

2012

Theory Building Through Praxis Discourse: A Theory- And Practice-Informed Model of Transformative Participatory Evaluation

Michael Allen Harnar
Claremont Graduate University

Recommended Citation

Harnar, Michael Allen, "Theory Building Through Praxis Discourse: A Theory- And Practice-Informed Model of Transformative Participatory Evaluation" (2012). *CGU Theses & Dissertations*. Paper 57.
http://scholarship.claremont.edu/cgu_etd/57

DOI: 10.5642/cguetd/57

This Open Access Dissertation is brought to you for free and open access by the CGU Student Scholarship at Scholarship @ Claremont. It has been accepted for inclusion in CGU Theses & Dissertations by an authorized administrator of Scholarship @ Claremont. For more information, please contact scholarship@cuc.claremont.edu.

Theory Building Through Praxis Discourse:
A Theory- And Practice-Informed Model Of Transformative Participatory Evaluation

By
Michael A. Harnar

A Dissertation submitted to the Faculty of Claremont Graduate University in partial fulfillment
of the requirements for the degree of Doctor of Philosophy in Psychology

Claremont, California

June 2, 2012

Copyright 2012 by
Harnar, Michael Allen

We, the undersigned, certify that we have read this dissertation of Michael A. Harnar and approve it as adequate in scope and quality for the degree of Doctor of Philosophy.

Dissertation Committee:

Dr. Christina A. Christie, Co-Chair

Dr. Stewart I. Donaldson, Co-Chair

Dr. Tarek Azzam, Member

Dr. David M. Fetterman, Visiting Examiner

Abstract of the Dissertation

Theory Building Through Praxis Discourse:

A Theory- And Practice-Informed Model Of Transformative Participatory Evaluation

By

Michael A. Harnar

Claremont Graduate University: 2012

Stakeholder participation in evaluation, where the evaluator engages stakeholders in the process, is prevalent in evaluation practice and is an important focus of evaluation research. Cousins, Donohue, and Bloom describe a utilization-focused form of participatory research and evaluation, where engagement is a means to increase use, and Cousins and Whitmore propose a bifurcation of participatory evaluation into the two streams of transformative participatory and practical participatory evaluation (T-PE and P-PE respectively). T-PE stems from a social justice perspective and P-PE has more of a use orientation. Transformative participatory evaluation is an underdeveloped evaluation theory with relatively low operational specificity. Case examples provide some understanding of it in practice, but comprehensive empirical support is still forthcoming. An increased focus on stringent internal validity and replicable causation advanced by the National Research Council of the National Academy of Sciences has the potential to shift the focus of evaluation further from a learning approach and more towards one of accountability. While accountability is an important aim of evaluation, it often sacrifices its potential to teach and build capacity, both of which are central to participatory approaches of evaluation. Participatory approaches are vital to the field and it is critical that they be further developed and defined to provide practitioners with substantiated alternatives to accountability models.

This study aims to develop a greater understanding of the participatory evaluation schema of P-PE and T-PE and to develop more practice-based and accessible operational specificity of transformative participatory evaluation by developing a logic-model like representation of T-PE informed by both theorists and practitioners. In the process, a set of 28 key T-PE variables and eight statements that help identify T-PE evaluators were developed with the help of an expert panel of evaluation theorists.

The sample engaged in the research was the American Evaluation Association's membership (N=6,615). Each was invited to an online survey where they were asked their agreement on eight statements related to participatory evaluation. If they were at all participatory in their approach to evaluation they were asked to model their evaluation practice using an online software. A total of 240 evaluators modeled their practice. A most-endorsed model was created from the drawings of those identified as T-PE evaluators (n=142) and a sample of them (n=21) commented on this model through webinars.

The model created in this research is quantitatively and qualitatively different from a model created by a group of practitioners identified as more utilization-focused (n=16). The T-PE model was more likely to have stakeholder involvement and community trust at its center and the comparison model was more action-oriented and outcomes driven. This theory- and practice-informed T-PE model, the set of variables expected to be key to T-PE, and the set of statements that might be used to identify T-PE evaluators from other practitioners provide for a more descriptive theory of transformative participatory evaluation and introduce a novel method for engaging practitioners in the theory development process.

Dedication

This dissertation is dedicated to my grandparents Helen and Fred Harnar. While they always supported my choice of joining the Navy after high school I know they would be even more proud to see me reach this milestone. Helen was an English teacher and both earned undergraduate degrees from Hiram College in the 1930s. From leadership roles in Rotary, the local Women's Club, and our church, to loving parents and grandparents, their example has been priceless.

Acknowledgements

How does one actually acknowledge all those who made something like this possible? One cannot acknowledge only those who supported the dissertation because the dissertation is the culmination of so many other endeavors that it would feel incomplete to just acknowledge the help in accomplishing this single product. I will try to do this piece justice, but it might get a bit long winded.

Dr. Deborah Vietze introduced me to evaluation in her class entitled Program Evaluation and the Logic of Inquiry. She provided a small window into what I've come to understand as a discipline, my discipline. Ten years later I'm publishing this dissertation. While it would be poor science to say she caused this moment, as with so many mentoring and teaching experiences, she was a major influence on my finding a place in this field. I feel that she can own a piece of this success.

There is no way that I could have accomplished this program without the supportive embrace of Dr. Christina Christie. During my first visit to Claremont eight years ago Tina presented an honesty and forthrightness that I wanted and needed from a mentor. She made the entire process accessible and provided clarity of scope, vision, and purpose. As a role model for evaluation research and practice she is an exemplar. One particular moment in this long journey stands out as critical to my success. At about halfway through my first year in the program I shared with her that I did not know if this was the right path for me. The classes were daunting and despite Dr. Dale Berger's wonderful teaching I felt like I had bitten off more than I was capable of digesting. She said two things that I will always be grateful for. "Stick it out till the end of the first year and reassess" and "of all the people in this program that I worry won't make it, you are not one of

them.” I stuck it out that first year and at that point felt like I had committed so much time and money that I decided I could not quit (her strategy worked there!). As for the latter encouragement, her confidence in me was priceless. For her to have that level of confidence in me was humbling and motivating. Truly, I could not have done this without you. Thank you!

This research would have been impossible without the support of the three theorists most familiar with this line of inquiry. As you will see as you read, all three provided their time, expertise, and insight willingly and abundantly. Also, Dr. John Gargani contributed a tremendous amount of time and effort to this research. He openly provided his software for my project and was so very patient in helping me understand the analyses best suited for the data. I am sure that Rasch modeling would have driven me crazy if not for John’s patient explanation(s). I look forward to where he takes this software next.

The faculty at Claremont have all been extremely supportive and made this journey possible. Dr. Crano helped me get my research legs with my Master’s thesis and his wife Suellen (another Dr. Crano) has been a wonderful cheerleader. Dr. Pezdek encouraged the athletic side of me and supported my starting of the CGU Sports Club so that we could play volleyball on the Pomona College sand courts. Dr. Donaldson and Dr. Azzam were regularly available for counsel and support. And Dr. Berger’s numbers wizardry always teased those parts of my brain that longed to think in equations (which admittedly is a small piece of my brain!). I will always be grateful for Dale’s open door and accessible mind. Though Paul Thomas and his wife have moved on to other adventures, his presence in SBOS was incredibly valuable. I arrived in Claremont just after my 40th birthday and Paul’s adult presence was appreciated around so many of the ‘fresh young faces’ of my cohort.

