the formulation and prioritization of questions engage stakeholder input to make the purpose agreeable, rational, and feasible?

#### *Assumptions in selecting evaluation designs and methods.*

Examining these assumptions gets to the following question:

- To what extent do the measures (methods, constructs, variables, comparisons) accurately depict the essential features of a program?

#### *Assumptions in determining measures, tools, and data collection*

Examining these assumptions gets to the following questions:

- How acceptable are the measures, tools, and data collection procedures
- To what extent does the data obtained from evaluation measures truthfully depict the program's features, dynamics and outcomes?

#### *Assumptions in analysis, inference, interpretation, and conclusions*

Examining these assumptions gets to the following questions:

- Are conclusions and inferences correctly derived from evaluation data and measures that generate this data?
- To what extent can the evaluation findings provide conclusions about other situations?

#### *Assumptions about evaluation use*

Examining these assumptions gets to the following questions:

- How will evaluation results be put to use?
- Will consequent decisions be in line with the conclusions of the evaluation?
- Does the evaluation address values to which major stakeholders (or audiences)

explication of stakeholder assumptions regarding the evaluand and evaluation.

But policies—both of client organizations and evaluation association (guiding principles) can be more intentional in promoting assumption-aware evaluation by Outlining and define important evaluand assumption categories that are of priority, e.g. Normative, diagnostic, prescriptive, causal and external assumptions. Policies could also be more normative in illuminating important evaluation assumptions that need to be examined during the conduct of evaluations. Foremost, assumptions made in framing the evaluation purpose are essential to the development of evaluation design and evaluation theory as well as the more practical steps of creating an evaluation terms of reference or in responding to the terms of reference in the inception phase. Nkwake (2015) proposes normative questions can be posed at each step of an evaluative process to ensure essential awareness of assumptions to improve the evaluation and promote program success (see Box 1 ).

## 2. Conclusion

In sum, assumption-aware evaluation practice is foundational to performing sound evaluations yet it remains a poorly developed area of evaluation theory and practice to date. Emergent approaches to working with assumptions, their challenges and potential need discussion and synthesis. Assumption awareness in program design and evaluation ability for stakeholders to ‘act on their values’ and evaluators to produce evaluations that more carefully consistently describe those changes that really mattered.

This volume has brought together a diversity of viewpoints on how assumptions impact evaluators work, particularly in complex contexts and with increasingly complex evaluands. Common themes emerge from each of the authors on the importance of having a solid awareness of the role of assumptions when selecting methodology, approaches, and tools. These are not choices that should be made consciously with full awareness of the underlying assumptions of those choices. There is also a tendency towards pragmatism – working with assumptions brings clarity to your evaluation and program designs that is a hallmark of program success. Finally, it seems that the evaluation policy environment and normative guidance have yet to catch up with the emergent practice in working with complex evaluands and in complex environments. Those closest to the stakeholders working with complex social interventions and on wicked problems know that surfacing and articulating assumptions is at the core of an evaluators evolving role. Mandated tools and approaches from funding partners, training and the literature have the potential to contribute to an environment more conducive to facing complexity if they put more focus on assumption awareness. This volume is only a first step towards a broader inquiry into the nature of

assumptions in evaluative processes and practice. Intentional program of research, innovation and comparison of more assumption aware evaluative methods will likely make a relevant contribution to program evaluation and design in the age of complexity.

## References

- Ahonen, P. (2015). Aspects of the institutionalization, of evaluation in Finland: basic, agency, process and change. *Evaluation*, 21(3), 308–324.
- Alkire, S. (2006). Measuring agency: issues and possibilities. *Indian Journal of Human Development*, 1(1) .
- Banaji, M. R. (2001). Implicit attitudes can be measured. In H. L. Roediger, J. S. Nairne, I. Neath, & A. Surprenant (Eds.), *The nature of remembering: essays in honor of Robert G. Crowder* (pp. 117–150). Washington DC: American Psychological Association.
- Banaji, M. R., & Greenwald, A. G. (1994). Implicit stereotyping and prejudice. In M. P. Zanna, & J. M. Olson (Eds.), *The psychology of prejudice: the ontario symposium: (Vol. 17. pp. 55–76)*. Hillsdale, NJ: Erlbaum.
- Banaji, M. R., & Hardin, C. D. (1996). Automatic stereotyping. *Psychological Science*, 7(3), 136–141.
- Bargh, J. A., & Chartrand, T. L. (1993). The unbearable automaticity of being. *American psychologist*, 54(7), 462–497.
- Buckley, J., Archibald, T., Hargraves, M., & Trochim, W. M. (2015). Defining and teaching evaluative thinking: insights from research on critical thinking. *American Journal of Evaluation*, 36(3), 375–388.
- Campbell, D. T., & Stanley, J. (1963). *Experimental and quasi-experimental designs for research*. Chicago: RandMcNally.
- Chatterji, M. (2005). Evidence of “what works”: an argument for extended-term Mixed-method (ETMM) evaluation designs. *Educational Researcher*, 34(6), 13–24.
- Chatterji, M. (2009). In K. B. Ryan, & J. B. Cousins (Eds.), *Enhancing impact evidence on the how global educational initiatives work: theory, epistemological foundations, and guidelines for applying multiphase, mixed method design* the SAGE International.
- Dyehouse, M., Bennett, D., Harbor, J., Childress, A., & Melissa, D. (2009). Comparison of linear and systems thinking approaches for program evaluation illustrated using the Indiana Interdisciplinary GK- 12. *Evaluation and Program Planning*, 32, 187–196.
- Mertens, D. M. (2008). *Transformative research and evaluation*. Guilford press.
- Meyer, J. W., & Jepperson, R. L. (2000). The ‘actors’ of modern society: the cultural construction of social agency. *Sociological Theory*, 18(1), 100–120.
- Midgley, C. (2003). *Systems thinking*. Thousand Oak, CA: Sage.
- Nkwake, A. (2013). *Working with assumptions in international development program evaluation*. New York: Springer.
- Nkwake, A. (2015). *Credibility, validity and assumptions in program evaluation methodology*. New York: Springer.
- Patton, M. Q. (2011). *Developmental evaluation: applying complexity concepts to enhance innovation and use*. Guilford Press.
- Peters, D. H. (2014). The application of systems thinking in health: why use systems thinking. *Health Research Policy and Systems*, 12, 51.
- Preskill, H., & Torres, R. T. (1999). *Evaluative inquiry for learning in organizations*. Thousand Oaks, CA: Sage.
- Sen, A. (2001). *Development as freedom*. Oxford Paperbacks.
- Sen, A. K. (2002). *Rationality and freedom*. Cambridge, Mass: Belknap Press.
- Trochim, W., Urban, J. B., Hargraves, M., Hebbard, C., Buckley, J., Archibald, T., Johnson, M., & Burgermaster, M. (2012). *The guide to the systems evaluation protocol (v2.2)*. Ithaca, N.Y: Cornell Digital Print Services.
- Tyler, R. W. (1942). General statement on evaluation. *Journal of Educational Research*, 35, 492–501.
- Tyler, R. W. (1950). *The basic principles of curriculum and instruction*. Chicago: University of Chicago press.
- Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.