

**In Name Only? An exploration of the operationalization of empowerment outcomes
in transformative participatory evaluations**

Joanne Tucker

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements
for the Master of Arts in Education

Faculty of Education
University of Ottawa

© Joanne Tucker, Ottawa, Canada, 2016

Abstract

This thesis presents an exploration of the operationalization of empowerment outcomes in research on transformative participatory evaluations, focusing on the context of international development evaluation. Covering a 15 year period from 1999 and 2014, through the examination of the empirical research literature, the study explores: 1) how empowerment outcomes are measured, 2) the extent to which these outcomes demonstrate empowerment principles, and 3) which factors and conditions appear to enable or detract from the attainment of these outcomes. I found that the current state of the empirical research on transformative participatory evaluation to be largely comprised of reflective case narratives that rely solely on scarcely documented qualitative methods. In general, transformative outcomes do tend to mirror empowerment principles such as ownership, inclusion, democracy, and social justice. Finally, I found that various factors and conditions are critical to the reported attainment of transformative outcomes, particularly in relation to the local program context, for example, reforms in local and international governments that support increased local control over resources and governance, organizational structures and priorities that are congruent with empowerment objectives, and previous experience with empowerment processes. I also highlighted deficiencies in the current empirical research and call on the evaluation community to improve research on transformative approaches to participatory evaluation by suggesting critical areas for practice and writing. These include strengthening research designs and the use of meta-evaluations, further defining and clarifying key terms, and providing rich detail to facilitate further learning in this area.

Keywords: participatory evaluation, transformative evaluation, empowerment

Acknowledgements

First and foremost, I would like to thank my thesis supervisor, Dr. J. Bradley Cousins for his consistent support throughout the research process. I truly consider myself to be quite lucky to have worked closely with him over the past few years. His door was always open for matters personal or technical, his patience and joviality unwavering.

I would also like to thank my committee members for their important guidance and inspiring discussion. Thank you Dr. Peter Milley, Dr. Richard Maclure and Dr. Elizabeth Whitmore.

Finally, I would like to thank my friends and family for their unfailing support and encouragement. This accomplishment would not have been possible without them.

Table of Contents

Abstract	ii
Acknowledgements	iii
Tables and Figures	vi
1. Introduction	1
2. A Review of Transformative Collaborative Approaches to Evaluation and Current Critiques	5
2.1 The concept of empowerment	5
2.2 International development evaluation	6
2.3 Collaborative approaches to evaluation	7
2.4 Critiques of transformative CAE	9
Summary	15
3. Theoretical Orientation	16
4. Research Questions	16
5. Methods	18
5.1 Sample Development	18
5.2 Sample description	20
5.3 Data Analysis	23
6. Empowerment Outcomes in Transformation Approaches to Collaborative Evaluations	24
6.1 Operationalization of transformative outcomes	24
Summary	25
6.2 Empowerment Principles	26
6.2.1 Political	28
6.2.2 Social	29
6.2.3 Economic	30
6.2.4 Agency	31
Summary	31
6.3 Factors and Conditions that have Contributed or Detracted from the Attainment of Transformative Outcomes	32
6.3.1 Pre-existing Factors and Conditions	33
6.3.1.1 Community and Program Context	33
6.3.1.2 Relationships	35
6.3.1.3 Evaluation Rationale	36

6.3.1.4 Organizational Experience with Evaluation	40
6.3.1.5 Evaluator Paradigm.....	42
6.3.1.6 Resources	46
Summary	46
6.3.2 Mediating Activities.....	47
6.3.2.1 Control Over the Process	47
6.3.2.2 Stakeholder Diversity.....	49
6.3.3 Depth of Participation	51
Summary	53
Conclusion	54
7. Discussion and Conclusion	55
7.1 Limitations of the Study.....	55
7.2 Thematic Review of Findings.....	56
7.2.1 Context is crucial	56
7.2.2 The politics of engagement.....	59
7.2.3 Clarity on the concept of empowerment.....	61
7.2.4 What constitutes good evidence?.....	61
7.3 Implications for Research	64
7.3.1 Small sample size.....	64
7.3.2 Variability of data	64
7.3.3 Wider Range of Research Designs	65
7.3.4 Authorship.....	65
7.3.5 Power Dynamics	65
8. Conclusion	66
References.....	68

List of Tables and Figures

Table 1. Study Sample	20
Table 2.1. Evaluation Rationales Across Cases.....	37
Table 2.2. Multiple Evaluation Rationales	37
Figure 1. Critical Factors in Achieving Empowerment Outcomes in Collaborative Approaches to T-PE.....	54

In Name Only? An exploration of the operationalization of empowerment outcomes in transformative participatory evaluations

1. Introduction

Participatory development planning and practice were embraced by many governmental and non-governmental agencies and organizations in the 1970s (Guijt & Shah, 1998). Today, in the interest of aid sustainability and equity, as well as an emerging focus on aid effectiveness and efficiency, many organizations have also embraced collaborative approaches to evaluation (CIDA, 2012). In addition, increasing numbers of funding bodies are requiring that organizations play a role in their evaluations, and even conduct evaluations in-house (Fitzpatrick, 2012).

Evaluation is a field of both study and practice. As such, various definitions of what evaluation is and how it is done have been put forth by a variety of scholars and practitioners. Greene (2013) provides a definition of evaluation that most closely reflects my own. She defines evaluation as the systematic process of gathering and interpreting empirical information in order to make a judgment of quality and worth about the program, policy, or practice being evaluated. Evaluation takes on many forms, ranging from randomized control trial (RCT) designs to more participatory approaches. Further, Greene (2013) states that evaluation is always “saturated with values”, and therefore always favours someone’s interests over the interests of others. Therefore, different evaluation approaches are chosen based on identifying whose questions will be addressed in an evaluation. Collaborative approaches to development evaluation, rooted in a social justice epistemology, have been utilized in the field for over several decades (Brisolara, 1998; Cousins & Whitmore, 1998). Deemed to counter the potential harmful effects of more scientific, positivistic approaches to evaluation (e.g., objectifying research subjects), collaborative approaches draw from participatory research and development approaches aimed to increase broader political participation of citizenry and community participation (Cornwall & Gaventa, 2009).

Collaborative approaches to evaluation (CAE) such as participatory evaluation (PE) as defined by Cousins and Whitmore (1998) implies that “when doing an evaluation, researchers, facilitators, or professional evaluators collaborate in some way with

individuals, groups or communities who have a decided stake in the program, development project, or other entity being evaluated” (p.87-77). Further, they stipulate that the knowledge produced through the evaluation is co-created by participants. There is now a wide variety of CAE being practiced worldwide, which can be categorized along three fundamental process dimensions: control of technical decisions, stakeholder selection, and depth of participation of stakeholders (Cousins & Chouinard, 2012; Cousins & Whitmore, 1998). Furthermore, CAE projects can be defined by their goals; for some practitioners, CAE are used to enhance evaluation utility among stakeholders (Practical PE), while others adopt them as a means for reallocating power in the production of knowledge and promotion of social change (transformative or T-PE)(Cousins & Whitmore, 1998).

These transformative approaches are the focus of this thesis. Empowerment evaluation, a notable transformative approach as defined by Fetterman over the years, is a process that facilitates the development of perceived and actual control over the fate of a community or people joined by their relationship to a social program (Fetterman, 2005). These approaches involve stakeholders in the design and conduct of evaluations and utilize transformative processes that facilitate empowerment outcomes at the individual, group and institutional level. Empowerment outcomes include increased critical awareness and an ability to analyze life circumstances that an individual or group of individuals wish to change, feel that they have the ability to change, and act in a way to enact such change. Institutionally, these empowered groups may be engaged in a network with similar institutions to influence change (Schulz, Israel, Zimmerman, & Checkoway; 1995).

As transformative approaches to development evaluation have permeated practice, various criticisms have emerged from both evaluation and development communities (Cousins, 2005; Gregory, 2000; Miller & Campbell, 2006; Patton, 1997, 2005; Scriven, 1997; N.L. Smith, 1999; Worthington, 1999). For example, Miller and Campbell (2006) outline criticisms such as conceptual ambiguity, lack of unanimity in practice, and a lack of documented evidence of success, such as increased evaluation capacity, high levels of evaluation use, and increased perceived and actual self-

determination. Further, Gregory (2000) raises important questions about the mainstreaming of participation and its paradoxical potential to increase marginalization. While many of these concerns specifically target empowerment evaluation, they can be applied to collaborative transformative evaluation as a whole.

Consistent with the evidence provided by Christie (2003) regarding evaluation in general, scholars (Cousins, Whitmore & Shula, 2013; Fitzpatrick, 2012) argue that distinguishing different models for CAE should not be our focus because many practitioners do not actually rigidly subscribe to any one approach and that it is important that evaluations are designed according to the problems or questions they are intended to address and the information needs arising from the context within which the identified intervention (or evaluand) is operating. Many of these critiques have been addressed over time; however, responses have not been particularly compelling. Clearly, there is a need to better understand how transformative processes work, and if and how they achieve their social justice goals. On the whole, while the theory of transformative evaluation is well documented, empirical evidence of transformative outcomes remains relatively unclear to many in the field (see, e.g., Cousins, 2005; Miller & Campbell, 2006).

In the context of participatory development programming, the notion of participation has come under fire. Who defines participation? Why is participation important and how does it impact participants? How can participation be critiqued and pushed further to produce less harm and greater benefit to participants? In my opinion, these questions are not investigated deeply enough in transformative evaluation practice. It is common for evaluators to ask “why” in collaborative evaluation, but these questions are often framed in an “evaluability lens”- ‘Is it feasible?’ ‘Is it credible/valid?’ ‘Will it increase use?’, and so forth. (Cornwall, 2012; Hay, 2010; Fitzpatrick, 2012). These questions stem from concerns about evaluation rigour and often reflect donor needs, as opposed to those of intended program beneficiaries and program managers.

If we view evaluation as an activity that takes place in a political context that passes value judgements on a particular program, and the epistemological basis for employing participatory methods is historically linked to social justice, then involving participants in evaluation should have the same social justice ends in mind (Mertens,

2009). This is particularly the case when evaluations claim to utilize a transformative framework. If transformative CAEs are deemed to be less harmful than conventional approaches it is important that the evaluation community thoroughly examine and understand the impacts that result in the implementation of these approaches and provide clear evidence of their processes, challenges and successes.

In the field of development evaluation, the context in which CAEs are implemented becomes more complex. In the present study I focus specifically on these approaches in international development. I have chosen this context for two reasons. First, I hold a bachelor's degree in international development and globalization and my interest in evaluation was first piqued during my year abroad where I experienced the disconnect between program theory and practice, sparking an interest in program evaluation and its role in understanding programs in their own context. Second, according to Brisolara (1998), many of the theorists that contributed greatly to the formation of T-PE have their roots in the international development scene of the 1960s and 70s. This stream of participatory evaluation has a much longer and richer history than practical participatory evaluation (P-PE), which is a relatively recent approach that developed predominantly in North America (Cousins & Whitmore, 1998).

This qualitative study aims to develop our understanding of transformational outcomes as achieved in development evaluations utilizing transformative CAE. The study, guided by current criticisms of CAE, takes a critical look at the operationalization and measurement of transformative outcomes and the processes engaged when evaluators report these outcomes. It provides a review and integration of the empirical literature that aims to explore how transformative outcomes are operationalized and how such representations might be improved. This study will answer the following questions:

1. How are transformative outcomes operationalized and measured in transformative evaluations that take place in development settings in the current research on evaluation literature?
2. To what extent are transformative outcomes demonstrating empowerment principles such as ownership, inclusion, democracy, and social justice being realized in transformative collaborative evaluation approaches in development settings?

3. What factors and conditions have contributed to the attainment of these outcomes?
What factors and conditions detract from the attainment of these outcomes?

2. A Review of Transformative Collaborative Approaches to Evaluation and Current Critiques

In the following section I present a succinct review of the literature on CAE and, specifically transformative outcomes in international development evaluation. The section identifies the criticisms of these theories and practices and outlines the current gaps in the literature, which the present study intends to address. I first review the concept of empowerment and then turn to development evaluation concepts. I then focus on CAE concepts and trends and finally identify critiques of transformative CAE in published scholarship on evaluation.

2.1 The concept of empowerment

Empowerment is a moving target. While aid agencies strive to empower beneficiaries of programs, they also strive to find better ways to measure and report on empowerment. In addition, empowerment is a contested concept- what it is, who defines it and who benefits are just a few of the questions that pepper the empowerment landscape.

Leading organizations in international aid have developed various rubrics to understand empowerment. Several rubrics have been developed to encourage the collection of evidence that demonstrates empowerment in concrete, tangible ways. Despite the fact that the empowerment lexicon has grown in increasing popularity, there is still no consensus on its definition. Most definitions seem to agree that empowerment requires that the traditionally marginalized groups gain access to decision-making powers, but the degree to which this ranges from radical state transformation to capacity-building is unclear (Jupp, Ali & Barahona, 2010). Furthermore, empowerment is sometimes defined as a process, and at others times as an outcome. Sometimes empowerment includes outsiders, and sometimes it excludes them.

Just as the development community has strived to attain and measure empowerment, the evaluation community is also working towards greater refinement in our own evaluations, not only to assess levels of empowerment experienced through

programming, but also to understand the impact of our own transformative practices. Empowerment can be seen to happen at two levels, as discussed briefly in the introduction: at an individual level, and at an institutional level. There are also different outcomes of empowerment, such as social empowerment, economic empowerment, political empowerment, and competency empowerment. Empowerment is also value-driven and is conceived of differently depending on social and cultural background. That being said, it is easy to understand a positive empowerment process and a negative process of disempowerment (Jupp et al., 2010).

Taylor (2000) tells us that the art of measuring someone else's empowerment is potentially disempowering and goes on to state the measurement of empowerment must not be allowed to become something that the more powerful do to the less powerful. Mayoux (2008) notes that 'the selection of any particular set of indicators... is inevitably based on underlying theoretical, and often political, understanding of what types of impacts are important'. For these reasons, self-reporting or appraisal is the best method to gather, monitor and assess data on empowerment.

As defined in the introduction, transformative CAE intend to produce empowerment outcomes. Therefore, understanding how empowerment happens and how we measure that outcome is an important task for evaluators embarking on this journey. Later I provide a brief history of transformative CAE and current critiques in the field.

2.2 International development evaluation

Notwithstanding its critics, participatory methods have gained prominence in development planning and programming (Cornwall, 2003). This process has gone beyond simply garnering the involvement of beneficiaries, but has led to addressing questions of citizenship, rights and governance.

Over several decades, development work had come under serious scrutiny over the quality of aid and its impact on the end-goals of development. In response, the Paris Declaration on Aid Effectiveness, endorsed in 2005, committed aid donors and recipients to five principles: ownership, alignment, harmonization, results and mutual accountability that were intended to reinvigorate development and reposition it in a place that was more

respectful of recipient nations and met the accountability needs of donor nations (Dabelstein & Patton, 2013). The Paris Declaration provided an action-oriented roadmap coupled with 56 commitment targets that were to be met by 2010. It represented an unprecedented commitment to do aid differently as well as a systematic way to ensure accountability on all sides.

The Paris Declaration ultimately impacted on the evaluation community as well, particularly in donor nation aid delivery organizations, such as the (former) Canadian International Development Agency (CIDA) (Essex, 2012). Aid effectiveness policies and strategies were implemented through an increase in efforts of the agency to focus their work and increase their reporting on results and transparency, in order to maximize accountability. Principles of ownership and alignment increased pressure for donor nations to charge recipient nations with the task of planning, implementing and evaluating their own development projects. Thus, a need for local individual and organizational evaluation capacity was created.

Criticisms of donor-controlled development are still being voiced today (Carden, 1997, 2010; Hay, 2010; Segone, 2006; 2008). In fact, the recent evaluation of the Paris Declaration shows that “improvements are slow and uneven in most developing countries and especially among donor agencies.” (Wood & Betts, 2012, p.103). Many evaluation scholars who work in the development sector continue to press for advancements in capacity building and view collaborative approaches to evaluation as a means to achieving this end (e.g.: Guijt, Eyben, Shutt, & Roche).

2.3 Collaborative approaches to evaluation

Aspects of CAE first appeared on the evaluation scene in the 1960s and 70s (Brisolara, 1998). As mentioned above, CAE, as defined by Cousins and Whitmore (1998) implies collaboration among stakeholders and evaluators in the production of evaluation knowledge. This is to say, the knowledge produced through the evaluation is co-created by participants. CAE is relatively new in the field of professional and academic evaluation; however, participatory methods have a rich history in the practice of social research. It was not until the 1970s that government and non-governmental agencies and organizations (NGOs), focusing on development and social betterment,

began to pay closer attention to approaches to programming that aimed to increase the acknowledgement of traditionally unacknowledged voices (Guijit & Shah, 1998).