Those of you who know me know that friends and family are critically important to my happiness. Here at Claremont, Kile and Kari Dyer and Max Freund and his wife Cynthia Luna, have been my ever-present stability, providing the loving kindness I needed. They shared in every joy and every disappointment of this journey. I treasure beyond words their loving support and their friendship, without which I would have fled Claremont ages ago and finished this PhD from afar, probably New York City. Speaking of New York, my NYC posse has been tremendously supportive of my ‘adventure to the left coast’ and I so appreciate all they have done to support me: Bill, Hillary, Eric, Donna, Bri, Katherine, Sonia. The list could go on. Thanks for remaining steadfast, despite my leaving you and our fair city. Meagan Murphy Forget, George Cole, Randy Kulp, your support and confidence in me, and your friendship are invaluable.

As for family, no relative has been more central to my success throughout my *life* than my uncle Jim. James Harnar is a force in his own right, a retired Navy Captain, Chief Operating Officer of a Foundation doing important healthcare work in Maine and a loving husband and father to two wonderful women. In all of this he has consistently made time for me. He participated in every major accomplishment in my Navy career and has been an avid supporter of this latest endeavor, always available for a supportive conversation. I think of him not as an uncle, but as a friend, a brother, and a role model. Jim, you truly do your parents proud and I am so grateful for your love and friendship.

My parents and their siblings, and my brothers and their families, while a little perplexed by this undertaking, support me and consistently lift me up rather than tear me down. Home truly is where the heart lies, and it lies with you.

Table of Contents

Abstract of the Dissertation	iv
Dedication	xii
Acknowledgements	vii
Table of Contents	x
List of Tables	xiv
List of Figures.....	xvi
CHAPTER 1	1
INTRODUCTION TO THE STUDY	1
Problem Statement.....	6
Research Questions.....	8
Purpose Statement	11
CHAPTER 2	12
REVIEW OF THE RELEVANT LITERATURE	12
An Introduction to Evaluation Theory	12
Development of Participatory Evaluation Theory.....	15
Control of Evaluation	18
Selection of Participants	19
Depth of Participant Involvement	20
Transformative Participatory Evaluation	20
Transformative Evaluation	22
Theory Driven Evaluation and Logic Models	25
Summary.....	31
CHAPTER 3	33
OVERVIEW OF METHODOLOGY AND PROCEDURES & PHASE ONE (EXPERT PANELISTS)	33
Overview	33
Overall Study Methodology	34
Mixed Method Theories, Mental Models, & Paradigms	34
Phase One: Expert Panelists	37
Phase One Methodology.....	37
Participants.....	37
Procedures.....	39
Variables Development Results	45
Principles	47
Activities.....	50
Outcomes	53
Identifying Statements Results	60

CHAPTER 4.....	62
PHASE TWO (SURVEY AND MODELING).....	62
Phase Two Methodology.....	62
Participants	62
Survey Sample	65
Model Sample	69
Procedures: Pilot Survey	80
Instruments: Survey	82
Instruments: Modeling Procedure.....	84
Analyses	85
Phase Two Results.....	86
T-PE Questions Results	86
Variables Results	93
Variable Endorsement By All Modelers.....	93
Variable Endorsement by T-PE Evaluators	95
Relationships Between Variables	95
T-PE Evaluators' Relationships Between Variables.....	98
Variable Usage by Practitioner Groups	99
Participant Proffered Variables Results.....	100
Model Results	103
T-PE Model.....	105
Comparative Analysis	108
CHAPTER 5.....	115
PHASE THREE (WEBINARS).....	115
Phase Three Methods	115
Participants	115
Procedures	121
Instrument.....	122
Webinar Results	122
Community Trust.....	122
Shared Understanding	124
Credible Findings	125
Build Capacity	126
Engaging Intended Beneficiaries.....	127
Social Justice	128
Impact of/on Evaluators	129
Summary of Webinar Comments	129
CHAPTER 6.....	132
RESULTS SUMMARY	132
Phase One Results.....	132
Phase Two Results.....	134
Phase Three Results.....	138
CHAPTER 7.....	140
DISCUSSION OF THE FINDINGS	140
Interpretation of Findings.....	142
Research Question 1: What are the key variables of T-PE?.....	142

Research Questions 2 and 3: What are the most important variables in evaluators' practice and how do they model it? How does evaluators' practice relate to the model?	142
Research Question 4: How does the T-PE model compare to other evaluation theories?	145
Connections to Existing Knowledge	147
Implications	155
Theoretical Implications	155
Practice Implications	157
Educational Implications	159
Limitations	159
Future Research	161
Conclusions	162
References	165
APPENDIX A:	182
Online Editing Instructions To Theorists	182
Front Page Instructions on Google Wiki for Editing both Variables and Statements	182
Instructions for Editing Variables Document on Google Wiki	182
Instructions for Statement Editing on Google Wiki	183
APPENDIX B:	184
Preliminary Variables And Definitions	184
APPENDIX C:	187
Revised Variables And Definitions (After Phase One)	187
APPENDIX D:	189
Preliminary Statements	189
APPENDIX E:	191
Revised Statements (After Phase One)	191
APPENDIX F:	192
E-Mail Invitation For Pilot Study	192
APPENDIX G:	194
Final Survey Instrument	194
Statements (After Phase One & Pilot Testing)	203
APPENDIX I:	205
Survey Invitation	205
APPENDIX J:	207
Survey Reminder	207
APPENDIX K:	209
Modeling Reminder	209
Modeling Interface	210
APPENDIX M:	214
Webinar Invitation	214
APPENDIX N:	215
Webinar Protocol	215
APPENDIX O:	220

Final Variables Set.....	220
APPENDIX P	222
Variables and Statements Editing Evolution	222

List of Tables

Table 1.	Number of Theorist Comments During Variable Development.....	43
Table 2.	Comment Types During Variable Development.....	43
Table 3.	Examples of Editing Process During Variable Development.....	44
Table 4.	Final 26 Variables and Definitions.....	46
Table 5.	Number of Theorist Comments During Statement Development.....	57
Table 6.	Comment Types During Statement Development.....	57
Table 7.	Example of Editing Process During Statement Development.....	58
Table 8.	Eight Identifying Statements.....	60
Table 9.	Survey Response Rates.....	64
Table 10.	Survey Participants' Characteristics.....	66
Table 11.	Three Statements Used to Identify Participatory Evaluators.....	70
Table 12.	Participants' Participatory Evaluation Categories.....	70
Table 13.	Modeling Participants' Characteristics.....	71
Table 14.	Added Variable Participants' Characteristics.....	76
Table 15.	Survey Pilot Comments.....	80
Table 16.	Participant Response to T-PE Statements.....	88
Table 17.	T-PE Items Correlation to Convergent Items.....	89
Table 18.	T-PE Items Coefficient Alpha.....	90
Table 19.	Differences in Theoretical Preference by PE Groupings.....	91
Table 20.	Comparisons Between PE Groupings on Two Theoretical Preferences.....	91
Table 21.	Proffered Additional Variables.....	101
Table 22.	Arrows in Models by Participant Grouping Categories.....	103
Table 23.	Percent of T-PE Modelers Selecting Top 21 Links.....	107

Table 24.	Ranking of P-PE Modelers Top 38 Links.....	109
Table 25.	Variables More Likely to be Endorsed by T-PE Modelers than P-PE Modelers...	111
Table 26.	Variables less likely to be endorsed by T-PE modelers over P-PE modelers.....	112
Table 27.	Webinar Dates and Participants.....	116
Table 28.	Webinar Participants' Characteristics.....	118
Table 29.	Final 26 Transformative Participatory Evaluation Variables.....	132
Table 30.	Transformative Participatory Evaluation Identifying Statements.....	133
Table 31.	Modelers' Participatory Evaluation Categories.....	135
Table 32.	Variables More or Less Likely to be Endorsed by T-PE Modelers than P-PE Modelers.....	136
Table 33.	Comparison of UCLA P-PE Model to Current T-PE Model.....	150
Table 34.	Variables Unique to Transformative Participatory Evaluation.....	151

List of Figures

Figure 1.	Dimensions Of Form In Collaborative Inquiry.....	18
Figure 2.	Mixed Method Design Elements And Participant Samples.....	36
Figure 3.	Endorsement of Principles, Activities, and Outcomes.....	94
Figure 4.	T-PE Evaluator Endorsement of Principles, Activities, and Outcomes.....	95
Figure 5.	All Modelers' Relative Endorsement of Principles, Activities, and Outcomes.....	97
Figure 6.	T-PE Relative Endorsement of Principles, Activities, and Outcomes.....	98
Figure 7.	Principles, Activities, and Outcomes Endorsement Percentage of Expected Value.....	100
Figure 8.	Histogram of Arrows in T-PE Models.....	104
Figure 9.	Histogram of Arrows in P-PE Models.....	104
Figure 10.	Histogram of Arrows in PE Models.....	105
Figure 11.	Transformative Participatory Evaluation Model.....	106
Figure 12.	Social Justice As Outcome Model.....	113
Figure 13.	Practical Participatory Evaluation Model by Hansen, Alkin, and Wallace.....	148

CHAPTER 1

INTRODUCTION TO THE STUDY

Stakeholder participation in evaluation, where the evaluator engages stakeholders in the process, is prevalent in evaluation practice (Fleischer & Christie, 2009; Preskill & Caracelli, 1997) and is an important focus of evaluation research (e.g., Cousins & Earl, 1992, 1995; Cousins & Whitmore, 1998; Cousins, Donohue, & Bloom, 1995, 1996; Weaver & Cousins, 2004). The Collaborative, Participatory, and Empowerment Topical Interest Group of the American Evaluation Association (AEA) includes approximately 25% of the more than 6,700-strong association membership (AEA TIG website, 2012). A recent Web of Science¹ title search for “participatory evaluation” revealed 94 articles published between 2002 and 2012 in journals as diverse as *Land Use Policy* and the *Journal of Cancer Education*. And a title search in the same period using Google Scholar, which casts a wider net, returned 702 books, articles, and other scholarly works.

Engaging stakeholders in research and evaluation serves at least three primary goals. Pragmatically, those involved will find the research design and findings more aligned with their beliefs and will feel more ownership of and be more likely to use the results of the evaluation (e.g., Patton, 1997, 2008; Weaver & Cousins, 2004). Politically, involvement of stakeholders who normally do not have a voice can strengthen their position and be emancipatory in its aims (e.g., Mertens, 2009). From an epistemological perspective, engaging stakeholders develops an understanding grounded in the lived experiences of those involved in the program being studied

¹ Web of Science searches these databases simultaneously: Science Citation Index Expanded (SCI-EXPANDED); Social Sciences Citation Index (SSCI); Arts & Humanities Citation Index (A&HCI); Conference Proceedings Citation Index- Science (CPCI-S); Conference Proceedings Citation Index- Social Science & Humanities (CPCI-SSH).

(e.g., Blackstock, Kelly, & Horsey, 2007; Brandon, 1998; Chen, Poland, & Skinner, 2007; Whitmore, 1994).

Cousins and Earl (1992), using an organizational learning perspective, described the process of collaborative evaluation and research through three key characteristics: (a) attention to who maintains *control* over technical decision-making of the evaluation, (b) the *selection* of key participants, and (c) the *depth* of their involvement. This framework was the foundation for other empirical research by Cousins, Donohue, and Bloom (1995, 1996) that describes a utilization-focused form of participatory research and evaluation where engagement is a means to increase use (see also Cousins & Earl, 1995). Cousins and Whitmore (1998) further built on this schema by using it to propose a bifurcation of participatory evaluation (PE) into the two streams of transformative participatory and practical participatory evaluation (T-PE and P-PE respectively). T-PE stems from a social justice perspective and P-PE has more of a use orientation.

To date, there is limited empirical research supporting Cousins and Whitmore's (1998) definition of transformative participatory evaluation. Cousins and Chouinard (2012) reviewed the evaluation literature from 1997 to 2011 and found support for the practice of PE in many of its forms, including T-PE (14% of those items reviewed). Of the 17 T-PE studies they reviewed, most (77%) were small-scale case examples of single evaluations and none provided insight into more than a few evaluations. Relative to other forms of PE (e.g., practical participatory, empowerment, collaborative), this is a paucity of empirical research. The definition for T-PE is therefore dependent upon only Cousin and Whitmore's supposition (1998). Beyond the Cousins and Choinard book, there is limited independent research focused on understanding T-PE in

practice. It is therefore a form of evaluation with a prescriptive theory that lacks substantial empirical support.

Cousins and Earl's (1992) schema of control, selection, and depth— notions Cousins and Whitmore (1998) used to help define differences between T-PE and P-PE—is relevant to participatory evaluation, but its utility in practice is limited because it must simplify very complex situations and overgeneralize multiple points of data that occur across various contexts (Alkin, 2004b). For instance, it would be difficult to separate the control one exercises in an evaluation from the selection or make-up of the group of stakeholders an evaluator is about to share control with; the capabilities of any given set of stakeholders selected to participate will have important implications for how deeply they are involved in the evaluation. In fact, Weaver and Cousins (2004), while using this schema to interpret their evaluation experiences, admit that these dimensions would be difficult to separate empirically, even though they serve well as heuristics to understand different dimensions.

Given these limitations, Weaver and Cousins argued for the selection of stakeholders dimension to be reconceptualized into three dimensions: the *manageability* of the different groups, the *power distribution* among stakeholders, and the relative *diversity* of participating stakeholders. Daignault and Jacob (2009), using Gerring's (1999) concept definition and Goertz's (2006) conceptualization framework, found this further hewing unnecessary; they posit that the simplified dimensions of depth, selection, and control adequately define the concept.

Schemas are valuable in the discipline's theory development (Akin, 2004a, in press). In the interplay between theory and practice, these more overarching descriptions of complex processes provide an opportunity to investigate the nuances between and among the application of theories

in practice. By providing heuristics that reflect much deeper meaning, they further establish a language of evaluation and add to the discipline's foundation.

Clarification of theoretical models and schemas is therefore important to discipline building (e.g., Alkin, 1979, 1991; Christie, in press). Ideally, theories are developed and refined in a recursive process, drawing on practice and studies of practice to inform theory building. For instance, Patton and colleagues (1975, 1977), in their study of federal health evaluation research usage, discovered an unexpected factor that is key to evaluation utilization: "the personal factor." Despite their attempts to gather information on 11 factors theorized to effect greater use, interviewees consistently mentioned the involvement of key personnel in the evaluation as the single most important factor. This unexpected finding catalyzed a proliferation of inquiry into the impact of personal involvement by both evaluators and key decision makers, and was the genesis of utilization focused evaluation (e.g., Patton, 1978, 1986, 1997, 2008).

Research into utilization also led to a *use* framework (instrumental, conceptual, political or symbolic) (Weiss, 1979), a discussion of evaluation *influence* (e.g., Henry & Mark, 2003; Kirkhart, 2000; Morabito, 2002), and the sensitizing concept of *process use* (e.g., Amo & Cousins, 2004; Harnar & Preskill, 2007; Patton, 2007; Preskill, Zuckerman, & Matthews, 2003). The recursive nature of theory building continues as these concepts move beyond the pages of journals into textbooks and guidebooks (e.g., Fitzpatrick, Sanders, & Worthen, 2010; Patton, 2008; Rossi, Lipsey, & Freeman, 2004) and eventually into practice.