In their seminal paper, *Framing Participatory Evaluation*, Cousins and Whitmore (1998) “explored the meanings of collaborative approaches to evaluation through the identification and explication of key conceptual dimensions” (p.87). According to the authors, engaging in PE requires the evaluation team to collaborate in some capacity with individuals, groups or communities that have been identified as stakeholders.

Stakeholders can be briefly described as individuals or groups of people that have a vested interest in the program that is being evaluated, or in the evaluation findings themselves. As noted by Cousins and Chouinard (2012), “it is a dialogue between these different groups that define the parameters of participatory practice and the knowledge that is ultimately constructed” (p.4).

Although the field of CAE is quite heterogeneous, Cousins and Whitmore (1998) focusing on PE have categorized approaches under two broad streams: P-PE and T-PE. The “core premise of P-PE is that stakeholder participation in evaluation will enhance evaluation relevance, ownership, and thus utilization.” (Cousins & Whitmore, 1998, p.88).

T-PE, as described by Cousins and Whitmore (1998), “invokes participatory principles and actions in order to democratize social change” (p. 89-90). According to the authors, the most fundamental issue in T-PE is the concept of creation and control of the production of knowledge. Based on the seminal work of Brazilian adult educator, Paulo Friere (1970), T-PE aims to empower people through their participation in the process of constructing and deconstructing their own knowledge, therefore understanding how their lived experiences connect with knowledge, power and control. Their second key concept concerns the process through which evaluation takes place. In T-PE, typically, it is expected that the researcher/ evaluator and stakeholders are all contributing collectively, initiating a dialogue that, according to Gaventa (1993), leads to mutual respect that can then facilitate the conditions through which empowerment can take place (Cousins & Whitmore, 1998). Finally, Cousins and Whitmore (1998) provide a closing critical reflection, which requires that all participants interrogate their lived realities, including

their own biases and assumptions.

Cousins et al. (2013) argue that it is through dialogue and deliberation between evaluators and various stakeholders that the context within which the evaluation will take place are understood and a collaborative approach can be deemed appropriate, or not. While P-PE may be motivated by demands for accountability and learning for improvement, T-PE is motivated by the need for capacity building through developing an evaluative culture, questioning assumptions and challenging current perceptions (Cousins et al., 2013).

Today, a variety of approaches are situated under the umbrella of CAE. Cousins and Whitmore (1998) established three dimensions for locating collaborative practice: control over technical decisions, stakeholder selection for participation, and depth of participation of stakeholders. This three dimensional framework was subsequently elaborated into five dimensions (Weaver & Cousins, 2004), but Daigneault and Jacob (2009) concluded after a thorough analysis that the initial three dimensional framework is the most comprehensive and parsimonious way to identify CAE. Cousins and Whitmore differentiated among forms of CAE using this framework associating them with the two strands of PE. Under the transformative strand, they included T-PE (Tandon & Fernandes, 1982, 1984; Fals-Borda, 1980; Gaventa, 1993), Democratic Evaluation (Macdonald, 1976; McTaggart, 1991), Empowerment Evaluation (Fetterman, 1994; 2011), and emancipatory action research (Carr & Kemmis, 1992; McTaggart, 1991). Today, CAE continues to proliferate, and critical debates in the evaluation and development community continue to enrich our understanding of these approaches (see e.g., Cousins, Whitmore & Shulha, 2013).

2.4 Critiques of transformative CAE

According to Miller and Campbell (2005), systematic evidence on evaluation practice is necessary for the further development of evaluation theory. The authors completed a systematic analysis of 47 case examples of empowerment evaluation to build upon the theory of empowerment evaluation, contributing a body of empirically derived knowledge designed to address the key points of disagreement in empowerment evaluation debates. This article is in response to the criticisms brought forth in the

evaluation community since 1994 over the lack of a systematic evaluation of the empowerment evaluation process and its consequences.

From the field of developmental evaluation, and more specifically empowerment evaluation, critiques have primarily concerned themselves with a lack of clarity of methods (Miller & Campbell, 2006). Many scholars (e.g., Cousins, 2005; Miller & Campbell, 2006; Patton, 2005; N. L. Smith, 1999) state that empowerment evaluation specifically is not very different from the variety of approaches that use the same lexicon and are rooted in social justice, and in practice, not very different from P-PE (Cousins, 2005; Miller & Campbell, 2006).

Further, many have identified a lack of specified mechanisms of change that leads to transformative outcomes, either at the individual or institutional level (Miller & Campbell, 2006; Patton, 2005; N. L. Smith, 1999; Worthington, 1999). This theoretical ambiguity and lack of unanimity in practice is important to mention as it has bearing on the focus of this study, which builds on the evidence presented by Miller and Campbell (2006, 2007) demonstrating that reported transformative outcomes more closely resemble outcomes related to more practical ends, such as evaluation use.

There appears to be an unclear link between transformative approaches and the attainment of empowerment outcomes, leaving the role of the evaluator and the specific context in which these approaches would actually lead to empowerment outcomes unknown. This lack of clarity leaves room for the criticism that of the adoption of the lexicon of empowerment without actually utilizing any empowering processes, which was suggested by McKie (2003) in her article concerning the mainstreaming of CAE. Building on this criticism, there is also a lack of evidence to show that utilizing CAE lead to empowerment, or social justice outcomes (Cousins, 2005; Miller & Campbell, 2006; N.L. Smith, 1999).

Miller and Campbell came to this statement by coding their cases of reported successes in empowerment evaluation two ways: by reported outcomes and by adherence to the principles of empowerment evaluation articulated by Fetterman and Wandersman (2005). As outlined by Miller and Campbell (2006), many of the transformative outcomes reported in their study are actually closely aligned with practical outcomes related to

evaluation utilization as opposed to transformative outcomes related to social justice. This study also intends to add to this discussion through exploring how transformative outcomes are operationalized in addition to interrogating who defines the term “transformative outcomes”. In transformative paradigms, attention must be paid to the researcher’s or evaluator’s relationship with the process of defining and creating knowledge (Reinharz, 1992). What is and is not transformative must be viewed through the lens of the participant and their lived experiences, not from an objective standpoint.

Miller and Campbell (2006) also found that nearly all authors reported what they believed to be indicators of success in empowerment evaluation with only four reporting failure in total or in part. In only seven of those successful cases were indicators of success empirically verified through systematic meta-evaluations. Thirty-nine cases reported only anecdotal impressions of the authors or others involved in the evaluation.

This finding strikes at the heart of why evaluators choose to engage in empowering processes. Miller and Campbell identify anecdotal impressions as inherently problematic, empirical evidence that is not valid or less reliable. I argue, however, that their terms of reliability and validity should expand to encompass the values of empowerment and transformational research, which do not view anecdotal evidence in a strictly negative light. There is a conflict of epistemology here that needs to be addressed. In an attempt to define the validity of evidence, I will be looking for evidence through the lens of post-positivist, emancipatory research- construct validity, face validity and catalytic validity (Lincoln & Guba, 1986). Construct validity, defined by Lather (1986) as a systematized reflexivity to ensure that observations represent the construct being investigated. Face validity refers to what Lincoln and Guba (1981) term “member checks”, an act that takes analysis back to at least a subsample of respondents, and catalytic validity refers to the “degree to which the research process re-orientes, focuses, and energizes participating in knowing reality in order to better transform it” (Lather, 1986, p.67).

Miller and Campbell coded these indicators of success in empowerment evaluation under a number of categories, ranging from data use and knowledge gained to accountability and sense of ownership. It is argued that the majority of these indicators

relate to evaluation use and process use, as opposed to the principals of empowerment evaluations espoused by Fetterman and Wandersman (2005). For example, the most common principles evident in case descriptions were community knowledge, organizational learning, and accountability. Ultimately, Miller and Campbell found that evaluations described as empowerment evaluations frequently did not embody the core principles of democracy and social justice (2006). Further, these outcomes were mostly attributed to program staff as opposed to a diversity of stakeholders as proposed by Fetterman.

Fetterman and Wandersman responded to many of these criticisms most recently in their 2007 article, positing that most issues that were raised by critics over conceptual ambiguity, methodology and validity have been addressed at length in previous volumes and panel discussions. Further, they believe that most of these critiques belong in empowerment evaluation's past, rather than its future. In my opinion, their response did little to actually address the continued efforts of the evaluation community to push empowerment evaluation forward, theoretically and practically, through empiricism. The article reads as a defensive, if not frantic response that did little to address their critics. They raised distinct criticisms of Miller and Campbell's work- specifically, their limited methodology, which did not include evaluation reports as valid cases.

Upon reflection, I have come to question this criticism. Evaluation reports, to the best of my knowledge, report on program outcomes as opposed to evaluation outcomes. Therefore, these reports would likely contain little information on empowerment outcomes related to the evaluation approach. Further, it is my understanding that evaluation reports are often delivered in a timely manner; usually within months after data collection takes place. Finally, I understand that for empowerment outcomes to have taken place, a certain amount of time would likely need to lapse before a researcher or evaluator could return to the site to fully grasp the transformative outcomes that have manifested even in part due to the use of transformative evaluation approaches. Therefore, I will not be including evaluation reports in my own sample of studies and instead will focus on empirical research on transformative evaluation.

It is also important to note that the majority of the criticisms levied by the

evaluation community point specifically to empowerment evaluation, and do not expand beyond this approach which does little to push transformative evaluation theory and practice forward. The present study expands our critical understanding of all CAE that specifically intend to promote transformative outcomes. Fitzpatrick calls for an end to theory debates that speak to the need to further distinguish between collaborative approaches to evaluation, and I agree with her. While the majority of these debates concern what Fitzpatrick (2012) refers to as the “nuts and bolts” and “nuanced bits” of collaborative evaluation, criticisms of empowerment evaluation or transformative evaluation that include a careful examination of power and control are not highlighted as frequently in the evaluation community. I believe that evaluators can do more to draw these issues to the forefront of our theory development and practice, and can do so by taking up the criticisms of participatory program planning.

Participatory planning approaches that have their roots in empowerment have come under fire for taking on the form of increasingly technical, management solutions to issues that are characteristically political (Cornwall, 2003; Gregory, 2000). These technical, quantitative approaches to evaluation accountability often side step or wholly ignore the politics of communities and the micro-politics of difference- gender, social economic status, sexual orientation, ability, and so forth. Even as more qualitative and quantitative impact assessment models are being used in the field, concerns with evaluation efficiency and economy often trump concerns related to empowerment.

Carden (2007, 2010) refers to the “real evaluation gap”, which is founded on the real lack of support for the development of an indigenous evaluation profession in developing countries that has the potential to not only infuse evaluation with new knowledge but provide balance between donor and recipient interests in evaluation. In his opinion, this requires involving citizens, researchers and evaluation professionals in building “indigenous evaluation cultures and capabilities to contribute to improved decision-making” (p.220). According to Carden, without this development donor control of the design and conduct of evaluations ensures that donor needs will continue to be put before recipients’ (including the host nation, NGO and program beneficiaries).

Framing evaluation within the above commentary, I see evaluation as a political

act. This viewpoint aligns with those of prominent evaluation theorists such as Greene (1997) and House and Howe (1999). Furthermore, in congruence with Greene (1997), it is my opinion that engaging in participatory methods is always about more than practical ends, and that every evaluation is driven by political, philosophic and pragmatic rationales. I align myself with Cornwall and Brock (2005), who state that commitment to participatory practices without awareness and attention to the politics of engagement can be damaging for the communities involved, the programs being evaluated and the field of evaluation at large. The politics of engagement, in this context, refers to the fact that evaluation findings make statements about the value and merit of a program, and the funding agent or other invested stakeholders who have vested interests in the program's performance often sponsor these evaluations.

Therefore, the contexts in which evaluations take place are inherently political. Political, in this context, refers to various questions posited by Greene (2009) that question the role of evaluation, and our position as evaluators- what purposes and whose interests does social inquiry serve? What values are advanced or repressed by this approach to inquiry? Failing to ask these questions masks critical questions relating to values, commitments, priorities and political control, power and agency.

Gregory (2000) argues that active participation, "while great in theory to ensure that power is balanced and people feel ownership, [effective participation] is not easily achieved"(p. 181). In other words, ensuring global participation does not guarantee equity. In fact, Gregory argues that enforcing full participation can contribute to inequality. There are many internal and external factors that inhibit or facilitate participation, including our own definitions of participation and evaluation. Furthermore, participation does not naturally mediate power dynamics and can also serve to facilitate their fortification if cultural norms are not fully investigated and addressed. As a community, evaluators must continue to problematize our ideas of how and why participation takes place.

The presence of dialogue between evaluators and participants is often viewed as a way to mitigate these power dynamics. As Mertens (2009) argues, the rationale for transformative paradigms hinges on utilizing community involvement and research

methodologies that acknowledge issues of power and discrimination. This involvement is required throughout the evaluation process, and especially when evaluation results are being communicated to those and power is being actively redistributed. Evidence of these sorts of processes that are deliberate efforts by the evaluator to position evaluation as a participant-oriented process, are not paid attention in Miller and Campbell's analysis.

Green (1998) furthers this point in his examination of constructs of participation as practice. He contends "any technique, approach or structure tends to have different meanings in differing geographical, cultural and temporal contexts" (p.74). As argued by Mathison (2009), procedures and techniques that are used in collaborative approaches are inherently Western-driven and are applied generically in cross-cultural contexts. For example, she cites the United Nation's development and evaluation of the Millennium Development Goals, specifically aligned to educational outcomes. Therefore, the outcomes of evaluations, and particularly those that are geared towards transformative ends as outlined above, must also be derived from Western ideologies, and may not be entirely appropriate if not identified and defined by stakeholders through a collaborative process. Evidence of attention paid to addressing cross-cultural issues in collaborative evaluation must also be investigated.

To date, the role of the evaluator in facilitating transformative evaluations has not been adequately examined through the lens of power. While Smith (1999) asks whether the role of the empowerment evaluator is to empower others or advocate for people or causes, or whether they are inherently self-serving (biased) (Cousins, 2005), little attention has been paid to the evaluator's role in disrupting power as an indicator of empowerment processes. Drawing on the work currently available on these themes (Chouinard & Cousins, 2015), this research will look for evidence of a shift in power that addresses the inherent imbalance between evaluator and participant.

Summary

This qualitative study aims to develop our understanding of transformational outcomes as achieved in development evaluations utilizing transformative collaborative approaches. Through providing a review and integration of the body of empirical research literature, this asks how we measure current indicators, the extent to which they

demonstrate empowerment principles and finally, which processes contribute to or detract from the attainment of these outcomes. The findings and conclusions of this study can inform future transformative evaluation practice and theory in international development setting.

3. Theoretical Orientation

The research design for this thesis draws from a social constructivist paradigm. In this worldview, individuals seek to understand the world in which they live by developing subjective meanings of these experiences (Creswell, 2007). Because these views are subjective, this paradigm implies the rejection of a singular, universal truth, but rather seeks to understand the complexity of different interpretations. Validity continues to be a central concept in social science methodology, particularly in terms of qualitative research, referring to the truth and correctness of a statement (Kvale, 1995). In qualitative research, this often means asking to what extent are our observations reflective of the phenomena under investigation. Therefore, my critical perspective on development and evaluation has influenced the knowledge created within this study.

In contrast to the criticism expressed by Miller and Campbell (2006), anecdotal evidence brought to bear by participants is seen as both reliable and valid. Through expanding the concept of knowledge as a mirror of reality to knowledge as a social construction, anecdotal evidence creates space to acknowledge all voices as co-creators of truth.

Research conducted with this understanding often begins with broad, general questions designed to explore and construct meaning. However, in this study I was also confined by pragmatist desires to move forward in the face of real-time constraints, modifying the vastly open-ended nature of an interpretivist approach and implementing a positivist approach when designing a structure to organize my data at the very beginning based on prior research conducted in the field.

4. Research Questions

While the theory of collaborative approaches to transformative evaluation has seen fair treatment in the evaluation community, there remains little empirical evidence

supporting the realization and operationalization of transformative outcomes, as outlined above. This study will endeavour to further understand whether claims of transformative outcomes are valid in the empirical literature available. Building on current empirical literature, notably Miller and Campbell (2006), and borrowing from the methods of Amo and Cousins (2007), this review and integration of the extant empirical literature will pay special attention to the social construction of validity and co-creation of knowledge in the reporting of transformative outcomes. It will also include an analysis of the integration of dialogue about the purpose and value of evaluation and participation in evaluation, using the concepts of validity espoused by Lather (1986)- construct validity, face validity, and catalytic validity. This research will employ a qualitative research methodology to gain insight into transformative evaluations that take place in development settings.