A schema's inherent complexity is also its shortcoming: it overgeneralizes complex processes, making it difficult to identify its application in practice (e.g., see Harnar & Preskill, 2007). Empirical study of schemas in practice helps connect them to reality. For instance, studies

of evaluation practice framed by these complex heuristics provide inductive and grounded lessons that make the broader categorizations more understandable (Birckmayer & Weiss, 2000; Jacob, Ouvrard, & Bélanger, 2011). In-depth focus can also show how weak the relationship between theory to practice can be (e.g., Fetterman, 1995; Miller & Campbell, 2006; Scriven, 1997; Stufflebeam, 1994). The metaphor that describes qualitative data as providing the “flesh and blood” for a quantitative data “skeleton” is useful here, as well: categorization schemas and theories provide frameworks around which we understand the discipline of evaluation; studies of practice organized around these frameworks are closer to the living practice and give the theory a heartbeat.

For students of the evaluation discipline, learning various theories and their nuanced differences can be a challenging undertaking. Schemas can help situate and frame both theories and practice in ways that serve to scaffold learning. Their simplicity is helpful because it puts very complex processes into manageable frameworks. And while experienced academics and theorists—and even advanced students—may work with complex schemas and theories with relative ease, busy practitioners generally have less time to stay immersed in the language and nuances of the evaluation literature. This makes it challenging for them to interpret and apply theoretical writings. More precisely, it can be difficult to identify and specify the factors that impact practice when one is deeply engaged in that practice (V. Magar, personal communication, 2009).

Using tools that practitioners are familiar with to describe theory could serve as a powerful strategy to relate theory to practice and provide a middle ground between the two. Modeling a program’s resources, activities, outcomes, and assumptions—commonly known as logic

modeling or a log frame—is a familiar undertaking in evaluation design (Coryn, Noakes, Westine, & Schroter, 2011; Rossi et al., 2004; Williams, 2010) because it offers ways to think about and conceptualize the intentions of program practitioners in an accessible format. This approach is useful for many steps in an evaluation, from designing an evaluation plan and determining the questions to be addressed, to documenting a project and how it worked. A program theory logic model might be used at the outset of an evaluation cycle to gain a shared understanding of the program among the various stakeholders. It might also be used at the end to reflect on what the program “looked like” at the beginning and to describe its evolution.

Because logic modeling is a tool used in one form or another by many evaluators (Frechtling, 2007) and is a fundamental tool of theory driven evaluation, the process and products are generally familiar to practitioners; most understand how to dissect a logic model into its constituent parts and operationalize them as necessary. Therefore, translating the skills needed to model program practice to what is needed to model evaluation practice can bring evaluation practitioners further into the theory-building world. In turn, this process offers practitioners a foundation from which to interpret and apply evaluation theories in their own practice.

Problem Statement

Transformative participatory evaluation is an underdeveloped evaluation theory with relatively low operational specificity (Miller, 2010). Case examples provide some understanding of it in practice, but comprehensive empirical support has yet to be developed. Evaluation is an applied discipline where theories are developed from practice, applied in practice, and further developed by studying practice. Empirical study of transformative participatory evaluation in practice is needed to inform our broader understanding of participatory evaluation.

Given the strong policy position taken on random assignment, qualitative methods, and the nature of trustworthy knowledge by the National Research Council of the National Academy of Sciences (NRC, 2002, as cited in Gersten & Hitchcock, 2009; “What Works Clearinghouse”, n.d.) program managers understandably pressure evaluators to take more experimental approaches to evaluating their programs. This increased focus on stringent internal validity and replicable causation shifts the focus of evaluation further from a learning approach and more towards one of accountability. While accountability is an important aim of evaluation, it often sacrifices its potential to teach and build capacity, both of which are central to participatory approaches (Cousins, Whitmore, & Shula, in press). Despite these shifts in educational research policy, evaluators continue their commitment to stakeholder participation (Cousins et al., 1996; Fleischer & Christie, 2009; Preskill & Caracelli, 1997). Therefore, participatory approaches are vital to the field and it is critical that they be further developed and defined to provide practitioners with substantiated alternatives to accountability models.

While current research is predicated on participatory evaluation having two streams, T-PE and P-PE, there is a third that must also be considered: transformative evaluation (Mertens, 2009; Mertens & Wilson, 2012). The intersection of transformative with transformative participatory is a neglected yet important crossing. For instance, participation with transformative foci is well rooted in cultures outside North America (e.g., Chambers, 2010; Hickey & Mohan, 2004; "Institute for Development Studies", n. d.) and participatory forms of evaluation are practiced and studied throughout the world under names as diverse as participatory action research, practical participatory evaluation, participatory rural appraisal, youth participatory evaluation, and participatory monitoring and evaluation. Despite the approach's importance to evaluation practice, it is neglected in the North American evaluation literature and practitioners here find

limited support for learning about transformative forms of participatory evaluation. With her recent work, Mertens and Wilson (2012) have begun to align the two, but more needs to be done to understand this paradigm in practice and explore its relationship to the existing participatory evaluation literature.

Practitioners are rarely engaged in the theory building process. Theoretical writing, heuristic categorizations, and limited access to publications make engaging in the process challenging for non-academics. The ideal theory-building process would engage practitioners in using evaluation theories in their practice and provide access for studying their use. Using a tool evaluators are comfortable with—specifically, logic modeling—may make evaluation theory more accessible and bring more practitioners into the theory-building process. Logic modeling an evaluation theory may also develop a product practitioners are more familiar with and give them a useful tool with which to describe their own practice. Models of evaluation approaches might serve to bridge the gap between practice and theory by helping evaluators discuss what they do with potential clients in a conversation grounded in evaluation theory.

Research Questions

To gain a greater understanding of the participatory evaluation schema of P-PE and T-PE and to develop more practice-based and accessible operational specificity of transformative participatory evaluation in particular, this study asked these questions:

- What does a theorist-informed T-PE theory look like? What specific variables and principles do key evaluation theorists who are most familiar with T-PE agree should be a part of any model of T-PE?

- What does a practitioner-informed T-PE theory look like? When asked to model their preferred practice using a set of T-PE variables, which variables, and what relationships between the variables, would T-PE evaluators use to create such a model? What variables and relationships between variables would be included in a most-endorsed model of T-PE practice?
- Once the models created by the T-PE evaluators have been condensed into one most-endorsed model, how do T-PE evaluators see their practice reflected in this model? How do they describe the model and the variables included?
- How does a most-endorsed model of T-PE theory compare to other evaluation theories?

To address these questions, three theorists who are most familiar with T-PE—J. Bradley Cousins, Elizabeth Whitmore, and Donna Mertens—were engaged in creating a list of variables expected to be included in a model of T-PE practice. Concurrently, in this first phase of the study, the theorists helped develop statements that could be used to distinguish T-PE practitioners from other participatory practitioners.

In the second phase of the study, the American Evaluation Association membership was surveyed and their levels of agreement with the above statements were used to categorize respondents as T-PE evaluators or other participatory evaluators. If participants reported involving stakeholders to some degree they were invited to take part in an online model-building exercise where the above-mentioned variables were provided for them so that they could build a practice-based model. All completed models were then combined to produce one most-endorsed, representative model.

In the third and final phase of the study, participants from this model-building sample who identified as ascribing to transformative evaluation practice were invited to participate in webinars where they discussed how the coalesced model reflects their practice. Finally, the model was compared to models of other evaluation theories. The product of this study is a theory- and practice-informed model of transformative participatory evaluation practice with a qualitative component that further explains some of its elements.

Two streams of evaluation literature guided this study's theoretical framework. First, the program theory-driven evaluation (TDE) work of Chen (1990, 2005), Rossi, Lipsey, & Freeman (2004), and Donaldson (2007) informed the choice of model building. Second, an a-paradigmatic mixed method framework (Greene, 2007) was used to build the research's knowledge claims. This framework, where mixed methods capitalize on the strengths of both positivist and constructivist paradigms (Greene, 2007; Guba, 1990), uses the strengths (and compensates for the weaknesses) of each paradigm. It is not tied to the dogmatic focus of one paradigm and uses methods based on how well they serve the research question (Greene, 2007).