It will address the following questions:

1. How are transformative outcomes operationalized and measured in transformative evaluations that take place in development settings in the current research on evaluation literature?
2. To what extent are transformative outcomes demonstrating empowerment principles such as ownership, inclusion, democracy, and social justice being realized in transformative collaborative evaluation approaches in development settings?
3. What factors and conditions have contributed to the attainment of these outcomes? What factors and conditions detract from the attainment of these outcomes?

This qualitative study, a form of meta-analysis, will further develop our understanding of transformational outcomes as achieved in development evaluations utilizing collaborative, transformative approaches. By building on the literature available, the research investigates current criticisms of collaborative evaluation by taking a critical look at the processes engaged when evaluators utilize these approaches, as well as attempt to identify which processes in particular lead to evidence of transformative outcomes. Further, this study intends to understand who identifies and defines what evidence constitutes transformative outcomes.

5. Methods

The following section provides an outline of the methodology undertaken to produce the findings and conclusions represented in this study. This research will employ a qualitative research methodology to explore the operationalization of transformative outcomes in the empirical literature of transformative evaluations that take place in development settings.

5.1 Sample Development

This thesis is based on a review of current empirical literature concerning the attainment of empowerment outcomes of evaluation in development settings. Cases (relevant articles and book chapters) were derived from utilizing keyword searches in online databases between 1998 and 2014, including Scholars Portal, EBSCO in addition to Google Scholar. Drawing from Miller and Campbell (2006) the following keywords were included: *empowerment evaluation, empowerment adj evaluation, collaborative approaches to evaluation, transformative evaluation, getting to outcomes, evaluation capacity, process use, transformative outcomes, empowerment outcomes, participatory evaluation.*

I chose the timeframe for the sample as beginning in 1998 as it coincides with Cousins and Whitmore's differentiation between practical participatory evaluation and transformative participatory evaluation. The year 2014 was chosen as the end date to ensure that the most up-to-date cases could be included, building on previous work completed in recent years. The sample for this study was finalized in March 2014.

I will search for the appearance of these terms in the abstracts, key words and bodies of articles and book chapters. In response to Fetterman's (2007) critique of Miller and Campbell, I considered including evaluation reports. This is an important distinction. Evaluations report on evaluation questions- namely the quality and worth of a program, policy or practice. However, research on evaluation examines the social practice of evaluation (Greene, 2013). As previously mentioned, I decided to continue to exclude evaluation reports, as they would not contain information regarding transformative outcomes as a result of the evaluation, but rather about empowerment outcomes achieved through the program. Linking these outcomes to the evaluation, in my opinion, would be

erroneous.

Reflective case narratives are admissible in this sample. While reflective case narratives do not include a methods section, I would concur with Cousins and Chouinard (2012) that they are 'empirical' by virtue of their being based on real experience and observation. As we will see, the majority of the research on evaluation available in this domain takes on the form of reflective case narratives. These cases are retrospective and provide the evaluation story as told from the perspective of the author. The authors in the majority of cases in this study are the lead evaluator or evaluation team, with a small selection co-authored by other evaluation stakeholders (e.g., members of the program community). This raises specific questions about authorship, voice and power, which are addressed in the discussion section of this study, which concur with Cousins and Chouinard (2012).

Building on the review completed by Miller and Campbell (2006), I searched the content of specific journals such as the *American Journal of Evaluation* [formerly *Evaluation Practice*], the *Canadian Journal of Evaluation*, *Educational Evaluation and Policy Analysis*, *Evaluation*, *Evaluation and Program Planning*, *Evaluation and the Health Professions*, *Evaluation Review*, the *Journal of Multidisciplinary Evaluation*, the *Journal of Multidimensional Evaluation*, *Development Evaluation*, *Canadian Journal of Development Studies*, *Development in Practice*, *Development and Change*, *Third World Quarterly*, *Action Research*, and *Community Development Journal*.

The following paragraph will outline the various selection criteria. Included studies are empirical in nature and must have been subjected to a peer review process or reviewed by an editorial board (book chapters, theses, for instance). Following this initial stage, the bibliographies of each identified article were examined to identify further sources. Sixty-five cases were initially included at this stage that outlined collaborative approaches to evaluation. These cases were then reviewed to ensure that the case took place in an international development setting and had explicitly stated that the evaluation undertaken and explicitly state that the evaluations conducted were intended to promote transformative outcomes. The second factor resulted in cutting the down the population size significantly. At the end of this stage, 24 cases remained. From this stage, cases were

categorized using an emergent theme coding strategy, loosely structured by the research questions. Table 1 provides a description of the final sample of studies according to context, design and methods and research objectives.

5.2 Sample description

Twelve of the 24 articles were published within the last 10 years, with only five published in the last five years. All cases took place in an international development setting, with the majority of them in Africa. Other projects took place in South America, North America (Mexico), and Asia. All cases were reflective case narratives, centred on the perspectives of the authors. These are classified as reflective case narratives because they contain no methods section detailing how conclusions (the evidence of empowerment outcomes as a result of evaluations) were drawn. Finally, the purpose of the majority of the cases was to illustrate a case example of participatory evaluation in development settings. The majority of cases reviewed single sites, and presented the description of the case and the evaluation process that was undertaken, while others served to provide illustrative examples of new methodologies undertaken, with guidance for future implementation. No studies employed a quantitative or mixed-methods approach.

Table 1. Study Sample

Study	Context	Design/Methods	Research Objectives
Newman, K. (2008)	Nigeria, multiple locations	Retrospective: Reflective case narrative provides details on the evaluation rationale, process used and evaluator reflections on the process used in single evaluation.	To explore the tensions between the desire to be participatory and conflicting needs of donor communities (outcomes vs. process). Places emphasis on understanding power relations in PE in relation to social justice/transformation. Pays explicit attention to open and honest dialogue.
Hart, R. A., and Rajbhandary, J. (2003)	Nepal, multiple sites	Retrospective: Reflective case narrative provides details on the evaluation rationale, process used and evaluator reflections on the process used in single evaluation	describes and evaluates participatory methods that were developed to enable children in Nepal to document and reflect on the democratic functioning of their own organizations. It concludes with an assessment of the potentials of participatory methods for self- monitoring by democratic organizations of young people.

Study	Context	Design/Methods	Research Objectives
Holte-McKenzie et al. (2006)	Kenya	Retrospective: Reflective case narrative provides details on the evaluation rationale, process used and evaluator reflections on the process used in single evaluation.	The discussion (Section 6) interprets these results in the light of the evaluation design and reflects on two questions: Did the process achieve participation? And did the process achieve empowerment? Lessons learned during the study are shared (Section 7), followed by the conclusion (Section 8).
Parkinson, S. (2009)	Uganda	Retrospective: PhD research completed as case study of specific project.	Examines the assumptions that underpin the design of participatory monitoring and evaluation using a case study of a rural development programme. Discusses mismatch between participant and donor expectations and assumptions, and discusses how heightened awareness of power dynamics and political factors are needed for the effective design of PM and E
Bi Niba, M and Green, J. M. (2005)	South Africa, various sites	Retrospective: 5 projects were studied that had completed participatory evaluations. Projects had similar aims and methods for attaining them, but demographically were different.	This article aimed at finding out if participatory processes (group discussions, enactments, and others) do make a valuable contribution in communication-based project implementation/ evaluation and the fight against HIV/AIDS. To identify values, the state of beneficiaries prior to and after project implementation/evaluation was compared.
Botcheva, L., et al. (2009)	Zimbabwe	Retrospective: Reflective case narrative provides details on the evaluation rationale, process used and evaluator reflections on the process used in single evaluation.	This paper describes a process-oriented approach to culturally competent evaluation, focusing on a case study of an evaluation of an HIV/AIDS educational program in Bulawayo, Zimbabwe.
Abes, R.V. (2000)	Philippines	Retrospective: Reflective case narrative provides details on the evaluation rationale, process used and evaluator reflections on the process used in single evaluation.	Chapter describes the initial experiences of development of a 3 year participatory impact evaluation with a focus on the first year.
Alzate, R.D.E. (2000)	Colombia	Retrospective: Reflective case narrative provides details on the evaluation rationale, process used and evaluator reflections on the process used in single context.	Chapter summaries the experience of the organization in designing and implementing and PM and E system
Blauert, J and Quintanar, E (2000)	Mexico	Retrospective: Reflective case narrative provides details on the evaluation rationale, process used and evaluator reflections on the process used in single context.	The paper addresses the problematic issues that arise with regard to participation and team dynamics in a sustainable agriculture program, focusing on the application of PAR methods.
Hamilton, C., et al. (2000)	Nepal	Retrospective: Reflective case narrative provides details on the evaluation rationale, process used and evaluator reflections on the process used in single context.	Review of the process of developing and implementing a locally relevant monitoring and evaluation system in rural Nepal with Forest User Groups.

Study	Context	Design/Methods	Research Objectives
Symes, J and Jasser, S. (2000)	Palestine	Retrospective: Reflective case narrative provides details on the evaluation rationale, process used and evaluator reflections on the process used in single context.	Review of the process of implementing a monitoring and evaluation system in rural Palestine for a farming extensionist program.
Lawrence, et al. (2000)	Two cases, Bolivia and Laos	Retrospective: provides details on the evaluation rationale, process used and evaluator reflections on the process used in single evaluation at two sites.	Provides an overview of the challenges in implementing a monitoring and evaluation system for impacts in new technology in two case sites.
Torres D, V.H (2000)	Ecuador	Retrospective: Reflective case narrative provides details on the evaluation rationale, process used and evaluator reflections on the process used in single context.	A review of the implementation of a culturally relevant monitoring and evaluation system in rural municipalities in Ecuador.
Motteaux, N, et al. (1999)	South Africa	Retrospective: Reflective case narrative provides details on the evaluation rationale, process used and evaluator reflections on the process used in single context.	Looking specifically at one program site, the paper provides opportunities for a range of participatory methodologies which aim to empower and provide an agenda for development
Vernooy, R., et al. (2006)	China	Retrospective: Reflective case narrative provides details on the evaluation rationale, process used and evaluator reflections on the process used in single context.	Examines the capacity building experiences of two teams in southwest China who used PM & E to strengthen development research. Aims to add to the scarce amount of literature available on the subject of community based resource management and PM & E.
Donnelly, J. (2010)	Melanesia	Retrospective: Reflective case narrative provides details on the evaluation rationale, process used and evaluator reflections on the process of using Appreciative Inquiry (AI) in single context.	Provides insights as to how AI approaches to evaluation can be used to empower program staff and beneficiaries and lead to more sustainable outcomes. Provides a case example where AI has been applied.
Brandao, D.B., et al.(2012)	Brazil	Retrospective: Reflective case narrative provides details on the evaluation rationale, process used and evaluator reflections on the process used in single context.	Chapter reports on the involvement of crime-involved youngsters in a participatory evaluation process in Brazil. Provides insights into a new methodology used, FRAMES, to elicit participation of youth.
Letichevsky, A.C., and Penna Firme, T. (2012)	Brazil	Retrospective: Reflective case narrative provides details on the evaluation rationale, process used and evaluator reflections on the process used in single context.	Presentation and discussion of an evaluation case from a program in a Brazilian slum with the intention of sharing the model with other organizations evaluating similar social programs with youth at risk.
Kariuki, J., and Njuki, J. (2013)	Kenya	Retrospective: Reflective case narrative provides details on the evaluation rationale, process used and evaluator reflections on the process used in single context.	Provides a case example highlighting the methodology of Participatory Impact Diagrams, used in a Kenyan development project. Provides recommendations to enhance its applicability.

Study	Context	Design/Methods	Research Objectives
Cornwall, A. (2014)	East Africa	Retrospective: Reflective case narrative provides details on the evaluation rationale, process used and evaluator reflections on the process used in single context.	A review of the use of participatory process evaluation in a development setting. Building on her experience, the author explores the use of visualization methods as empowering methods.
Brandon, P.R., et al. (2014)	South Africa	Retrospective: Reflective case narrative provides details on the evaluation rationale, process used and evaluator reflections on the process used in single context. Authors were interviewed by the journal to provide an exemplar case narrative.	A reflective review of a multi-year on-going monitoring and evaluation system in an international development program.
Ward, P. (2000)	Zambia	Retrospective: Reflective case narrative provides details on the evaluation rationale, process used and evaluator reflections on the process used in single context.	Chapter outlines the process that the organization underwent to achieve a shift towards a more learning-oriented and participatory approach to development.
Mompoti, T, and Prinsen, G. (2000)	Botswana	Retrospective: Reflective case narrative provides details on the evaluation rationale, process used and evaluator reflections on the process used in single context.	Discusses the understanding and practice of participatory development and the role that ethnicity can play in the involvement of different ethnic communities in development planning and decision making processes.
Dageid, W., and Duckert, F. (2007)	South Africa	Retrospective: Reflective case narrative provides details on the evaluation rationale, process used and evaluator reflections on the process used in single context.	Describes the process of evaluating a capacity-building initiative for HIV positive South Africans.

5.3 Data Analysis

I analyzed the data in NVivo software. Articles were input into the software as PDF sources and codes were applied directly to the articles themselves. The original framework for the analysis consisted of three parent nodes: Pre-existing Conditions, Mediating Activities and Outcomes. Pre-existing conditions include the community and program context as well as the role and location of the evaluator. The mediating activities section is comprised of the actual practice of transformative evaluation as recorded in the cases. These include the methods used, loosely connected to the fundamental dimensions of participatory practice as defined by Cousins and Whitmore (1998): Control of the evaluation process, stakeholder diversity, and depth of participation of stakeholders. Outcomes refer to the consequences of the evaluation process. I then coded each category

through an emergent-theme coding strategy. As coding proceeded and writing began, I integrated and separated various nodes as the data took shape.

6. Empowerment Outcomes in Transformation Approaches to Collaborative Evaluations

The following sections provide answers to the research questions presented in the introduction. First, the evidence collected to describe how transformative outcomes have been operationalized and measured are presented, followed by the evidence collected to further understand what empowerment principles are present in the identified outcomes, ending with a discussion on what processes have contributed and detracted from the attainment of transformative outcomes in the cases studied. Each section is followed by a brief conclusion as well as an introduction to the emergent concerns that I examine in the discussion chapter.

6.1 Operationalization of transformative outcomes

This section will answer the first research question: ‘How are transformative outcomes operationalized and measured in transformative evaluations that take place in development settings in the current research on evaluation literature?’.

Once a definition of empowerment exists, research requires that it must be operationalized through evidence. Six of 24 articles presented quotations from beneficiaries as evidence. Self-report is one of the best ways to understand empowerment; therefore, quotations in and of themselves represent good evidence when used properly. For example, Abes (2000) reported on how the methods used opened up new areas for the community members to understand one another and utilized a quotation from a participant: “It is good to know these things. We had no time and opportunity to discuss these in the past because we were busy at work. But now, we are here and have a deeper understanding of our community”.

Holte-McKenzie, Forde and Theobald (2006) admit that they never asked participants if they felt empowered through the process, however, they presented various quotes from their informal evaluation of the process such as: “I wasn’t shy. I had a lot of confidence and I felt very happy. I was able to do something that some girls said I couldn’t so I want to do more FGDs to develop my skill.” (p.373).

The majority of the evidence presented in the cases constituted of anecdotal points described by the authors. For example, one article presented a list of the skills acquired from an evaluation, listing items such as learning via activities, showing sympathy, listening, problem-solving, negotiation and decision-making abilities. As a reader, it is difficult to interpret the presence or increase of empowerment in this population without a more fulsome discussion of the perceived changes.

Eight articles presented case study boxes with detailed instances where changes occurred as a result of empowerment developed through the use of participatory monitoring and evaluation strategies. These cases provide the most valid evidence because they provide rich examples of transformation. For instance, these cases presented a narrative including quotations, which described women's involvement in community decision making before and after engaging in participatory monitoring and evaluation activities. Four other articles presented short lists of outcomes as described by the author of the study. While many anecdotal examples of transformative outcomes were presented in the cases there is a clear lack of precision on the concept of empowerment, who defines it, and how it is measured.

Descriptive words such as “strengthened”, “more” and “additional” were employed to describe the advancement of skills, attitudes, opportunities and behaviours as a result of participating the evaluation process; however the evidence is not clear as to how this was measured or defined. In some cases the evidence was used to infer the possibilities going forward for the beneficiaries. For example, Hart and Rajbhandary (2003) discuss that based on the discussions that took place amongst the children, the methods used could be valuable as on-going tools for decision making within the clubs.

Motteaux, Binns, Etienne and Rowntree (1999) presented the only case that completed a follow-up workshop six months after the participatory process took place. Parkinson's 2009 study examined the assumptions that underpinned the design of participatory monitoring and evaluation programme using a case study of a rural development programme. The remaining cases did not mention specific timelines.

Summary. The majority of cases present a retrospective summary of accounts from the perspective of the author, who was largely the lead evaluator with few exceptions.

Many of these articles were written based on observations made throughout the evaluation process with only a minority of these cases utilizing quotations as evidence; most cases employed descriptive words such as “strengthened”, “more” and “additional”, which leaves the reader questioning how outcomes were measured and defined by the author. Most cases are classified as reflective case narratives and do not include a methods section; further, none of the studies utilized mixed-method or quantitative methods.

The strongest evidence was presented by eight cases, all from the same book, which provided rich, detailed case studies of the outcomes of program monitoring and evaluation processes. A few articles inferred evidence that suggest how the immediately observed outcomes could result in further empowerment.

Overall, the studies provided little context to understand how outcomes were measured, leaving much of the interpretation up to the reader. Moving forward, additional work must be done to further contextualize the outcomes described in research on evaluation to fully understand our impact. I provide a further discussion of these concepts and suggestions for the future in the Discussion chapter.

6.2 Empowerment Principles

This section presents the findings related to my second research question: ‘To what extent are transformative outcomes demonstrating empowerment principles such as ownership, inclusion, democracy, and social justice being realized in transformative collaborative evaluation approaches in development settings?’

As previously discussed in the literature review, the concepts of transformation and empowerment are subjective and value-based in definition; therefore they are generally measured through the observation and measuring of proxies. According to Fetterman and Wandersman (2005), for individuals this can manifest as an increased critical awareness and an ability to analyze life circumstances that one wishes to change, one has the ability to change, and one acts in a way to enact that change. For institutions, transformative outcomes are evidenced by empowered groups engaging in a network with similar people to influence change.

My initial data analysis divided outcomes into two categories: empowerment and

evaluation capacity building. The evidence provided in this section speaks specifically to the data coded under empowerment. While the research initially sought to link outcomes to the empowerment principles of ownership, inclusion, democracy, and social justice, this language was not present in the majority of cases; therefore an emergent coding framework was used. These emergent categories align directly with empowerment principles.

The findings below represent some examples of proxies of empowerment outcomes, as identified through this meta-analysis. It is important to note that transformative outcomes do not occur only as a direct result of evaluation interventions, but rather are the consequence of a variety of factors and conditions, which are further developed in question three.

After the initial classifications of political action, social inclusion and equity, and economic were developed a significant amount of data fell into another broader, unidentified category. These data are comprised of the development of skills in individuals and communities that were not only evaluation capacities but new skills that could act as the foundation for further development that could result in empowerment. In short, these represented access to new opportunities that were not political but were characterized as a result of the learning that took place inside the evaluation. At this point, I returned to the literature to review previous efforts to measure and categorize empowerment outcomes. In my search, I found a framework developed by Jupp et al. (2010) that complemented my emergent themes and used it to supplement my own work. The final broad categories developed for transformative outcomes were: political, social, economic, and capability, as described below:

- *Political*. Group and individual feelings of enhanced power, ability to present their own views and negotiate for their own ends in formal and informal decision-making.
- *Social*. Outcomes to do with mutual support, trust, respect and equity.
- *Economic*. Outcomes concerning access and use of economic resources (savings, technical resources, etc.).
- *Agency*. Outcomes pertaining to a group or individual's own capability and

independence.

Below follows a discussion of the most compelling evidence for each category. It is important to note that these themes have been separated here for clarity, but that they are interrelated and evidence may fall into more than one outcome category.

6.2.1 Political. Political outcomes are defined as group and/or an individual's feelings of enhanced power, ability to present their own views and negotiate for their own ends in formal and informal decision-making.

Some authors state that the instantaneous access to findings prompted participants to assess the results and produced approaches to advocate for changes to improve current interventions in communities. For example, Alzate (2000) employed Participatory Rural Appraisal (PRA) processes that have resulted in the entire community becoming actively involved in the development process. Members are decision makers and work collaboratively towards their goals. At the end of the process, they reportedly held their organizations to account and, at the time of writing, were more aware of how local governments allocate resources, allowing communities to take advantage of emerging opportunities. Further, as communities grew in their decision maker roles, traditional power relationships were disrupted and horizontal linkages were established across the region. Communities also used monitoring and evaluation data to negotiate and establish alliances with the national government and the private sector for their mutual benefit. In addition, they were using the data to advocate for more appropriate policies favouring the interests of Indigenous communities, including healthcare, education, natural resource management and food security. The Torres (2000) case describes similar outcomes.

Further examples of this type of outcome include: the stimulation of community action as a result of PRA activities (Motteaux, Binns, Etienne & Rowntree; 1999), the publication of a book describing the evaluation process and results which incited political debate about youth in conflict with the law in the presence of public authorities (Brandão, Silva & Cudas, 2012), and the creation of a new decentralized organizational structure to better meet the needs of forest users who were traditionally marginalized from full participation (Hart & Rajbhandary, 2003).

The political empowerment outcomes outlined above provide evidence of group

empowerment through increased self-advocacy, increased access to opportunities, and implemented changes that enabled greater participation in program management and activities. Findings in this category align with the principles of ownership, inclusion, democracy and social justice.

The following presents a discussion of social empowerment outcomes.

6.2.2 Social. Social empowerment outcomes consist of increased evidence of community-wide support, trust, respect and equity. Evaluation teams utilized a variety of methods to promote social empowerment outcomes, often providing new avenues for communities to communicate in new ways. As a result, new knowledge of each other was obtained which shifted participant perceptions. For example, through the use of PRA activities, Hart and Rajbhandary (2003) note that the children's evaluation discussion was the first time they had spoken to their issues inclusively and openly, potentially leading the way to a more democratic way of functioning in their organizations. Similarly, collaborative relationships and the sharing of experiences led Bi Niba and Green (2005) to conclude that social empowerment had taken place in program beneficiaries. The process allowed vital program issues to arise, and one participant remarked that the evaluation helped her re-evaluate her stand on HIV/AIDS prevention.

In other cases, increased communication and collaboration led to increased trust (Letichevsky & Penna Firme, 2012; Vernooy, Qiu & Jianchu, 2006). For example, Vernooy et al. (2006) spoke to the added value of the performance monitoring and evaluation work as follows: increased trust in one another and recognition of each other's roles, strengths and weaknesses; increased cooperation among stakeholders (researchers, farmers, government officials); increased farmer participation, particularly women, in the research and change process; and contributed to building a stronger community identity through collective efforts.

Evaluation activities also served to provide space for traditionally marginalized populations to have a voice. In Lawrence, Haylor, Barahona and Meusch (2000), the study found that the PRA process eventually led to staff appreciation of the different views of men and women and that each was equally valid. This led to a valuing of women's knowledge by program staff, which was absent before. Similarly, in Palestine,

interactive workshops were held with women and men separately, and as a result, women were able to meet together and present their perspectives, assert their views, and define their priorities (Symes & Jasser, 2000). Blauert and Quintanar (2000) report similar outcomes in their work with farmer extensionists.

Several cases also reported on how conflict within the evaluation process led to increased trust and cooperation among community beneficiaries and organizations (Alzate, 2000; Hamilton, Rai, Shrestha, Maharjan, Rasaily & Hood, 1998; Torres, 2000). The participatory monitoring and evaluation process used in the Torres case (2000) brought about conflict and resentment, which in turn caused others to work against the process making it difficult for one group to dominate. In some instances, the process weakened the traditional influence and power of politicians and young people are now emerging as important leaders as they take on greater coordination and monitoring and evaluation roles.

In summary, evidence of social empowerment as outlined above includes increased understanding of the community and the diverse needs represented, increased levels of trust, and a strengthening of roles and responsibilities within communities. Findings here align most closely with the empowerment principles of inclusion and democracy.

6.2.3 Economic. Evidence in this category reflects increased access and use of resources, including natural resources. There was little evidence found in the cases that fit this category.

Motteaux et al. (1999), Hamilton et al. (1998) and Vernooy, et al. (2006) infer an increase in access and management of natural resources. The authors speak to the communities' new knowledge about their environment and the resources within, and spoke about how the participatory monitoring and evaluation process helped identify potential solutions to identified problems in resource management. Having an increased understanding of the use and management of state resources permitted them to seize opportunities that could improve their daily lives.

This section provides the least amount of evidence across cases. Evidence suggests that increased access to local resources will improve the daily lives of participants. Outcomes here align closely with the empowerment principle of ownership.

6.2.4 Agency. Evidence in this category refers to an individual or an organization having acquired a new skill, new knowledge, or new awareness. High levels, or advanced stages of agency show evidence of a sense of independence. Low levels, or early stages of agency are evidence of the attainment of new skills, such as learning to collect data (Jupp et al., 2010).

Some of these outcomes related to influencing leadership skills or an increased sense of independence (Brandão et al., 2012; Hamilton et al., 1998; Hart & Rajbhandary, 2003; Holte-McKenzie et al., 2006). While some of the outcomes described in this section seem insignificant, each has the potential for long term empowerment. For example, Holte-McKenzie et al. (2006) discussed how participating “felt good to have an important position, to be looked up to, to prove myself in front of my friends. I want to teach others what I have learned, share the knowledge I have” (p. 373).

New skills were also identified as empowerment outcomes if the skills had the ability to influence leadership (Hamilton et al., 1998; Lawrence et al., 2000; Motteaux et al., 1999). For example, evidence of capability from Motteaux et al. (1999) was the community's assistance in the research process. Children and community members assisted in collecting data about the river, which instilled a new understanding of environmental management. This new knowledge was then applied to a community “environmental day” where learners shared the information with the larger community. Further, taking part in the research activities generated conversations among villagers and visiting scientists, which led to improvements in water management in concrete ways, such as preventing soil erosion and how to detect bad water, and good water, which has had impacts on community health.

Evidence of agency empowerment outcomes ranged from increased self-esteem, and learning new knowledge and skills, which was then applied to participant's daily lived experiences. Empowerment outcomes here do not align explicitly with the empowerment principles but rather act as a foundation to further ownership, inclusion, democracy and social justice.

Summary. While the studies did not reflect the language used in the empowerment evaluation literature, they did provide ample discussion of examples of political, social,

and agency transformative outcomes, while providing limited examples of economic transformative outcomes. Examples ranged from new insights into future community decision making to the creation of policy discussion papers to government lobbying. Importantly, the definition of empowerment outcomes was expanded by this study, including the category of agency. Agency defined outcomes refer to an individual or an organization having acquired a new skill, new knowledge, or new awareness. This takes a step further from evaluation capacity building outcomes, because the new skill knowledge or awareness must be applied to visioning and enacting change in alignment with a desired future outcome.

Much of the evidence presented infers future, sustained empowerment. Considering the majority of cases were reported during or shortly following the completion of the evaluation process, this can perhaps be attributed to the timeframe of reporting. Unfortunately, this significantly limits claims towards sustained empowerment outcomes because there is no longitudinal evidence. In the discussion of transformative outcomes, the timing of reporting has huge impacts on trustworthiness of data, which I speak to in depth in the Discussion section.

6.3 Factors and Conditions that have Contributed or Detracted from the Attainment of Transformative Outcomes

My final research questions focus on factors and conditions observed to have led to empowerment outcomes; specifically, ‘What factors and conditions have contributed to the attainment of these outcomes? What factors and conditions detract from the attainment of these outcomes?’.

I have organized the findings for this section under two main headings: Pre-existing Factors and Conditions, and Mediating Activities. Pre-existing conditions include the community and program context as well as the role and location of the evaluator. The mediating activities section is comprised of the actual practice of transformative evaluation as recorded in the cases. These include the methods used tied into the fundamental dimensions of participatory practice as defined by Cousins and Whitmore (1998): Control of the evaluation process, stakeholder diversity, and depth of

participation of stakeholders. In the interest of readability and efficiency, the following sections contain illustrative examples of the findings.

6.3.1 Pre-existing Factors and Conditions. Pre-existing factors and conditions include the community and program context as well as the role and location of the evaluator. I have found that several of these factors have impacts on the attainment of transformative outcomes.

6.3.1.1 Community and Program Context. This section speaks to the characteristics of the program community found within the cases. Drawing from Cousins and Chouinard (2012), these consist of demographic, social, cultural, economic and historical aspects of the beneficiary community. These may also include local micro-politics in addition to macro-political processes taking place in the national and international context. These factors combined and in isolation may have an effect on collaborative evaluation processes and outcomes. Not all cases provided complete information on the characteristics of the program community and there was great variability in the data available in each case.

Various cases highlighted great times of change in either the national or local context. These changes in the state's organization provided openings in the institutional climate of the nation for increased local ownership over development planning (Alzate, 2000; Hamilton, 2008; Symes & Jasser, 2000; Torres, 2000; Vernooy et al., 2006).

For example, Torres (2000) describes how mass decentralization of the state led to rural parishes' active involvement and inclusion in local economic activities and national resource management through the proliferation of rural organizations and coalitions. The new development of these institutions has demanded the use of participatory methodologies to strengthen local capacity. These include local development committees, rural assemblies, round tables, and indigenous parliaments.

Processes of this kind were reported in conflict environments as well. Symes and Jasser (2000) speak to how the occupation of Palestine greatly influenced the organization's ability to not only operate in the unstable region (moving from emergency relief to community development), but how it influenced their focus on continuing to develop participatory methods and the challenges that they face because of the instability.

The occupation severely limits the control people have over their lives and the environment makes it difficult for people to see beyond the problems they feel powerless to change, an effective barrier to mobilizing and empowering people through collective action and peaceful negotiation. At the same time, the authors describe how the work of the organization is rooted in this struggle and the grassroots organizing that has arisen.

These cases share similar characteristics. There is a collective experience of national macro political processes that have led to the decentralization of power, either formally or informally, which has led to new opportunities for local communities to seize control over resources and means of production. In all of cases exhibiting these attributes, transformative evaluation processes led to empowerment outcomes.

Other cases describe how national and local contexts can potentially restrict program communities and collaborative practice (Brandão et al., 2012; Brandon et al., 2014; Hart et al., 2003; Holte-McKenzie et al., 2006; Kariuki & Njuki, 2013; Lawrence et al., 2000; Letichevsky & Penna Firme, 2012; Mompoti & Prinsen, 2000; Motteaux et al., 1999; Newman, 2008).

Of the ten cases coded with these attributes, 40 percent did not demonstrate the attainment of empowerment outcomes. For example, following a long legacy of colonization, slavery and government instability, Letichevsky and Penna Firme (2012) characterize the last two decades in Brazil as a location for significant efforts in improving education and social programming. However, social inequalities remain and the emphasis on these problems has resulted in a state of social panic and near hopelessness for the urban population. Frequency of violent episodes created areas of high risk. Compounding this unpredictable environment, there is also a law of silence, an attitude adopted out of fear of repercussions if one chooses to speak out of line to the authorities, including evaluators, and also as a means to preserve their identity and dignity (Letichevsky & Penna Firme, 2012).

National and local context appear to be critical factors in enabling empowerment. National political changes resulting in increased access and agency in local communities appear to positively influence empowerment whereas movements towards more restricted agency nationally can negatively impact local empowerment. The following section

outlines how relationships between stakeholders may have impacted the attainment of transformative outcomes.

6.3.1.2 Relationships. Cases described the organizational structure of the program but often did not provide a value assessment of the characteristics of the relationship: supportive, conflictual, or even neutral. In addition, these relationships were described as in a vacuum, or were not considered consequential for reporting. Few cases described the relationship as having interacted with the successes or challenges experienced in facilitating a transformative collaborative evaluation. In fact, twelve articles did not provide any evidence of the existing relationship between the donor organization, the delivery organization and the beneficiaries.

Empowering relationships are described as a relationship where power and context are recognized and the organization collaborates openly with beneficiaries. Beneficiaries have access to decision making, and the organization itself has a commitment towards empowering the local population. Five cases were coded to this node and 100 percent of these cases demonstrated the attainment of empowerment outcomes (Alzate, 2000; Bi Niba & Green, 2005; Blauert & Quintanar, 2000; Symes & Jasser, 2000). Authors describe the relationship between the organization and the population served as a point of access for real political power.

For example, in the case located in Palestine, the context of occupation limited local participation in consultations. At time of writing, Symes and Jasser (2000) report the organization's extensive work to promote participation. They view this as a long-term process and are approaching it gradually, training staff in participatory concepts, developing methods and providing practical experience to staff. Through this process, they have developed a culture of participation that recognizes the wider benefits to participatory decision-making.

Some relationships actually caused conflict in the area due to the opposing values of the evaluation team or program and participant communities. Eighty percent of these cases produced empowerment outcomes, indicating that a conflictual or transitioning relationship does not necessarily detract from the attainment of transformative outcomes (Hamilton et al., 1998; Holte-McKenzie et al., 2006; Newman, 2008; Parkinson, 2009).

The case explored by Holte-McKenzie et al. (2006) provides an interesting perspective. The organization, established in the area in 1999, promotes and teaches football to girls externally but also contributes to those players' life skills and capacity to contribute actively in community development. Throughout the paper, the authors discuss how the organization's "hidden" mission was at odds with local understandings of girls, participation and education; however, the case reported evaluation outcomes.