By treating evaluation practice as one might treat a program to be evaluated, this study mirrors the TDE procedure of explicating a program's theory through logic modeling. First, a literature review and discussions with theorists make explicit the implicit understandings of how T-PE works, producing the foundation for a prescriptive theory that includes principles, activities, and outcomes. Then, turning to practice, we see if the variables articulated by researchers are supported in practice when practitioners are asked to build a model of their practice using these variables. Finally, practitioners compare a most-endorsed model of practice with their actual practice to describe the final model. The produced T-PE model will be

consequently grounded in both theory and practice. This is an expanded and deepened logic modeling practice that mirrors how it might be done for programs in the field.

Purpose Statement

Transformative participatory evaluation has been defined, in theory, by researchers. But theorists have not empirically described it, and no more than a few practitioners have described it with anything other than case examples. Therefore, the ultimate objective of this research is to produce an empirical model of transformative participatory evaluation practice that contributes to the theoretical specificity (Miller, 2010) of participatory evaluation in a manner that is more accessible to evaluation practitioners and theorists alike. In short, this research will use input from theorists and practitioners to build a more complete theory of transformative participatory evaluation.

CHAPTER 2

REVIEW OF THE RELEVANT LITERATURE

This research seeks to empirically describe a theory of evaluation practice by using a theory-driven modeling process grounded in a mixed method approach. Thus, this literature review focuses on the development of participatory evaluation theory and how modeling and a mixed method approach can inform this development. To contextualize the summary of the literature that follows, a brief introduction to evaluation theory and an overview of the development of participatory evaluation and its relationship to transformative participatory evaluation are first provided.

An Introduction to Evaluation Theory

Researchers in general focus on creating and understanding basic knowledge as it *might* be applied to real world problems. Applied researchers (e.g., evaluators) take this a step further and develop *new* knowledge in the direct pursuit of solving those problems. Evaluation is therefore a practitioner-based discipline and evaluation theory is derived from practice (Shadish, 1998). For instance, Shadish, Cook, and Leviton (1991) describe the five important components of the evaluation of social programs through *post hoc* analyses of theorists' writings and practice case examples; utilization-focused evaluation (Patton, 1978, 1986, 1997, 2008) was developed from case studies of evaluations, including the groundbreaking 1977 study that surfaced the "personal factor" of evaluation use (Patton et al., 1977). These theories describe evaluation practice, but because few are empirically supported, they are prescriptive in nature (Alkin, 2004a). That is, they provide guidelines for practicing evaluation in some manner depending on certain contextual issues, such as the questions guiding the evaluation or the potential for using the evaluation findings.

Evaluation theories do not say “in 95 out of 100 evaluations if you provide this type of data at this point in the evaluation to this particular set of stakeholders you will get this type of outcome.” That requires extensive empirical study that takes into account all of the various issues and variables that interact within any given evaluation context so that prediction models might provide support for all potential outcomes of each choice. While this may be an ideal definition of descriptive or contingency theories, the discipline does, in fact, intend to move in that general direction so that it becomes a well-described and documented field. Shadish (1998) provides clear reasons why evaluation theory is important—it is what we talk about, it is what drives our conversations, it is the nomenclature that gives us a framework to guide practice, and it is what serves the researcher interests in many of us. Without evaluation theory, evaluation practice would be “too scattered, too ill-defined, and too vulnerable to poaching by the many other people who also claim that they can do evaluative work as well as we can.” (Shadish, 1998, p. 13).

A deeper understanding of a full range of variables, issues, and contexts can aid in judging the merit of professional versus novice approaches in evaluation practice (Shadish, 1998). Those who are aware of the potential applications of various theories based on contextual variables will be more experienced and competent evaluators, comfortable with various applications. More novice evaluators may not have the necessary experience. When the evaluation discipline develops its theories to the point that all (or most) contingencies in the field of practice can be controlled for, then teaching new evaluators will be an easier feat, and the practice of evaluation more standardized and replicable (Miller, 2010).

One strand of research on evaluation aims to better describe the discipline through empirical study of the practice. It has historically included various types of studies, including surveys of practitioners to understand their work, retrospective case examples as first hand evidence of evaluation practice, elaboration on these examples by theorists, and categorization efforts to make sense of various research efforts. For instance, Smith (1993) listed numerous empirical studies of evaluation practice; Shadish, Cook, and Leviton (1991) studied the writings of theorists through the lens of five “necessities” of good social program evaluation theory; Stufflebeam (2001) classified 22 evaluation models and subsequently rated nine of them using the Joint Committee’s Standards for Evaluation of Educational Programs, Projects, and Materials (Joint Committee, 1994); and Alkin and Christie (2004; Christie & Alkin, in press) used a tree metaphor to categorize the primary stance of evaluation theorists as either values-, methods-, or use-focused.

As these views integrate and consolidate broad evaluation practice and theory, they refine our understanding of the discipline, create boundaries and frameworks, and, subsequently, invite us to go back into the field to observe practice and to theorize about the boundaries and connections that support or refute these categorizations.

Of the many ways that theories are developed, Alkin (1991) describes the value of theorists comparing their own theories with others’ interpretations (such as the practitioner who cites a theorist’s work in a way that conflicts with how the theorist sees her own work), exploring how others’ theories relate to their own, and being categorized in a schema like the “evaluation theory tree” (see Fetterman, 2004). He also points to the importance of one’s own research or field experiences and/or personal interactions with others about the theories. This theory building

process provides us with relatively prescriptive theories that might be more appropriately called evaluation models or approaches (Alkin, 2004a).

A model is a mental representation of a phenomenon of interest that uses some elements or a framework familiar to users. For instance, scientists' planetary model of the atom, with electrons orbiting a central nucleus, uses spheres in three-dimensional space that reflects our understanding of the earth and sun's movements. While this may not be the exact shape or relative dimensions of an atom, the model uses familiar ideas to provide some understanding of a much more complex theory. In the same way, evaluation theories simplify the very complex nature of a program, policy, or product evaluation, and place it into a recognizable framework.

Development of Participatory Evaluation Theory

Early evaluation practice (i.e., pre-1980s) was grounded in a positivist search for effective solutions to social problems (Greene, 1987; Shadish et al., 1991). From this perspective, stringent application of research methodology (e.g., Campbell's "Experimenting Society") (Campbell, 1991) was used to produce evidence of a program's success. Successful programs could then be replicated and transferred to other problems or contexts and those not proven successful would be terminated (Cronbach & Associates, 1980). Frequently, this expectation of program termination went unfulfilled as programs continued despite negative evaluation findings (Patton, 1997; Shadish et al., 1991). Further, these evaluation experiments often proved difficult to sustain and rarely provided valid data (Alkin & House, 1992; Fitzpatrick, Sanders, & Worthen, 2003; Cronbach & Associates, 1980; Patton, 1997; Shadish et al., 1991; Weiss, 1972).