Based on the evidence, I have determined that relationships are required between beneficiaries, program staff and donors; however, the nature of these relationships does not appear to impact the attainment of empowerment outcomes provided that dialogue is present. The vast majority of cases that provided evidence about relationships reported transformative outcomes. However, evidence was only collected for half of the cases in this study. Further study is required into this factor to fully understand its impact.

6.3.1.3 Evaluation Rationale. In order to be included in this study the evaluations presented in the empirical literature had to specify that the evaluation approach taken was intended to produce empowerment outcomes. However, the underlying rationale to take up evaluation in the different organizations often originates from a multitude of needs.

Evidence of the rationale for evaluating the program was coded into the following emergent categories: accountability, capacity building, research purposes (or as part of a research program), to promote further understanding of a program either by program staff, donors or beneficiaries, or empowerment purposes.

The following table presents a breakdown of the evidence pertaining to evaluation rationale. One study (Dageid & Duckert, 2007) did not disclose a rationale or motivation for the evaluation. It is important to note that many evaluations have various rationales that fell under multiple categories and were therefore cross-coded accordingly, presented in Table 2. A discussion of each category follows the descriptive tables.

Table 2.1. Evaluation Rationales Across Cases

Evaluation Rationale	# of studies
Accountability	12
Capacity Building	15
Research Purposes	5
Promote Understanding	7
Empowerment	10
No stated rationale	1

Table 2.2. Multiple Evaluation Rationales

Evaluation Rationale	Accountability	Capacity Building	Promote Understanding	Empowerment
Accountability	11	5	1	1
Capacity Building	5	15	2	1
Research Purposes	1	2	5	3
Promote Understanding	1	4	1	
Empowerment	1	4	1	10

Accountability. As described in the above table, eleven evaluations were driven by accountability needs, with the majority driven by donor requirements. Most of this work concentrated on reporting the attainment of program objectives and generally speaking to external donors. Several had cross-cutting drivers, with the majority of these being evaluation capacity building (ECB).

Five of the cases studied were driven by accountability and capacity building needs. It should be noted that accountability needs were always mentioned first, and in some cases, capacity building was a secondary goal and seen as a way to lessen organizational reliance on external institutions and evaluators (e.g.: Hamilton et al., 1998; Ward, 2000).

Only one case had both accountability and research drivers. The evaluation sponsors sought to understand how the program was meeting its objectives as well as

contributing to the public debate on the relations between youth, crime and the strategies being employed (Brandão et al., 2012).

Two studies had both accountability and understanding drivers. Again, it should be noted that donor interests trumped educational purposes (Letichevsky, 2012; Symes & Jasser, 2000).

Several evaluations in the study had both accountability and empowerment drivers (Brandon, 2014; Cornwall, 2014; Hamilton et al., 1998; Letichevsky, 2012; Symes & Jasser, 2000). Symes and Jasser (2000) explain that much of the work on planning, monitoring and evaluation systems had been developed by the donor agencies and designed with their own reporting needs in mind. As an extension, much of the incentive for the host organization to enhance their monitoring and evaluation stemmed from their own administrative needs. It was only at the time of the study, and in fact another driver of this study, that they recognized that empowerment is an important aspect of the work they do and that the process can empower participants. Eight out of 11 of these cases demonstrated evidence of transformative outcomes. Three of these cases focused solely on accountability drivers.

Capacity building. Capacity building was the most common rationale driving evaluations in this study. Capacity building referred directly to the organization's ability to do and use evaluations internally.

Nine of the 15 studies coded in this category also explicitly stated that ECB was not only to occur at the organizational level, but at the beneficiary level as well. This is likely due to the participatory and self-evaluation nature of the evaluation approaches selected. While a majority of the ECB coded studies had ECB of beneficiaries as an explicit goal, there were a variety of depths of participation of beneficiaries, which had a direct impact on their ability to develop capacity in this area. For example, in some cases ECB was directed at developing indicators for evaluation and monitoring purposes (Holte-McKenzie et al., 2006), while others intended to build ECB through broader participation in all stages of the evaluation process (Abes, 2000; Blauert & Quintanar, 2000; Hart & Rajbhandary, 2003; Motteaux et al., 1999; Symes & Jasser, 2000; Ward, 2000). Two cases were coded as both capacity building and for research purposes

(Holte-McKenzie et al., 2006; Motteaux, 1999). Four cases were cross-coded as both ECB and to promote understanding and cooperation although these two categories are closely linked in purpose, as further described below.

Only three evaluations were driven by capacity development efforts and empowerment. Letichevsky and Penna Firme (2012) report that the organization wished to promote innovation in evaluation theory, methodology and practice, as well as contribute to the social and economic development of the communities through cooperation aimed at self-determination, social justice, capacity building and transformation. Similarly, Motteaux et al. (1999) reported objectives related to improving the community's ability to review and evaluate its action and impact on the environment and collectively seek solutions. Only one of the 15 studies coded here did not demonstrate transformative outcomes (Cornwall, 2014). Therefore, ECB can be seen to reinforce transformative outcomes.

Research purposes. Several of the evaluations (5) had a research focus. The impetus for this research came from government (Brandão et al., 2012; Mompati & Prinsen, 2000; Motteaux et al., 1999;), or unknown sources (Bi Niba & Green, 2005; Holte-McKenzie et al., 2006). Research was conducted either on evaluation theory and practice or to gather specific information about populations and environments to inform policy. Only in the case presented by Motteaux et al. (1999) were research and empowerment the evaluation drivers. The authors describe, "the aims of the study were for the researcher and community members to identify and evaluate Indigenous perceptions of people-environment relationships in the local area, with a view to generating community action which could lead to an improvement in environmental conditions and community welfare" (1999, p.263). Four of five studies coded here demonstrated evidence of transformative outcomes; therefore, research drivers do not deter from the attainment of transformative outcomes.

Promoting understanding and cooperation. It is important to note that most coded cases referred specifically to their ability to learn from their actions, which is difficult to differentiate from evaluation capacity building. They wanted to learn what their program was doing. One study was cross-coded as promoting understanding and empowerment.

Lawrence et al. (2000) sought to implement a participatory approach that facilitated a two-way learning process between farmers and technical researchers, designed to lead to local knowledge creation. All six cases coded here demonstrated transformative outcomes. In some definitions of empowerment, promoting understanding and cooperation among beneficiaries and organizations is in fact empowerment. Therefore, it is not surprising that these evaluations also reported the attainment of transformative outcomes.

Empowerment. Ten studies were coded as driven by empowerment purposes, including social and political empowerment. Half of the evaluation rationales acknowledged that disadvantaged populations could benefit from the evaluative process and were also co-owners of the organization and therefore, the evaluation (Hamilton et al., 1998; Lawrence et al., 2000; Parkinson, 2009; Symes & Jasser, 2000; Vernooy et al., 2006). Eighty percent of the cases coded to empowerment rationales demonstrated evidence of transformative outcomes. This wasn't surprising to see in the data because

Taken as a singular element in the larger context of evaluation, there do not seem to be any straightforward patterns in the data in terms of the impact the evaluation rationale has on the attainment of transformative outcomes. All rationales produced transformative outcomes. This is likely influenced by the inclusion criteria for this study as all studies had to be explicitly transformative in nature.

6.3.1.4 Organizational Experience with Evaluation. Located under the pre-existing conditions node, and the parent node of organizational attributes, experience with evaluation indicates the amount of experience with evaluation, or the concept of program evaluation, prior to the intervention studied in the case.

Six of 24 cases presented no evidence of previous evaluation experience. It should be noted, however, that only one case explicitly mentions the organization having no previous experience with evaluation.

The remaining 17 cases demonstrated evidence that the organization had prior experience with evaluation or evaluation-like activities. Evaluation and evaluation-like experiences were described in the cases in a variety of ways, ranging from participating in traditional community consultation processes to formal evaluation training and

facilitation, to be expanded on below.

One case reported how evaluation-like activities were integral to the ethos of the organization. Newman (2008) describes the Reflect approach, where learners are active participants in constructing knowledge, and the (usually local) teacher supports that process. In fact, this strong organizational culture based on discussion, reflection and analysis challenges the validity of external evaluators who may misinterpret, and report only what power holders wish to hear. They viewed external processes as limiting in participation and influence. This case reported empowerment outcomes.

Much of the previous experience in evaluation for many of the cases is an indirect result of participatory development planning and priority setting (Alzate, 2000; Hamilton et al., 1998; Hart & Rajbhandary, 2003; Symes & Jasser, 2000). These processes made space for monitoring and evaluation systems with active community involvement.

Symes and Jasser (2000) describe how the prevailing situation in Palestine led to popular struggle and grassroots organizing that required quick action and short-term objective setting. It is important to note that, in contrast to the cases described above, the authors state that this setting led to evaluation activities but did little to promote participatory methods in planning, monitoring and evaluation. All four of these cases reported transformative outcomes.

Mompati and Prinsen's (2000) discussion of ethnicity and participatory development discusses how the prior existence of traditional community consultation and decision-making in Botswana had a detrimental effect on the evaluation process. This consultation process, led by the kgosi (chief), systematically excluded ethnic minority groups and in truth generally imparted information or the issuing of instructions rather than hearing the concerns of local citizens. This is one of the only cases that analyses the impact that previous evaluation experience had on the conduct of their study. This history impacted the plurality of voices present in the evaluation. This case did not report any transformative outcomes.

There were few examples of formal training and participation in evaluation (Abes, 2000; Cornwall, 2014; Torres, 2000). They describe the experience of the organization in informal and largely undocumented evaluation-like activities happening before the

studied evaluation.

Other cases describe the semi-regular use of baseline activities and monitoring and evaluation strategies (Holte-McKenzie et al., 2006; Lawrence et al., 2000; Letichevsky, 2012; Vernooy et al., 2006). Often times, the results of these activities served as the foundation for future formal evaluation activities and were often limited in participation. Seven of eight of the cases coded here report transformative outcomes, suggesting that formal evaluation training and experience could result in the attainment of transformative outcomes.

Only one case inferred a lack of experience with evaluation in the organization. Even then, it is unclear whether the organization had previous experience with evaluation; what is clear is that the participants had no prior experience with the evaluation system being implemented. Parkinson (2009) describes the NGO as having introduced PM and E to farmer groups at the same time as a loan program. The implication of this, in Parkinson's view, is that compliance with the PM and E was widely understood by farmers to be a condition of the loans. This case did not report any empowerment outcomes.

The studies provided rich detail in this category. The evidence suggests that evaluation experience may have an impact on the achievement of empowerment outcomes. Experience with evaluation or program planning and priority setting can facilitate the achievement of empowerment outcomes, and demonstrates that the nature of the experience (empowering vs. consultative) is also important. In cases where evaluative activities were not collaborative, evidence of empowerment outcomes was not present.

6.3.1.5 Evaluator Paradigm. Twenty-two cases exhibited evidence of evaluator paradigm and location. The following presents a discussion of the different themes that describe the evaluator's paradigm in regard to evaluation and participation. The section also includes a discussion of the composition of the evaluation team, including information on their background, profession, experience with evaluation and transformative processes and the local context. A great deal of interpretation was required to populate this node, specifically in reference to evaluator paradigm, which was rarely discussed openly in the text.

Evaluation as emancipatory describes an evaluator, whether external or internal, that is committed to collaborative approaches to evaluation not only as a means for increasing participation in the evaluation, but also that the participation itself is a means to a social justice end such as decreasing reliance on external organizations, interrogating power and powerlessness. Of 24 total cases, 14 were coded to this node.

Evaluation for learning still involves participation, but participation is seen as a learning tool. Participants learn about the organization's decision making processes and evaluation, but their involvement is not explicitly aimed at changing the structure of organizations or questioning the decision making processes. Importantly, those that viewed evaluation as an emancipatory process also regarded it as a location for learning.

It is important to note here that the evaluator paradigm did not necessarily directly translate into actually implementing an empowering approach. This could be for a variety of reasons (time and resource constraints, needs of the organization or donor, etc.), including some cases where the evaluator spoke of empowerment as a goal, but the methods used did not reflect this goal.

Many of the articles were written by internal evaluators, and in these cases, it is assumed that the evaluator shared the same understanding of evaluation and participation as the organization, unless explicitly stated otherwise. Testing of these assumptions was not noted in the evidence, however. For example, Alzate (2000), Symes and Jasser (2000) and Ward (2000) presented cases where the organization envisioned PM and E as a way to empower disadvantaged beneficiaries and applied techniques designed to encourage appropriate levels of participation to attain that goal. This was demonstrated by the evaluators simply stating that they believed PM and E could be empowering, but also in the ways they questioned the role of evaluation in the organization. For example, Symes and Jasser (2000) explain that they faced constraints to promoting participation because of the Palestinian context. They did this through opposing a "culture of occupation" and working towards overcoming people's sense of powerlessness. All of the cases demonstrated transformative outcomes.

A handful of cases exhibited a critique of participation as an empowering concept (Bi Niba & Green, 2005; Mompati & Prinsen, 2000; Newman, 2008; Parkinson, 2009;

Vernooy et al., 2006). This does not mean that they recognize the real-world constraints of participation in terms of resources and time broadly, but that they question the add-value of adopting a participatory approach. In these cases, the evaluator made deliberate decisions about when and how participation might occur by taking into account the context of the evaluation, including but not limited to: the history of conflict, historical relationships between the organization and the participants, existing power structures within the organization and the community, the ability (time, skills, interest) of the participants, and whether participation would have predominantly positive or negative effects on participants. Additionally, evaluators who were critical of the concept of participation often were critical of their own role in the process and of evaluation in general.

Three studies wrote explicitly about the authors' role as an evaluator and how their participation produced a power dynamic that needed to be interrogated and understood as the evaluation progressed. For example, Botcheva et al. (2009) spoke at length about their role as outsiders emphasizing continued adjustment and adaptation to "achieve the best fit between evaluation process and program cultural context" with evaluators striving for the "highest level of cultural competence possible..." (pp.179-179).

Many of the cases presented above demonstrated evidence of empowerment outcomes indicating that having a critical perspective of evaluation can facilitate empowerment. More importantly, the alignment of the evaluator's actions with their critical paradigm appears to be a significant factor.

Motteaux (1999) presents an interesting case. At the time of writing, the author was a research student with a special interest in community. She commenced her research on water quality and safe-drinking water sites initially from a traditional, positivist approach; however, as time progressed she became involved in the community and realized her role was changing from positivist, detached researcher to that of a facilitator and co-researcher working closely with the community. As a result of this experience, she adopted Participatory Rural Appraisal methods to integrate indigenous technical knowledge.

A number of studies involved teams of evaluators from slightly different

backgrounds (Brandão et al., 2012; Hamilton et al., 1998; Hart & Rajbhandary, 2003; Letichevsky & Penna Firme, 2012; Mompati & Prinsen, 2000; Newman, 2008; Vernooy et al., 2006) and located in different areas, normally one in academia or government and one closer to the field. Most of the work reflected a critical understanding of race, ethnicity, power and exclusion.

Eighty-six percent of these cases demonstrated evidence of transformative outcomes. They share a common characteristic of including program insiders as leaders of the evaluation effort. These cases are differentiated from the remaining cases, described below, because they demonstrate the participation of lead evaluators who were internal to the program, the evaluation sponsor or possessed specific knowledge or experience of the program or local context. Each case also exhibited an understanding of participation as both a functional and an empowering process. The remainder of these cases were written by external evaluators.

Some cases coded to this node were led by evaluators with experience working with traditionally marginalized populations and research or evaluation (Botcheva et al., 2009; Dageid & Duckert, 2007). The authors adopted a participatory approach to openly address empowerment and issues of power in evaluation and as a way to validate data.

None of the evaluations conducted by external evaluators demonstrated evidence of empowerment outcomes as defined by this study. The evidence presented above presents interesting insights into a larger question regarding an external evaluator's ability to influence the empowerment of others; however, the method of data collection for this thesis does not provide sufficient evidence to answer this important question.

A minority of cases that demonstrated evidence of empowerment suggest that evaluation and participation were utilized for practical purposes, including evaluation capacity building, demonstrated evidence of transformative outcomes. These cases involved mixed teams, and included program delivery staff to enhance the use of local knowledge in the evaluation.

In summary, in a great majority of cases where the evaluator held a critical understanding of evaluation and participation, the process resulted in empowerment outcomes. This factor was particularly potent when the evaluation methodology reflected

this stance. Adopting a view of participation as a means to practical ends had a milder effect on empowerment outcomes. This section also looked at evaluator location. Further research is required to fully understand why the evidence suggests that cases in which the evaluations were led by external evaluators did not demonstrate empowerment outcomes.

6.3.1.6 Resources. Few articles mention resources made available for the evaluation. Only one article describes the percentage of the total budget given to evaluation (11%) (Brandon et al., 2014). With such limited data, it is difficult to ascertain whether program resources facilitated or inhibited the attainment of transformative outcomes.

Summary. This subsection provided data to suggest that a variety of forces are at play in the context of an intervention before an evaluation begins. These range from local, national and international political movements, local relationships, organizational history with evaluation and the composition and paradigm of the evaluation team itself. These factors need to be understood and appreciated in the design and implementation of a transformative evaluation. It is also important to understand that these influences are continually in flux and do not occur in a vacuum but rather are a complex system of inter-related factors.

Due to the variability of available evidence, it is difficult to suggest which factors lead to the attainment of transformative outcomes. However, the data suggest that local context does impact transformative outcomes. It was clear that openings in national political systems provided new opportunities for local control of resources and program planning which can enhance evaluation practice and the attainment of empowerment outcomes. In these cases, it would seem that empowerment as a part of the evaluation process was part of a natural progression, led by local stakeholders. The data also suggest that the organizational rationale must be compatible with empowerment processes. Experience with evaluation or evaluation-like activities is also a critical factor. Interestingly, experience with traditional modes of evaluation or consultation seem to negatively influence the attainment of empowerment outcomes while a history of collaborative evaluation and evaluation-like activities have a highly positive influence on the attainment of outcomes. The evidence also suggests that evaluators who hold a critical perspective on evaluation can positively influence the achievement of

empowerment outcomes whereas evaluators holding a perspective more aligned with the practical ends of participation are less likely to produce empowerment outcomes.

6.3.2 Mediating Activities. The factors and conditions presented in the preceding chapter inherently influence the practice of transformative evaluation. In the following section I present a sample of the evidence collected from the cases on the fundamental dimensions of participatory practice as outlined by Cousins and Whitmore (1998): control over the evaluation process, stakeholder diversity and depth of participation in the evaluation knowledge production.

6.3.2.1 Control Over the Process. All cases provided data that could be coded to this attribute; however the data quality varied across cases. Cases were ranked from high, indicating a collaborative or shared evaluation process to low, indicating little to no evidence of collaboration between the evaluator and other stakeholders such as program donors, program staff, beneficiaries, and the larger community.

Many cases exhibited very high levels of collaboration that made use of elected representatives, travelling assemblies, and large training programs (Alzate, 2000; Newman, 2008; Torres, 2000). Authors describe utilizing a simplified methodology in order to work directly with community and facilitate local understanding. Their process enabled communities to review the expected results, adjust goals, formulate new strategies and learn how to carry out specific evaluation and monitoring activities, including interpreting data.

The evidence from these studies suggests that collaborating extensively with stakeholders in meaningful ways appears to enhance the attainment of empowerment outcomes. This requires that decisions that guide the evaluation are made collectively and that the evaluator must give up a certain degree of power over the course that the evaluation will take.

There were a few cases that described limited participation in decision-making (Abes, 2000; Hamilton et al., 1998, Hart & Rajbhandary, 2003; Holte-McKenzie et al., 2006; Letichevsky & Penna Firme, 2012; Motteaux et al., 1999; Ward, 2000). In these cases training was provided on a variety of evaluation processes however community involvement and training was not consistent and there was a tendency for staff to control

the process. For example, evaluation questions and data collection methods were predominantly decided upon by the evaluators. These decisions were sometimes attributed to the limited facilitation capacity of the field team.

Several cases provided little detail outlining the actual responsibilities of the various stakeholder groups (Brandão et al., 2012; Bi Niba & Green, 2005). For example, Brandão et al. (2012), describe how the youngsters fulfilling their program requirements were included in decision making about the evaluation, although it is unclear how this took place and to what extent.

The majority of cases described above demonstrated evidence of transformative outcomes. These outcomes are characterized by fluidity in evaluator role, openness to questioning the process and involving beneficiaries in some official capacity. However, the main decision making power rests in the hands of the evaluator.

The following cases demonstrated high levels of evaluator control (Parkinson, 2009; Dageid & Duckert, 2007). In some cases, the intention was to involve participants in decision making but in practice this was not achieved (Parkinson, 2009).

Some evaluators controlled the process to facilitate fuller participation from different groups (Cornwall, 2014; Kariuki & Njuki, 2013; Mompati & Prinsen, 2000). Four of six cases here demonstrated the attainment of transformative outcomes.

Mompoti and Prinsen (2000) provide an example of how lower control over decision on the part of participants can benefit participation. The case highlights how traditional chiefdoms exerted power over the PRA process put in place in Artesia, Botswana. In order to circumvent the power imbalance, the PRA team organized separate sessions for different ethnic minorities. While effective in the beginning, the second day, villagers noted that an elder was noting who spoke up against the established order, culminating in the locals being trained to threaten to quit that same evening, for fear of getting in trouble. In another locality, a woman from a subordinate ethnic group spoke out against discriminatory practices, and was removed from the meeting at the request of the chief. The meeting continued as if nothing had happened. Consequently, future meetings were held at a distance from central locations. Temporary and outsider interventions can rarely change power balances directly due to the above-mentioned

constraints. Even when participation is achieved, this may still be highly constrained by the subordinate's internalized self-concepts. In addition, reprisals for participating may be enacted as soon as the outsiders leave; therefore the authors argue that careful facilitation is required to minimize risks for subordinate groups.

Overall, the majority of cases reporting on the decision making process in their evaluations demonstrated the achievement of empowerment outcomes. Importantly, three cases provided evidence to suggest that high evaluator control can facilitate greater participation, and therefore perhaps transformative outcomes. This conclusion is interesting as it raises questions about whether the outcome of empowerment trumps the means to achieving it. Balanced or shared control has been argued (Cousins & Chouinard, 2012) as best and most democratic, which seems to be evidenced in data presented above. However, others (Daginault & Jacob, 2009) would argue that higher stakeholder control is actually participatorier. I will discuss this in further detail in the Discussion section.

6.3.2.2 Stakeholder Diversity. Collaborative transformative evaluation requires stakeholders work as collaborators in the evaluation process. Another dimension of participatory practice is the diversity of voices present in the evaluation. Evidence collected by Cousins (2001) demonstrated that the median number of none evaluator stakeholders involved in an evaluation is six. In evaluation literature, there are both pros and cons noted in involving a diversity of stakeholders in the process. In the case of transformative approaches to evaluation, I believe that a wide variety of stakeholder voices are required to elicit lasting transformative change. We require the participation of all levels to create an institutional climate that is open to transformation.

Various cases included community members beyond beneficiaries in their evaluation process. For example, Abes (2000) included the community of leader-graduates as well as a comparison group, comprising community leaders who had not participated in any of the courses. In total, seven different groups of people have been involved in the evaluation process in a variety of ways and the program is currently working to enhance the roles of those involved.

Other authors describe involving hundreds of individuals reaching far into the community and external institutions and NGOs as data collection agents and sources, as well as data interpreters (Alzate, 2000; Blauert & Quintanar, 2000; Holte-McKenzie et al., 2006; Lawrence et al., 2000; Torres, 2000). Eight of nine of these cases demonstrated the attainment of transformative outcomes, providing a basis to the argument that a variety of stakeholders, including those external to the intervention, can lead to or enhance empowerment.

Various evaluations involved program staff and beneficiaries in the evaluation process and decision making (Brandão et al., 2012; Hamilton et al., 1998; Hart & Rajbhandary, 2003; Motteaux et al., 1999). For example, Hamilton et al. (1998) involved joint discussion and planning between project teams and their beneficiaries, who were also divided by caste or ethnic group, with each developing their own goals and indicators, as well as providing an assessment of current activities. Combined indicators were then presented to general assemblies that agreed on a strategy for future monitoring and evaluation. 100 percent of these cases exhibited transformative outcomes.

A variety of cases involved very few stakeholders, and a number of cases provided very little information on the variety of stakeholders present at the decision making table (Botcheva et al., 2009; Symes & Jasser, 2000). For example, Symes and Jasser (2000) speak about the community and their involvement, but there is no further descriptive information about who the community is, and how they were involved in the process.

One case discussed how the stakeholder group was broad but overpowered by the dominant group. Mompati and Prinsen (2000) discuss how elected participants in the evaluation process who were to take over when the PRA team left the site were invariably members of dominant ethnic groups and further, impacted how subordinate groups participated in the larger process. Perhaps in this case a limited stakeholder group would be beneficial to the participatory process. Half of these cases exhibited evidence of transformative outcomes.

Some cases did not provide ample information on who was involved in the evaluation to a level that was able to contribute to this study. These studies referred to the

community or to stakeholders broadly and did not specify who they were or how they participated. More attention to these details would have enhanced this section.

Involving a variety of stakeholders appears to be a critical factor in the achievement of empowerment outcomes. The majority of cases in this study involving a wide variety of stakeholders proved to enable empowerment outcomes. However, in one case the variety of stakeholders proved detrimental to the participation of subordinate group members. As was discussed by Cousins and Chouinard (2012), the more diverse the stakeholder group the more likely issues of power and privilege will come forward. The question then becomes: Is evaluator intervention, such as limiting certain powerful stakeholder groups' participation to increase the meaningful participation of less powerful stakeholders justifiable? Evaluators recognize that the involvement of visibly powerful stakeholders is often required to affect change and have evaluation recommendations acted upon.

However, evaluators also recognize that powerful stakeholders often take up more space, influence the less powerful groups from participating meaningfully and unfairly bias evaluation findings to their own benefit. Power operates in these contexts on three levels: visible, hidden and invisible. The danger of evaluator intervention as described above is manifested in these three levels. Often, evaluators are only privy to limited information about community context and often only address issues of power that are obvious and often coloured by their own cultural values, which can fail to understand the complex power dynamics of communities, families and individuals. In the end, this approach to intervention creates an artificial context and can upset and undermine the on-going power struggles of a community and have lasting impacts on empowerment. To fully understand this phenomenon, a deeper understanding of community dynamics and the role of the evaluator are needed. This in turn can influence evaluators in facilitating patterns of dominance and dependence in complex community interventions.

6.3.3.3 Depth of Participation. The final fundamental dimension of participatory practice is depth of participation. Drawing from Cousins and Whitmore (1998), deep participation requires involvement throughout the evaluation process, from the design to the implementation and reporting on evaluation outcomes whereas a low depth of

participation, and what Cousins and Whitmore (1998) referred to as consultation, would be any case where stakeholders had no decision making control or responsibility. My elected frame of reference for understanding the participation of stakeholders is centered on the experiences of the beneficiary communities.

Before describing the findings, a quick note on the use of PRA as it relates to depth of participation. The use of PRA techniques requires that data collection and analysis often occur at the same time; however, unless it is explicitly mentioned that the general findings were brought back to the community for further analysis, I consider this sort of analysis to be a part of the reflection process and further, one that would occur naturally if properly facilitated in a focus group. Therefore, the use of PRA did not necessarily register as deep participation.

An example of deep participation is the case by Abes (2000). Abes utilized participant researchers, leader researchers and program staff to select sample sizes, develop indicators, the design of data collection and data gathering and analysis. As each evaluation cycle runs its course, more opportunities for further involvement of stakeholders are introduced. The work by Alzate (2000), Blauert and Quintanar (2000), Hamilton et al. (1998) Symes and Jasser (2000) and Brandão et al. (2012) also provide evidence of deep participation of stakeholders. 100 percent of these cases demonstrated transformative outcomes and indicates that deep participation of participants in an evaluation contributes to empowerment.

Many cases had stakeholders involved in data collection and analysis, as was the case of Cornwall (2014). This was most common because the vast amount of cases utilized PRA techniques that by virtue of design involve beneficiaries in reflecting and analyzing their own data (Newman, 2008; Torres, 2000; Vernoooy et al., 2006). As mentioned above, I decided that simple involvement in PRA techniques did not equate to participants being involved in the full analysis of the data, unless the authors explicitly mention that the general findings were brought back to the community to analyze further.

In other cases, stakeholders were involved in designing indicators and collecting data (Dageid & Duckert, 2007) or in developing the program logic and providing input

on the development of data collection instruments (Brandon et al., 2014; Ward, 1998). Just over half of these cases demonstrated the attainment of transformative outcomes.

The cases coded in this section provided evidence to suggest that stakeholder participation was primarily as data collection sources. Cases that were coded as low participation were: Bi Niba and Green (2005), Botcheva et al. (2009), Donnelly (2010), Kariuki and Njuki (2013), Motteaux et al. (1999) and Parkinson (2009). Six of eight of these cases demonstrated transformative outcomes.

Letichevsky and Penna Firme (2012) provide an interesting case. Based on the evidence provided, the team made the majority of the decisions around evaluation design and approach, and even carried out the majority of data collection, analysis and reporting. Using rapid reporting, they provided findings and feedback to specific community members on a regular basis for them to validate. Even though participation based on evidence provided would be considered low, based on outcomes it would appear that the process was appropriate and transformative for the beneficiaries.

A variety of cases were unclear as to who was participating at various times throughout the evaluation process. For example, involvement of stakeholders was mentioned but there was a lack of clarity on what roles and responsibilities the various participants (Holte-McKenzie et al., 2006; Vernooy et al., 2006). Both cases demonstrated transformative outcomes.

In summary, the evidence suggests that depth of participation may not correlate in a linear way with the attainment of transformative outcomes and that context is key in understanding appropriate levels of participation. While cases that exhibited deep levels of participation among beneficiaries demonstrated empowerment outcomes, it was also determined that shallow participation also led to empowerment outcomes. Again, more evidence is needed to fully understand the complexity of this factor as various cases did not supply ample details to fully understand who was involved and to what level.

Summary. The above subsection outlined some of the evidence provided by the cases on the mediating factors in an evaluation based on the framework provided by Cousins and Whitmore (1998). As with the pre-existing conditions, there was also a variety in the evidence provided by authors for each case. For this reason, it is difficult to ascertain

patterns within the data that suggest what factors and conditions lead or detract from the attainment of transformative outcomes. An examination of the evidence suggests that there is great variety in how participatory processes unfold and that a great deal of flexibility is required on the part of the evaluation lead. It is also clear that these three mediating factors are critical in the evaluation process and must be handled with appropriate care by the evaluator as they can lead to empowerment outcomes but can equally limit empowerment if not applied appropriately.

Conclusion. The question as to which processes and conditions led to the attainment of empowerment outcomes proved to be a difficult question to answer.

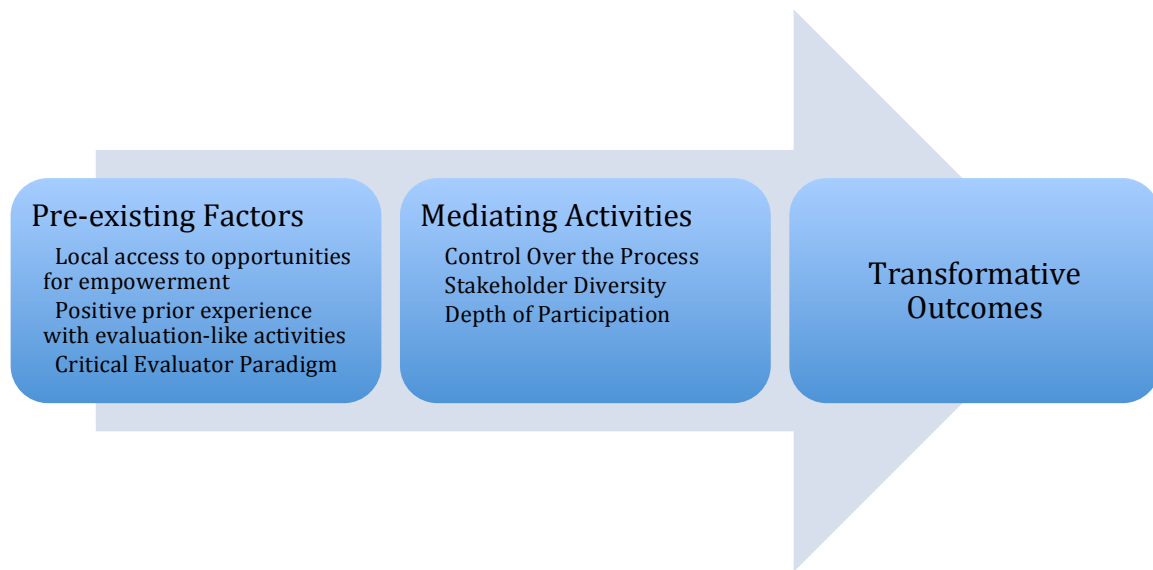


Figure 1. Critical Factors in Achieving Empowerment Outcomes in Collaborative Approaches to T-PE

Although rudimentary, Figure 1 represents the broad findings of this research. Participatory practice interacts deeply with the pre-existing conditions outlined in the previous section of this chapter to produce transformative outcomes. The data demonstrates great variability in the evidence provided by authors for each case. For this reason, it is difficult to ascertain patterns within the data that suggest what factors and conditions lead or detract from the attainment of transformative outcomes.