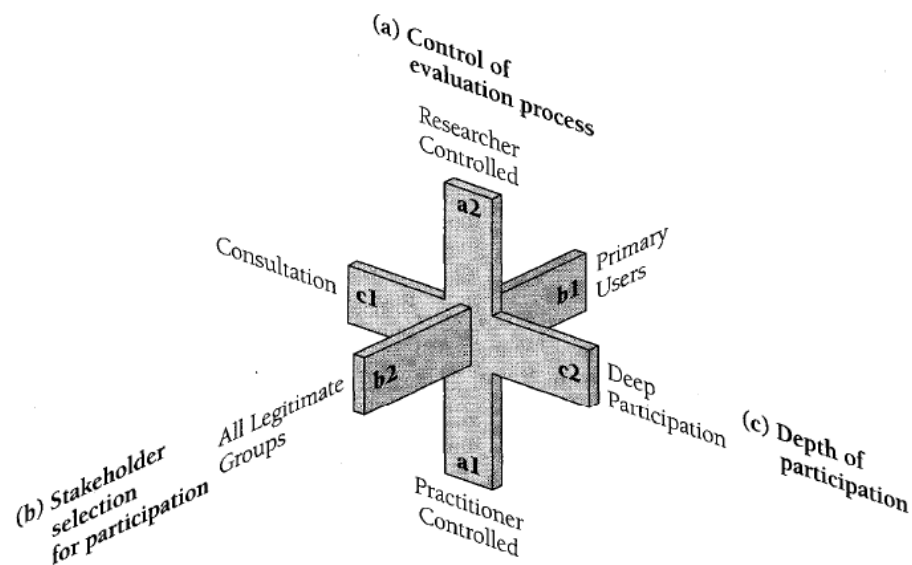
Because evaluation findings were used so infrequently, an era of research on evaluation's value ensued (Patton, 1997; Weiss, 1972). Researchers focused on the variables that seemed to mediate its use. Stakeholder involvement emerged as a predictor variable (Alkin, 1985; Alkin & Patton, 1987; Alkin, Dailik, & White, 1979; Alkin, Kosecoff, Fitz-Gibbon, & Seligman, 1974; Cousins & Leithwood, 1986; Greene, 1987, 1988a, 1988b; Patton, 1997; Patton et al., 1977). Later research on evaluation use also supported the notion of participation of program practitioners, program participants (beneficiaries), and program evaluators as key factors in utilization (Johnson, 1998; Turnbull, 1999).

By the mid-1980s, evaluators began to realize that public funders of evaluations probably did not hold the broadest value perspectives and should not be the sole arbiters of program value (Shadish et al., 1991). Also, Cronbach and Associates (1980), after writing extensively on experimental designs using individual outcomes as indicators of program success, argued for recognition of contextual confounds that may be best tapped into through engagement with stakeholders. To connect with broader values, the field expanded participation to include engaging stakeholders closest to the program (Mark & Shotland, 1985). At the same time, researchers like House (1980), Bryk (1983), Stake (1975), and Cronbach and Associates (1980) were advocating for more responsive, inclusive evaluation designs that involved those with the most informed stake in the evaluand. For example, Bryk, Gold, and Murray, as representatives of the National Institute for Education, used a stakeholder-engaged model of evaluation with two highly visible national programs—Cities-in-Schools and Jesse Jackson's PUSH-Excel program (Bryk, 1983; Gold, 1983; House, 1991; Murray, 1983; Weiss, 1983a, 1983b). Though these evaluations were considered unsuccessful attempts at cooperation and engagement (Murray, 1983), they represented a positive move toward active stakeholder involvement (Weiss, 1983b).

From the confluence of this disparate theorizing, a use-focused participatory model of evaluation emerged that involved key stakeholders as participants. A 1995 exploratory survey of North American evaluators reflects this stakeholder-based evaluation approach (Cousins et al., 1995, 1996). Survey respondents reported evaluation participants were involved at the early and later stages of the evaluation—designing research and evaluation questions and interpreting and disseminating results. Generally, participants were limited to those with a vital stake in the program and did not include program beneficiaries (Sabo Flores, 2008). While evaluators reported maintaining most of the evaluation's technical control, some described stakeholders engaged in instrument design and data analysis.

As illustrated in Figure 1, Cousins et al. (1996) posited three dimensions of collaborative research and evaluation: a) the evaluator's level of control over methodological rigor (full evaluator control to stakeholder-held); b) the selection of stakeholders (from key decision makers to all legitimate groups); and c) the depth of involvement of these stakeholders (from data sources to full partners in the research).

Figure 1. Dimensions Of Form In Collaborative Inquiry (Cousins & Chouinard, 2012)



Control of Evaluation

Control can rest fully in the hands of evaluators, entirely in the hands of participating stakeholders, or somewhere in between (Cousins & Earl, 1992). It may involve evaluation decision-making as well as control over the evaluation's resources and timeliness (T. Azzam, personal communication, 2010). Greater technical control in the hands of participants produces enhanced learning about the program and the application of evaluative thinking to other parts of their lives (Coupal & Simoneau, 1998). Unlike evaluators who focus on empowerment, participatory evaluators generally balance this dimension to allow adequate participation without sacrificing methodological rigor.

In separating a T-PE practice from other participatory practices, one would expect to see more control shared with practitioners in order to enhance ownership, empowerment, and potential use (Cousins et al., 1995, 1996; Cousins & Whitmore, 1998; Fetterman, 1994). Furthermore, technical control is not stagnant. It often begins in the hands of the evaluator and is

divested to stakeholders as capacity builds (Naylor, Wharf-Higgins, Blair, Green, & O'Connor, 2002).

Organizational power structures (or context) may impact how control is managed within a participatory evaluation. Weaver and Cousins (2004) place power differentials and manageability as separate dimensions of collaborative evaluations. But, Daigneault and Jacobs (2009) label these as facets of control. From this perspective, control is a broader dimension that encompasses both who has control of the evaluation and the manageability of that control. The current research prefers this more comprehensive definition of control.

Selection of Participants

The selection dimension of collaborative inquiry can be viewed as the number and diversity of participant stakeholders or stakeholder groups. There is a general acceptance that diversity of participation is preferable (Daigneault & Jacob, 2009; Weaver & Cousins, 2004). In their effort to better conceptualize measurement of participatory evaluation, Daigneault & Jacob (2009) simplified this dimension from selection to diversity. Specifically, measuring selection by how many different stakeholder groups are engaged (diversity) indicates how participatory the process is and probably informs how manageable the process will be. While Weaver and Cousins (2004) split this dimension into manageability, power differentials, and diversity, Daigneault and Jacob (and the present research) see limited value in this differentiation.

With respect to practical and transformative participatory evaluation, the key selection difference relates to the scope of participants. P-PE usually engages key decision makers because their positional authority has the potential to make the greatest use of evaluation findings (Cousins & Earl, 1992; Cousins, et al., 1995, 1996). T-PE opens participation to all interested

stakeholders and focuses particularly on involving program beneficiaries (Brisolara, 1998; Cousins & Whitmore, 1998). The transformative approach, described by Mertens (2009), would purposefully involve those with historically less voice for the purposes of shifting power structures. For Daigneault and Jacob (2009), participation by any non-evaluative stakeholder constitutes a point of diversity and indicates a participatory evaluation.

Depth of Participant Involvement

Preskill and Caracelli (1997) surveyed AEA's Use TIG and reported 80 percent of their survey respondents agreed that evaluators should take responsibility for involving stakeholders in evaluation processes. In a reiteration of Preskill and Caracelli's survey, Fleischer & Christie (2009) found that 98 percent of their respondents agreed with this assertion. Involvement of stakeholders in an evaluation is a methodological choice that has gained significant ground.

Daigneault & Jacob (2009) propose measuring participation using a dichotomous variable (yes/no) for each of four stages of an evaluation. On a scale of 0 to 1, involvement at a single stage by one member of any non-evaluative stakeholder group constitutes 25% involvement. In their effort to narrow the conceptualization of this complex dimension, Daigneault and Jacob detach from the nuances of participatory research and this scale was not used in the current study.