In regard to pre-existing conditions, the evidence suggests that national and local contexts that provide opportunities for access and agency of the community enhance

empowerment. It is also clear that experience with evaluation and evaluation-like activities that promote collaboration also influence empowerment in a positive way. On the other hand, prior experience with traditional evaluation and consultative practices can hinder the attainment of empowerment outcomes. The evidence also suggests that evaluator paradigm can have an effect on empowerment outcomes with a critical perspective on evaluation and participation having a positive influence on empowerment outcomes and more practical participatory paradigms having a less positive effect on empowerment outcomes.

Control over the process, variety of stakeholders and depth of participation seem to be critical to understanding how collaborative approaches to evaluation actually take place in the field and more information is needed to better understand these complex processes. Evaluations in which evaluators provided ample opportunity for beneficiaries to impact the decision making in the evaluation tended to more positively impact the attainment of empowerment outcomes. In addition, a wide variety of stakeholders also appear to bolster the achievement of empowerment outcomes. However, it is important to note that in each of these categories outliers exist. Tighter control over the evaluation process and a more controlled variety of stakeholders did result in empowerment outcomes in some cases, which raises a questions such as whether the intervention of the evaluator in these situations truly constitutes transformative participatory practice. Simply put, is the outcome more important than the means of achieving it. In the following discussion section, I shall expand on these issues and provide guidelines for moving forward.

7. Discussion and Conclusion

In this chapter I provide a brief overview of the findings of the study and then outline some of its limitations. I then review the findings thematically relative to contemporary issues identified in the literature review. The section will conclude with some considerations for an agenda for further research on transformative CAE.

7.1 Limitations of the Study

This study should be understood through the lens of the following limitations. There were a limited number of cases that met the inclusion criteria; therefore

discovering patterns within the analyzed data proved difficult. To compound this issue, the data within the cases varied greatly in quality and depth. For example, some cases provided detailed information in some areas, such as pre-existing conditions, and did not provide ample information pertaining to mediating activities and outcomes, resulting in only a few complete cases with detailed information across all nodes (and questions of interest to the research). I give consideration to each of these limitations below as they relate to recommendations for moving forward. In utilizing various search terms in my sample selection, I attempted to identify as many cases of transformative CAE as possible. While outside disciplines were not specifically excluded, my search results were not from outside the evaluation and international development world.

Finally, I alone completed the data collection and analysis. As a result, no verification procedures were employed to assure coding quality outside of discussions of discrepancies and analysis with my thesis supervisor. Yet I can say that my level of comfort with coding the studies grew as I worked my way through the data. In addition, the sample was created from exclusively peer-reviewed sources, therefore enhancing the quality and trustworthiness of the data examined. I found relatively few instances where I had difficulty coding important material in the sample articles. With the exception of the few cases sourced from the edited collection *Learning from Change*, all cases were downloaded directly as PDFs into NVivo to diminish researcher selection bias.

Preliminary findings were also presented at the *European Evaluation Society's 2014 Biennial Conference* in Dublin for discussion. The presentation resonated well with those in attendance and directions for further research in this study were developed as a result.

7.2 Thematic Review of Findings

The following section presents a discussion of the findings in relation to the current critiques of empowerment evaluation as discussed in the literature review. The evidence collected in this study provides further evidence to support the key arguments posited by a range of evaluation scholars, specifically: Cousins (2005); Gregory (2000); Miller and Campbell (2006); Patton (1997, 2005); Scriven (1997); N.L. Smith (1999); and Worthington (1999).

7.2.1 Context is crucial. We have learned this lesson again and again (Cousins,

Whitmore & Shula, 2013; Rog, Fitzpatrick & Conner, 2012). Yet, when the research on evaluation is reviewed it is clear that the focus remains on the details of the methods chosen rather than the context within which the evaluation takes place, including the community's history and experience of research and evaluation. We know from previous research that externally driven research and evaluation has not been beneficial to marginalized populations and we must address these issues in our writing to inform how we as evaluators can and should act in such contexts to cause as little harm as possible. Only two of 24 cases (Symes & Jasser, 2000; Letichevsky, 2012), for example, make mention of the history of colonization and the legacy this history imparts on the organization and its initiatives, including evaluation.

We also understand that for empowerment to take place, we need to understand the institutional climate within which the organization lives and breathes. In their experience in measuring empowerment, the World Bank (2005) has identified categories of institutional climate that can impact the occurrence of empowerment outcomes and these can be applied to transformative evaluation contexts. These categories include: access to information, inclusion and participation, accountability, and local organizational capacity. Within these categories, there are also levels of attainment. In organizations where these indicators are scored highly, there is more opportunity for empowerment. In organizations that score low on the following indicators, there is less opportunity for movement of individuals and collectives. These categories can provide a framework for evaluators when discussing evaluation context in transformative approaches to evaluation in the research. A brief description of the four categories follows:

- *Access to information:* Includes a two-way communication between institutions and beneficiaries on the rules and rights to services, information about the organization's activities and performance, and opportunities to be involved. An informed citizenry is able to take advantages of opportunities, access better services and exercise their rights.
- *Inclusion and participation:* Intended beneficiaries are given the opportunity to co-produce knowledge and have some authority over decisions and resources.

This process usually requires a change in the way an organization works, creating space for traditionally marginalized groups to have a voice at the table.

- *Accountability*: In order to make change, intended beneficiaries must believe that they can hold organizations to account, both state and private. A certain amount of trust in the organization must exist.
- *Local organizational capacity*: Intended beneficiaries are able to work together to organize themselves and resources to solve their problems.

Further, giving full treatment to the organizational and evaluative context has particular impacts on data quality. As a reviewer, it was difficult to discern whether transformative outcomes were the result of the natural progression of the community (e.g., Abes, 2000; Alzate, 2000) or as a result, however indirect, of the transformative evaluation process. This is likely due to weak research design, which is addressed in a later section.

Further to issues of context, the evaluation sponsor can have significant impacts on how the evaluation is undertaken and the expected outcomes. Few articles provided information on relationships between organizations and evaluation sponsors, be they external organizations, government entities or the host organization and intended beneficiaries themselves. These relationships are often fraught with tension, a result of on-going power struggles and economic resource allocation conflicts. As detailed in the case presented by Parkinson (2009), evaluation sponsors and host organizations often have differing and competing priorities. For example, the state may sponsor CAEs because they are *en vogue*; however, participation is significantly limited because the state is actively engaged in sustaining the status quo or retaining their power. Evaluations lead by civil society groups or the local organization, characterized by a shared interest in the community's journey towards increased power and voice, provide perhaps a greater opportunity and likelihood for local empowerment.

CAE are utilized in transformative contexts for these reasons and also because we recognize that outsider perspectives (such as an external evaluator) lack the ability to fully understand a local context and that local knowledge holders must be integrated to have a chance at being transformative. The way we do research on evaluation should also

reflect this.

7.2.2 The politics of engagement. Little evidence was found in the cases that describe in detail how decisions are made in the organization undertaking the evaluation process. Further, few details are provided on how decisions were made and by whom throughout the evaluation process. Carden (2007) challenges us to pay this part of the process stricter attention, both as a basis for understanding the mechanisms of change, but also to increase local capacity building which can act as the foundation upon which empowered change can take place (earlier referred to as ‘agency outcomes’). To further understand this portion of the process, we must report on the role of the evaluator, who defines it, and how and why it changes throughout a process.

As a part of interrogating the politics of engagement, we must also ask ourselves whether the ends of empowerment ever justify the means. We must decide whether the intervention of an evaluator in a conflict, whether sparked by the evaluation or a long-time dynamic, is appropriate. For instance, in the case presented by Letichevsky and Penna Firme (2012), the evaluation team intervened and re-oriented a process where marginalized groups were being displaced from the conversation. To intervene seems like second nature to a democratically focused evaluator who sees value in the presence and opinion of all participants. However, in making this choice we are altering the evaluation outcome and potentially disrupting current power struggles.

We are also potentially acting against several core tenets of empowerment. The intersection of inclusion, democracy, ownership and social justice is complex. Transformation is only truly attained when every party is empowered to give their assets and resources together to achieve a common vision.

The role of municipalities, the state and civil society cannot be understated. In many cases that demonstrated empowerment outcomes, clear links can be drawn between local and national government openings that are suggestive of allowing space for empowerment. Evaluation is a blip in a larger system of complex relationships. The moment must be ripe for evaluation to complement natural process of empowerment. Evaluation cannot act as a means towards empowerment in a vacuum. We have a responsibility to understand where empowerment approaches to evaluation are feasible

and warranted. Are empowerment approaches to evaluation always appropriate or can they cause more harm than good when placed in the larger context?

Post-positive approaches to research were abandoned by many because of the tendency to objectify research subjects and exclude them from the research process, thus deciding the fate of programs and perhaps populations without meaningful involvement of these populations. We know now that simply engaging populations does not free us of these responsibilities. As discussed in the literature review, Gregory (2000) states that participation does not naturally mediate power dynamics and can also serve to facilitate their fortification if cultural norms are not fully investigated and addressed. As a community, evaluators must continue to problematize our ideas of how and why participation takes place and ensure that the community takes the leadership in this area.

Collaboration and censorship are not mutually exclusive and, according to Turnbull (2000), are inextricably linked. Co-creating knowledge raises more questions than provides answers. These questions are political, practical, and ethical. The practical considerations of empowerment approaches to evaluation have been well documented; however, as was recognised in this thesis, the ethical and political considerations are not given equal treatment in the current empirical evidence.

Evidence of the complexity of these questions also lies in the case of Mompati and Prinsen (2000) as well as the case presented by Letichevsky (2012) where participating stakeholders were limited by the evaluation team to allowing for more meaningful participation of the traditionally powerless in the community. Few details are given on how these decisions were made, who made them, and what the consequences were of those choices. In accordance with the argument put forth by Smith (2007), the evaluator role must be further investigated. We must also consider the impacts of limiting the participation of the powerful populations in CAE. Generally speaking, the powerful people in communities are not likely to give up their positions as decision makers, and any lasting change must also include their buy-in. When we engage in evaluations that take place over a short amount of time that exclude the powerful we are potentially ensuring that the empowerment outcomes attained are not sustained.

7.2.3 Clarity on the concept of empowerment. As discussed in the Findings section, there is great variation in definitions of empowerment. Empowerment is usually never defined as a concept, but discussed through the definition of a proxy. Few included cases mention a discussion of empowerment with beneficiaries prior to engaging in evaluation activities. Therefore, there is no evidence to suggest that evaluators and intended beneficiaries ever discussed a working definition of empowerment throughout the evaluation.

Based on this research, conceptual clarity is most ambiguous in differentiating between capacity building and empowerment. The evidence that empowerment has taken place is when a subordinate group effectively advances their own priorities or interests through their own choice and action. I decided to remove evaluation capacity building and evaluation use from the category of empowerment if outcomes only pertain to evaluation; however, in instances where evaluation skills or evaluation use led to a subordinate group effectively advancing their own priorities or interests through their own choice and action, I included it as a transformative outcome. In the end, only 18 of the 24 studies had outcomes that were classified as transformative.

In the study by Bi Niba and Green (2005), the evidence presented of concrete empowerment takes the form of capacity building and knowledge acquisition; communities acquired new skills and knowledge about water contamination and erosion and applied this knowledge to solve their problems. In fact, their case was the only study that revisited the project area later on to see if changes had taken root in the communities and concluded that transformative outcomes had indeed persisted. The absence of other examples of follow-up is astonishing given the interest in writing about and commenting on transformative outcomes.

7.2.4 What constitutes good evidence? Once a definition of empowerment exists, research requires that it must be operationalized through evidence. The majority of the evidence presented in the cases consisted of anecdotal experiences and narratives described by the authors. Using self-report often raises questions of validity and rigour and this discussion strikes at the heart of why evaluators choose to engage in empowering processes. In their study, Miller and Campbell (2006) identify anecdotal impressions as

inherently problematic, empirical evidence that is non-valid or less reliable and argued for the use of meta-evaluation to verify successful cases. I argue that their terms of reliability and validity should expand to encompass the values of empowerment and transformational research, which do not view anecdotal evidence in a strictly negative light. There is a conflict of epistemology here that needs to be addressed. In an attempt to define the validity of evidence, I address validity through the lens of post-positivist, emancipatory research- construct validity, face validity and catalytic validity. Construct validity, defined by Lather (1986) as a systematized reflexivity to ensure that observations represent the construct being investigated. As previously noted, Lincoln and Guba's (1981) face validity refers to "member checks" and catalytic validity refers to the research's ability to energize participants to know and transform their reality (Lather, 1986).

In accepting that anecdotal evidence is valid and reliable, evaluation research must document the evidence that construct validity, face validity and catalytic validity have been achieved. This evidence should be identified and defined by those who benefit from the transformative process, including rich examples that differentiate program outcomes from evaluation outcomes. In this study, only four cases utilized participant quotations as evidence.

However, all anecdotal evidence should not be received in the same way. Empiricism, as defined by Cousins and Chouinard (2012), is about making sense of observation. Creswell (2007) developed a set of criteria when assessing the quality of qualitative research, including prolonged engagement and persistent observation, triangulation, peer-reviewing, negative case analysis, investigation into researcher bias, member-checking, rich descriptions and external audits. A few of these methods have been outlined above. It is pertinent that not one of the included cases provided evidence of meeting any of these criteria. For example, in most of the empirical evidence presented in this thesis, anecdotal evidence is provided from the position of the lead researcher or evaluator. Anecdotal evidence can also alienate and obfuscate truth. While anecdotal evidence allows for local knowledge to inform general understandings, individual impressions can be presented as a collective opinion.

Anecdotal evidence presented to emphasize the successes of an evaluation, reported by the lead evaluator or researcher, do not bode well for trustworthiness. We must demand more of our colleagues and ourselves. As Cousins and Chouinard (2012) emphasize, one way we can improve our reflective case narratives and the anecdotal evidence provided within is to better frame our narratives. We must be more explicit in our methods, write from the point of view of multiple perspectives and use evidence from the focal evaluation to support claims. The authors argue that these criteria do not necessarily ensure trustworthiness but they can provide a context within which to interpret the cases.

Further, Cousins and Chouinard (2012) emphasize the author/researcher as instrument and therefore suggest that the positionality of the narrator or author be further defined in studies. In this process, writers must also describe their own validation processes. Angen (2000) provides us with two more types of validation: ethical and substantive validation. The term substantive validation indicates that the understanding one's own understandings of the topic, the understandings derived from other sources, and the realization of this process in the written study is also important. Self-reflection, in other words, contributes to the validation of the work. Further, ethical validation means that research agendas must question their underlying moral assumptions, their political and ethical implications and the equitable treatment of diverse voices.

I agree with Miller and Campbell's (2006) and Cousins' (2005) position that further use of meta-evaluation activities and follow-up research are required to move the field forward. Empowerment is defined as an increased critical awareness and an ability to analyze life circumstance that one wishes to change, has the ability to change, and act in a way to ensure that change occurs over time. It is simply not sufficient to run through a process and call it empowerment without analyzing the actual transformation in individual and community livelihoods after sufficient time has passed.

This final point leads us to consider the timing of the cases. As previously indicated, one of the included cases report revisiting the evaluation site to investigate empowerment claims. The remaining cases reported on evaluations utilizing anecdotal evidence collected throughout and directly following the evaluation process. This raises

the question: at what point do skills gained or insights experienced translate into actual demonstrated empowerment? And if we do not allow for sufficient time to pass, how do we know that what we have recorded as empowerment was sustained over time? In the development context, there are countless stories of diminishing outcomes; counted outputs that no longer exist five years later. Meta-evaluations of these interventions are an excellent way to understand empowerment in context. I believe that to fully understand the operationalization and attainment of empowerment outcomes we require that evaluation sponsors, program staff and evaluators understand that meta-evaluation is standard practice. A new model is needed.

7.3 Implications for Research

The following section details how the evidence, coupled with the current empirical literature on evaluation research, provides a guide towards better practice in research on CAE. These implications for research reflect the demonstrated weaknesses in the current research on evaluation literature, as identified by this study. Overall, I am calling for an increase in research on transformative evaluation that provides rich descriptions that include deeper engagement with transformative processes, including the ecosystem within which they function. Further, I discuss the evidence demonstrated and suggestions for improving quality of data and research design, with attention paid to complex power dynamics, which push our own practice towards more transformative ends.

7.3.1 Small sample size. In order to truly understand this complex phenomenon, we would benefit from the availability of a wider range and volume of studies. A small sample of cases fit my broad search criteria so as to make it difficult to understand each case in its own complexity and connect together with the other cases. More research on transformative evaluation would provide a more solid basis for understanding patterns that would potentially provide insight into the mechanisms for change. As transformative practice continues to expand in the field of evaluation, evaluation community members should engage in more evaluation research to further our understanding of the complexities involved.

7.3.2 Variability of data. As was made clear in this study, evaluations take place in complex environments. The ecosystem is rich with detail, and one must be conscious that

an evaluation journal article is not an anthropological study. However, there remains work to be done to further extrapolate and understand the real mechanisms for change in an evaluation environment. We do have evidence from the field of evaluation research as well as the field of participatory development research that suggests reasonable areas for our attention. These areas must be brought to the forefront of evaluation writing so that we can continue to improve transformative practice in the field.

7.3.3 Wider Range of Research Designs. All of the above supports the understanding that while various forms of evidence can and should be accepted, we must strive for better research design when undertaking empirical studies to strengthen our knowledge base and improve practice in the field. Reflective case narratives must be strengthened to ensure that interpretive concepts of validity are satisfied. Further, in this study no use of mixed-methods was reported in gathering data on evaluation outcomes. In the field, we must continue to experiment and try new ways of understanding our own work to compliment and reinforce our empowerment goals.

7.3.4 Authorship. If we are to interrogate the evaluation process and our writing of it, we must also address the question of power and control in the writing process. With the exception of the book chapter cases from *Learning From Change*, the lead evaluators, who were not representative of the organization or the participant group, authored the majority of the cases. We must recognize that while there are challenges to co-authoring academic papers, if we desire to push our transformative practice further, we must be revolutionary in our methods and ask ourselves what we can do to further the principles of ownership, democracy, inclusivity and social justice. Failing to address these broad issues of power and control have the potential to further marginalize people (Gregory, 2000), leading us directly away from our stated purposes when engaging in these transformative practices.

7.3.5 Power Dynamics. Most importantly, this study provides a rationale to be more inclusive of power dynamics in the way that we conduct evaluations and research on evaluation. Several factors and conditions surfaced questions that interrogate our own understanding of empowerment and our role in that process. We must continue to search for these answers in our own research practice by engaging in the difficult questions of

democracy, inclusion, ownership and social justice through a lens that honours our empowerment values and the multitude of voices we invite to the table when we pursue this work, particularly those of the traditionally marginalized. How do the choices and actions of an evaluator operating in a post-colonial environment interact with these principles and a community's ability to define and act out empowerment on their own terms?

8. Conclusion

In this thesis, I presented evidence to support the development of our understanding of the operationalization of transformational outcomes as achieved in development evaluations utilizing transformative approaches to collaborative evaluation. As a point of clarification, this research suggests that we seek further understanding of our own engagement in these approaches to evaluation and the results of this work not to create a singular toolkit to guide the practice of transformative CAE, but rather to better understand the consequences of our own work. Meta-evaluations serve to cultivate better practice through increasing our understanding of complex interventions, particularly when we pay specific attention to the areas indicated above. To return to the argument put forward by Carden (2007), the goal of an empowering practice is to encourage and support the creation of a sustained local culture of evaluation that reflect culturally-relevant understandings of empowerment. This thesis supports understanding our own work so to ensure this locally viable culture of evaluation.

Overwhelmingly, this research suggests that to improve our practice we must define and describe transformative outcomes with program stakeholders and document our processes more clearly to enhance learning. Through defining empowerment, we are more capable of assessing our impacts through our literature in a way that demonstrates validity that corresponds with the aims of transformative evaluation and are aligned with the objectives of empowerment. As presented in the discussion section, in order to push the research on transformative evaluation further, we must continue to write about these processes with a focus on local context, and provide rich descriptions of the process, including the changing roles of the evaluator and participants and the issues of power found within.

The dearth of information on local context, the evaluation process (excluding methods utilized), and the contextualization of outcomes found in the studies solidify the need for further work. As a result of these findings, important questions have been asked in this thesis concerning how research on evaluation can inform better practice such as: If we desire to empower stakeholders and build what Carden defines as an indigenous evaluative culture, how can we use research on evaluation to inform a practice that supports this outcome? I believe that the answer to this questions lies in asking ourselves several key questions as brought forward by this thesis: Are CAE that are intended to empower individuals viable in local contexts that do not currently support empowerment? With that answer in mind, what is the role of transformative CAE in these contexts? Who must be involved in the empowerment process to ensure long-lasting empowerment outcomes are attained? What is the role of the evaluator in decision making, especially in regard to stakeholder diversity and participation, in the context of politically charged environments? How can we further integrate the principles of empowerment into our reporting on the research, both in terms of the evidence presented and authorship?

Based on this study, I believe that the work that lies ahead for evaluators resides in further understanding our roles and responsibilities in these challenging contexts. Answers to the questions posed above and a fulsome discussion of our roles and how our work impacts local power dynamics is largely missing from the current literature. It is imperative for the evaluation community to engage with these issues to ensure that we are indeed moving toward the positive empowerment of those who entrust us with these processes.

References

- Abes, R.V. (2000). Strengthening citizen participation and democratization in the Philippines: ELF's impact evaluation. In Estrella, M., Blauert, J., Campilian, D., Gaventa, J., Gonsalves, J., Guijt, I., . . . Ricafort, R. (Eds.). (2000). *Learning from change: Issues and experiences in participatory monitoring and evaluation*. Ottawa: International Development Research Centre.
- Alzate, R.D.E. (2000), Monitoring and evaluating local development through community participation: The experience of the AICNC, Colombia. In Estrella, M., Blauert, J., Campilian, D., Gaventa, J., Gonsalves, J., Guijt, I., . . . Ricafort, R. (Eds.). (2000). *Learning from change: Issues and experiences in participatory monitoring and evaluation*. Ottawa: International Development Research Centre.
- Angen, M.J. (2000). Evaluating interpretive inquiry: Reviewing the validity debate and opening the dialogue. *Qualitative Health Research*, 10, 378-395.
- Bi Niba, M and Green, J. M. (2005). Major Factors Influencing HIV/AIDS Project Evaluation. *Evaluation Review*, 29(4), 313–330. doi:10.1177/0193841X05276654
- Blauert, J & Quintanar, E (2000). Seeking Local Indicators: Participatory Stakeholder Evaluation of Farmer-to- Farmer Projects, Mexico. In Estrella, M., Blauert, J., Campilian, D., Gaventa, J., Gonsalves, J., Guijt, I., . . . Ricafort, R. (Eds.). (2000). *Learning from change: Issues and experiences in participatory monitoring and evaluation*. Ottawa: International Development Research Centre.
- Botcheva, L., Shih, J., & Huffman, L. C. (2009). Emphasizing Cultural Competence in Evaluation: A Process-Oriented Approach. *American Journal of Evaluation*, 30(2), 176–188.
- Brandao, D.B., Silva, R.R., & Cudas, R. (2012). Youth Participation in evaluation: The Pro-Menino Program in Brasil. In S. Kushner & Rotondo (Eds.), *Evaluation voices from Latin America. New Directions for Evaluation*, 134, 49-59
- Brandon, P.R., Smith, N.L., Ofir, Z. & Noordeloos, M. (2014). Monitoring and evaluation of African women in agricultural research and development: An exemplar of managing for impact in development evaluation. *American Journal of Evaluation*, 35(1). 128-143.

- Brisolara, S. (1998). The history of participatory evaluation and current debates in the field. In E. Whitmore (Ed.), *Understanding and practicing participatory evaluation. New Directions in Evaluation, No. 80* (pp. 25-41). San Francisco: Jossey Bass.
- Carden, F. (2007). The real evaluation gap. *Alliance Magazine*, 12,4
- Carden, F. (2010). Introduction to the forum on evaluation field building in South Asia. *American Journal of Evaluation, 31*(2), 291-221.
- Christie, C.A. (2003). What guides evaluation? A study of how evaluation practice maps onto evaluation theory. In C.A. Christie (Ed.), *The practice-theory relationship. (New Directions for Program Evaluation, No. 97)* (pp. 7-36). San Francisco, CA: Jossey-Bass.
- Cornwall, A. (2014). Using Participatory Process Evaluation to Understand the Dynamics of Change in a Nutrition Education Programme. *IDS Working Paper, 2014*(437).
- Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation, 1998*(80), 5-23. doi:10.1002/ev.1114
- Cousins, J.B., Goh, S., Clark, S., & Lee, L. (2004). Integrating evaluative inquiry into the organizational culture: A review and synthesis of the knowledge base. *Canadian Journal of Program Evaluation, 19*(20), 99-141.
- Cousins, J. B. (2005). Will the real empowerment evaluation please stand up? A critical friend perspective. In D. Fetterman & A. Wandersman (Eds.), *Empowerment evaluation principles in practice* (pp. 183-208). New York: Guilford.
- Cousins, J.B & Chouinard, J.A. (2010). A Review and Integration of Empirical Research on Participatory Evaluation. University of Ottawa, Canada.
- Cousins, J.B & Chouinard, J.A. (2012). Participatory Evaluation Up Close: An Integration of Research-based Knowledge. IAP: Charlotte
- Cousins, J. B., Whitmore, E., & Shulha, L. (2013). Arguments for a Common Set of Principles for Collaborative Inquiry in Evaluation. *American Journal of Evaluation, 34*(1), 7-22. doi:10.1177/1098214012464037
- Creswell, J. (2007). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage.

- Dablestein, N., & Patton, M.Q. (2013). Lessons Learned and the Contributions of the Paris Declaration Evaluation to Evaluation Theory and Practice. *Canadian Journal of Program Evaluation*, 27(3), 173-200.
- Dageid, W., & Duckert, F. (2007). The process of evaluation a capacity-building support initiative for HIV positive South Africans. *Evaluation and Program Planning*, 30, 381-391.
- Donnelly, J. (2010). Maximizing participation in international community-level project evaluation: a strength-based approach. *Evaluation Journal of Australasia*, 10(1), 43-50.
- Fetterman, D., & Wandersman, A. (2007). Empowerment Evaluation: Yesterday, Today, and Tomorrow. *American Journal of Evaluation*, 28(2), 179-198.
doi:10.1177/1098214007301350
- Greene, J. C. (2013). On rhizomes, lines of flight, mangles, and other assemblages. *International Journal of Qualitative Studies in Education*, 26(6), 749-758.
- Gregory, A. (2000). Problematizing Participation: A Critical Review of Approaches to Participation in Evaluation Theory. *Evaluation*, 6(2) 179-199.
- Guba, E.G., & Lincoln, Y.S. (1989). Fourth generation evaluation. London: Sage.
- Hamilton, C., Rai, R.K., Shrestha, R.B., Maharjan, M, Rasaily, L & Hood, S. (1998). Exploring visions: Self-monitoring and evaluation processes within the Nepal-UK Community Forestry Project. In Estrella, M., Blauert, J., Campilian, D., Gaventa, J., Gonsalves, J., Guijt, I., . . . Ricafort, R. (Eds.). (2000). *Learning from change: Issues and experiences in participatory monitoring and evaluation*. Ottawa: International Development Research Centre.
- Hart, R. A., & Rajbhandary, J. (2003). Using Participatory Methods to Further the Democratic Goals of Children's Organizations. *New Directions for Evaluation*, 2003(98), 61-75. doi:10.1002/ev.85
- Holte-McKenzie, M., Forde, S., & Theobald, S. (2006). Development of a participatory monitoring and evaluation strategy. *Evaluation and Program Planning*, 29(4), 365-76.
doi:10.1016/j.evalprogplan.2006.08.007

- Jupp, D., Ali, S. I., & Barahona, C. (2010). *Measuring Empowerment ? Ask Them Quantifying qualitative outcomes.*
- Kariuki, J., & Njuki, J. (2013). Using participatory impact diagrams to evaluate a community development project in Kenya. *Development in Practice*, 23(1), 90–106. doi:10.1080/09614524.2013.753031
- Lather, P. (1986). Issues of validity in openly ideological research: Between a rock and a soft place. *Interchange*, 17(4), 63–84. doi:10.1007/BF01807017
- Letichevsky, A.C., & Penna Firme, T. (2012). Evaluation with at-risk communities: Learning from a social program in a Brazilian slum. In S. Kushner & Rotondo (Eds.), *Evaluation voices from Latin America. New Directions for Evaluation*, 134, 61-76
- Lincoln, Y. S., & Guba, E. G. (1985) *Naturalistic Inquiry*. New York: SAGE.
- Lincoln, Y. S., & Guba, E. G. (1986). But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. In D. D. Williams (Ed.), *Naturalistic Evaluation, New Directions for Program Evaluation*, No. 30 (pp. 73-84). San Francisco: Jossey-Bass.
- Miller, R. L., & Campbell, R. (2006). Taking Stock of Empowerment Evaluation An Empirical Review. *American Journal of Evaluation*, 27(3), 296–319. doi:10.1177/1098214006291015
- Miller, R. L., & Campbell, R. (2007). Taking stock again results in the same conclusions: A reply to Fetterman and Wandersman's defense of empowerment evaluation. *American Journal of Evaluation*, 28(4), 579-581.
- Mompati, T., & Prinsen, G. (2000). Ethnicity and participatory development methods in Botswana: Some participants are to be seen and not heard. *Development in Practice*, 10(5). 625-637.
- Motteaux, N, Binns, T., N. Etienne, and Rowntree, K. (1999). Empowerment for Development: Taking Participatory Appraisal Further in Rural South Africa. *Development in Practice*. 9(3), 261-273.
- Newman, K. (2008). Whose view matters? Using participatory processes to evaluate Reflect in Nigeria. *Community Development Journal*, 43(3), 382–394. doi:10.1093/cdj/bsn015

- Parkinson, S. (2009). Power and perceptions in participatory monitoring and evaluation. *Evaluation and Program Planning*, 32, 229-237.
- Patton, M.Q. (2005) Toward distinguishing Empowerment Evaluation and placing it in a larger context: take two. *American Journal of Evaluation*, 26(3). 408-414
- Rog, D.J., Fitzpatrick, J.L., & Conner, R.F. (Eds.) (2012). Special Issue: Context: A Framework for Its Influence on Evaluation Practice. *New Directions for Evaluation*, 2012(135).
- Segone, M. (Ed.). (2006). New trends in development evaluation. *UNICEF Working Papers*, Issue #5.
- Segone, M. (Ed.). (2008). Country-led monitoring and evaluation systems: Better evidence, better policies, better development results. *UNICEF Evaluation Working Papers* #19.
- Scriven, M. (1997). Empowerment evaluation examined. *Evaluation Practice*, 18(2), 165-175.
- Schulz, A. J., Israel, B. A., Zimmerman, M. A., & Checkoway, B. N. (1995). Empowerment as a multi-level construct: Perceived control at the individual, organizational and community levels. *Health Education Research*, 10(3), 309-327.
- Smith, N. L. (1999). A framework for characterizing the practice of evaluation, with application to empowerment evaluation. *Canadian Journal of Program Evaluation* [Special Issue], pp. 39-68.
- Smith, N. L. (2007). Empowerment evaluation as evaluation ideology. *American Journal of Evaluation*. 28(2), 169-178.
- Symes, J & Jasser, S. (2000). Growing from the grassroots: Building participatory planning, monitoring and evaluation methods in PARC. In Estrella, M., Blauert, J., Campilian, D., Gaventa, J., Gonsalves, J., Guijt, I., . . . Ricafort, R. (Eds.). (2000). *Learning from change: Issues and experiences in participatory monitoring and evaluation*. Ottawa: International Development Research Centre.
- Taylor, C. (2004). *Modern social imaginaries*. Durham, NC: Duke University Press.
- Torres D, V.H (2000). Monitoring local development with communities: The SISDEL approach in Ecuador. In Estrella, M., Blauert, J., Campilian, D., Gaventa, J.,

- Gonsalves, J., Guijt, I., . . . Ricafort, R. (Eds.). (2000). *Learning from change: Issues and experiences in participatory monitoring and evaluation*. Ottawa: International Development Research Centre.
- Turnbull, A. (2000). Collaboration and censorship in the oral history interview. *International Journal of Research Methodology*, 3(1), 15-34
- Vernooy, R., Qiu, S., and Jianchu, X. (2006). The power of participatory monitoring and evaluation: insights from south-west China. *Development in Practice*, 16(5).
- Ward, P. (2000). Getting the right end of the stick: Participatory monitoring and evaluation in an organizational context. In Estrella, M., Blauert, J., Campilian, D., Gaventa, J., Gonsalves, J., Guijt, I., . . . Ricafort, R. (Eds.). (2000). *Learning from change: Issues and experiences in participatory monitoring and evaluation*. Ottawa: International Development Research Centre.
- Wood, B., & Betts, J. (2012). Results of the Paris Declaration Evaluation. *Canadian Journal of Program Evaluation*, 27(3), 101–126.
- World Bank. (2005). Measuring Empowerment. Eds. Deepa Narayan. Washington DC.
- Worthington, C. (1999). Empowerment evaluation: Understanding the theory behind the framework. *Canadian Journal of Program Evaluation*, 14(1), 1-28.