Transformative Participatory Evaluation

As noted earlier, Cousins and Whitmore (1998) parsed participatory evaluation into two broad categories—transformative participatory evaluation (T-PE) and practical participatory evaluation (P-PE). P-PE is seen mostly as a North American practice (Brisolara, 1998; Cousins & Whitmore, 1998) focused on stakeholder involvement to foster greater relevance, ownership,

and use (Cousins & Earl, 1992; Greene, 1988a, 1988b; Patton, 1997). T-PE uses many of the same processes as P-PE, but intends to produce social change by empowering the disempowered (Brisolara, 1998; Burke, 1998; Cousins & Whitmore, 1998). It is more aligned with participatory action research's (PAR) focus on power redistribution (Brisolara, 1998; Estrella & Gaventa, n.d.; Greenwood, Whyte, & Harkavy, 1993; Sabo, 1999; Suarez-Balcazar & Harper, 2003) but differs from empowerment evaluation (Fetterman, 2005) in its evaluator role. While both T-PE and empowerment evaluation focus on empowering the disempowered, a T-PE evaluator maintains more technical control and is more engaged in managing and directing the evaluation than an EE evaluator.

T-PE empowers participants through varied data collection strategies that encourage joint knowledge creation (Burke, 1998; Meyer, Park, Grenot-Scheyer, Schwartz, & Harry, 1998; Whitmore, 1988). This is rooted in a southern hemisphere developmental perspective, where non-governmental organizations (NGOs) work in evolving economies (Brisolara, 1998; Goulet, 1989). NGOs see evaluation as an opportunity to further engage their constituencies in an educational experience, while evaluators see a need to give the local communities a voice in the process. T-PE addresses both (Brisolara, 1998; Goulet, 1989; Lau & LeMahieu, 1997; Monkman, Miles, & Easton, 2007). Likewise, in the youth development field, evaluation is another opportunity to offer youth more ways to take ownership of their lives and development (Sabo, 1999; Sabo Flores, 2008; Checkoway & Gutierrez, 2006; Checkoway & Richards-Schuster, 2003)

The role of human agency in knowledge creation also helps define T-PE. Evaluation participants produce socially constructed knowledge through dialogue. As knowledge informs

the evaluation, its creators are empowered by seeing their knowledge being used (e.g., Brisolara, 1998; Cousins & Whitmore, 1998; Sabo, 1999). As a result, participatory evaluators and evaluand stakeholders develop close relationships, mutual respect, and deep understandings (Gaventa, 1993; King, 1998). Also, participants gain a greater sense of control and agency when they see their knowledge put to use in a respectful, team-oriented manner.

While it is important to parse empowerment evaluation (EE) from T-PE, it is not an obvious distinction. One problem is that EE seems to hold multiple intentions and may still be an unfocused theory (Cousins, 2005; Miller & Campbell, 2006; Shadish, 1998). Both EE and T-PE intend to empower and transform participants, but T-PE encourages sharing of decisions among stakeholders and the evaluator, while EE decision-making is often abdicated to the participants so that they feel the impact of their decisions, and are subsequently transformed by the experience (Fetterman, 2005; Cousins & Whitmore, 1998). Key differences can also be seen in the role of the evaluator. An EE evaluator is more of a program advocate with allegiance to the success of the program (Fetterman, 2005). On the other hand, a T-PE evaluator advocates for the voice of the people undertaking the evaluation. EE is also more constructivist, in that all the power is left in the hands of the participants to create their own evaluations within their own realities (D. M. Fetterman, personal communication, 2008). In T-PE, the evaluator manages decision-making and technical control and divests both as the capacity to appropriately apply them evolves.

Transformative Evaluation

Transformative evaluation (TE) has a developing theoretical base through the efforts of Donna Mertens (e.g., Mertens, 2009; Mertens & Wilson, 2012). Mertens draws TE from multiple

theories, including feminist, queer, critical race, postcolonial, and indigenous theories. Its focus reflects a respect for contextually-bound cultural norms and places high value on increasing social justice and promoting human rights. Embedded in a critical theory framework, TE is aware of the role of privilege, its power in defining what is real, and the consequences of making decisions from a position of power. The multiple realities of a transformative approach are therefore defined by the various identities of the culture within which the evaluation is performed (e.g., gender, ethnic, economic, cultural, etc.).

As Mertens and Hopson (2006) note, a transformative paradigm offers a “theoretical umbrella” that allows evaluators and stakeholders “to explore the philosophical assumptions and guide methodological choices for approaches to evaluation that have been labeled inclusive, human-rights-based, democratic, constructivist, and responsive.” Moreover:

The transformative paradigm extends the thinking of democracy and responsiveness by consciously including identification of important dimensions of diversity in evaluation work and their accompanying relation to discrimination and oppression in the world (p. 48).

The evaluator’s role in TE is “consistent with the evaluator as advocator of democracy and democratic pluralism” (Mertens & Hopson, 2006, p. 42). Knowledge is attached to and situated in the process by which it is created, including the perspectives of the evaluator as well as all the other stakeholders engaged in the evaluation. Socially constructed knowledge, such as local program experience, is as valid as empirically produced knowledge. For instance, a program manager has a perspective on the inner workings of the program that, while perhaps different from others further from the process, is grounded in the program’s active existence.

Because of its power, knowledge is not neutral; knowledge reflects the interests of humans and their desire for power. Those historically without power find less equality and justice because their voice is rarely heard. Those with power are generally interested in maintaining the status quo. The attributes of utility, feasibility, propriety, accuracy, and evaluation accountability from the *Program Evaluation Standards (3rd ed.)* (Yarbrough, Shulha, Hopson, & Caruthers, 2011) provides for the attention to stakeholders, a negotiated purpose, and a responsive and inclusive orientation. These standards are intended to mitigate a lack of attention to contextually responsive questions and evaluation focus so that the voices of the historically subverted are heard.

Methods within a transformative approach are selected to provide close contact with the context. While quantitative and mixed methods are acceptable, qualitative, dialogic methods ensure grounded perspectives. Participants are involved interactively in developing the research's purpose and focus. "Methods are adjusted to accommodate cultural complexity, especially as they relate to discrimination and oppression" (Mertens, 2009, p. 49)

Transformative evaluation's relationship to T-PE has only recently been articulated in the literature (Mertens & Wilson, 2012). The primary difference between the two may be connoted in their names: Transformative evaluation is rooted in a social justice perspective aimed at balancing power structures and transforming social conditions. Here, stakeholder involvement is utilitarian in its transformative impact on the program and the power of situating design within the community being studied. Transformative participatory evaluation might then be an application of the transformative philosophy of transformative evaluation (Mertens & Wilson, 2012).

Theory Driven Evaluation and Logic Models

Theory driven evaluation puts emphasis on creating a clear, mutually agreed upon description of the evaluand, including its implicit and explicit theories, and using that knowledge to design an evaluation that tests assumptions (e.g., Chen, 1990, 2005; Chen & Rossi, 1992; Donaldson, 2007; Rossi et al., 2004; Sheirer, 1987). Weiss (1997) situates the roots of theory driven evaluation in the 1960s, and Wholey (1979, 1987) later popularized it as part of an evaluability assessment and program description studies (see also Frechtling, 2007). Some authors have adapted the name, adding “program” at the beginning to emphasize the importance of using a program’s theory (and not only social science theory) (Chen, 2005), or science to the end to emphasize the importance of using scientific methods in the evaluation (Donaldson, 2007). Regardless, the basis of theory driven evaluation is the same across interpretations (Coryn, et al., 2011) and the names are used interchangeably in the present study.

When practicing theory-driven evaluation, the evaluator works with key stakeholders to describe the program’s theories, formulate and prioritize evaluation questions, and collect data to answer questions (Donaldson, 2007). Understanding a program thoroughly at the earliest phases helps keep evaluators from rushing into an evaluation using the design at the top of their toolbox. As well, enlisting stakeholders in the process of defining their program, as in the Centers for Disease Control’s Framework (Milstein & Wetterhall, 1999), gives an opportunity to nurture ownership of the evaluation process and sometimes affect a new, improved, shared understanding of a program.

An important first step in a program theory-driven evaluation is to make explicit the target program’s theory. Different evaluation theorists express this step in different ways. For instance,

Chen (2004, 2005) talks about identifying both prescriptive and descriptive theories. Chen's prescriptive program theories describe the actions occurring in a program that hopefully lead to alleviation of the ill for which the program was designed. His descriptive theories describe the change processes that occur as a result of the program activities and help identify the outcomes associated with the program.

Rossi and colleagues (2004) and Donaldson (2007) break the program theory into two general pieces—process-oriented theories and impact theories. Process-oriented theories include organizational and program theories that explain how the intervention should function. Impact theories explain the proximal and distal outcomes that should occur as a result of the program's process. These impact theories are critical to the efficacy of any program and need to be considered in depth when describing a program theory (Donaldson, 2007). Program theories often are more conceptual than actual and need to be corroborated with the true workings of the program before being finalized (Rossi, et al., 2004).

After stakeholders are engaged, and a draft of a model is presented and agreed upon, evaluators do a plausibility check to see if the theories believed to be at work are substantiated in previous research. Each element in the model is probed to increase the model's specificity (Donaldson, 2007) before a final model is created. The evaluation team then identifies parts of the program theory that would be critical to outcomes and determines where the focus of the evaluation should be placed. For instance, if a key change theory in a program's model says that, by engaging in a community of learning, underprepared community college students will transfer to four-year colleges faster than a similar group of students not involved in a learning community, then an evaluation question might be "How many more learning community

students transfer to four-year colleges than mainstream underprepared students?” This question would probably initiate a conversation around valuing outcomes. For example, is transfer a valuable enough outcome for the program, or is it an output leading to a larger outcome of upward social economic status mobility? Without the explication of this change theory, evaluators and program managers might struggle for an unnecessarily extended amount of time trying to identify the learning community’s centrality to the program’s change theory.

The most common way of explicating program theories, through the use of logic modeling processes, has gained its popularity through the program theory work done by the scholars mentioned above (e.g., Chen, Donaldson, Rossi, Wholey) (Frechtling, 2007). A logic model, as defined by this research, is a collection of elements on a graph that make explicit the activity and causal theories of a program. Many start with assumptions on the left and move left to right, providing activities, outputs, outcomes and impact. Some logic models have context and values running across the bottom and some are built with feedback loops to represent the non-linearity of some program theories. They have been called log frames in other fields (Hummelbrunner, 2010), but evaluation practitioners mostly know of them as logic models. Developing logic models of program theories is an established procedure in the evaluation toolkit.

The logic model is a useful tool not only for managing and evaluating programs, but also for promoting and facilitating organizational learning and development. First, the logic model makes clear and explicit how a program is supposed to work and what activities and outputs should lead to what outcomes. Because both the elements of a program theory and the connections between the elements are made explicit, program staff knows not only what they should do, but also why

it is important to do so. Organizational plans cease to be bureaucratic to-do lists and instead allow program staff to feel like they are part of the team, because they are.

For the evaluator, developing and being guided by a program theory takes some of the element of surprise out of the evaluation. Rather than evaluating a program with assumptions about what the program is intending to do, a logic model provides the groundwork to be more informed about the program and its activities. A well-constructed logic model provides deeper understandings of a program's workings and avoids superficial assumptions about how activities link to outcomes. This informs program staff about what the evaluation will be looking for and provides evaluators with a roadmap for assessment.

From an organizational perspective, making explicit connections between program objectives, program activities, and resources simplifies program process monitoring. This can become part of the program's operational documents and guide ongoing decision-making. In the absence of such a model, many organizations lack clear focus on how goals will be achieved. Indeed, the W.K. Kellogg's Logic Model Development Guide (2004) notes, "According to many funders, grant applications frequently lack solid descriptions of how programs will demonstrate their effectiveness.... Conducting an activity is not the same as achieving results from the accomplishment of that activity.... Specifying program milestones as you design the program builds in ways to gather the data required and allows you to periodically assess the program's progress toward the goals you identify" (p. 16).

Explicating a theory through logic modeling also provides the groundwork for a second phase: testing that theory in action. The first step in this phase is to conduct a literature review to identify if and how this type of theory has been applied in other places and times. If it has been

applied elsewhere, the available information might help frame an understanding of how the program should work and how others have viewed the relationships between intervening variables. It grounds the testing of relationships between the intervening variables in the program to see where each affects program outcomes.

For the present research, modeling program theories in program theory-driven evaluation science is equated with the modeling of implicit theories in one's evaluation practice. Attempts at this are seen in the work of a group of University of California Los Angeles researchers. Wallace and Hansen, (2010), Vo (2010), Dillman (2010), Luskin (2010) and Ho (2010) have all drawn comparisons between three evaluation theories using logic model-like representations of the theories. They each compared their models of emergent realist evaluation, value-engaged evaluation, and practical participatory evaluation using Mark's (2008) framework for research on evaluation. They chose these theories from the three branches of Alkin & Christie's (2004) theory tree (i.e., use, methods, value) to ensure some diversity in the produced models. The prevalence of recent research on these three bodies of theory provided a strong foundation to study and a relatively short and accessible list of key authors. It is telling that they undertook depicting P-PE, but avoided T-PE. The dearth of T-PE research no doubt informed its omission.

Creating the models within existing frameworks (i.e., Mark, 2008; Shadish et al., 1991; "Logic Model Training Module", n.d.) provides empirical credibility and certainly contributes to research on evaluation. But, this process is grounded only in the published literature. It may or may not reflect practice. Until their models are tested in practice, and found to adequately reflect *both* theory and practice, they are still prescriptive. As such, the current research takes the next step by engaging practitioners in modeling their own practice to produce a model grounded in

both theory and practice. The resultant model is a more complete reflection of T-PE theory in practice. It might then be compared to other theories to further develop those theories and continue the cyclical process of theory development.

Models of evaluation theory provide an opportunity for interactive engagement with theory in a two-dimensional space. They place activities on a timeline, make assumptions explicit, and provide a space for outcome expectations. By making these explicit and showing the links between them, we can provide those attempting to understand theory a visual representation of what are otherwise abstract ideas.

The systems theory concepts of perspectives, boundaries, and relationships provide another perspective in the discussion of evaluation theory modeling. Specifically, *perspectives* are represented by a theory's values and assumptions, which are critical to the application of an evaluation theory to practice and inform the selection of participant stakeholders, the depth of their involvement, and the control divestment. *Boundaries* are valuable in describing the elements of evaluation practice. For example, when operationalizing the activities and outcomes of practice, it is important to reflect back on the values and assumptions undergirding the theory. And *relationships* in modeling evaluation practice are exemplified in the causal attributions attached to resources, activities, and outcomes. All of these are informed by the barriers placed on the activities, the values attached to their outcomes, and the looping nature of theories (Williams, 2010). In the process of developing evaluation theory models from practice, ample opportunities arise to discuss key mediator and moderator variables in a causal chain. By making them explicit and discussing the relationships and operationalization of the variables in a model, the evaluation theory and practice is made clearer.

Making explicit the interrelationships in an evaluation theory logic model guides the testing of that theory. For instance, if stakeholder engagement is expected to create a shared understanding, one can test this relationship and begin to describe the mediating and moderating variables that affect that relationship. Finally, transferring the understanding of modeling program theories to that of modeling evaluation theories creates a deeper understanding of the modeling process. The understanding can then be transferred back to the program theory modeling process, improving the evaluator's future use of the practice with a client's program.

Summary

Evaluation theory development, compared to other, more established social science disciplines, is in an adolescent stage. Our prescriptive theories of practice are ripe for analysis. The discipline's categorization schemas provide broad understandings that need to be supported or refuted by empirical study of the different theories in practice. The resulting descriptive theories will pave the way for creating a more respected, replicable, and rewarding discipline that can more easily be taught and practiced.

Participatory evaluation is a well-researched evaluation domain. In its evolution, some of the founding theorists have posited a bifurcation into a more practical, use-focused application and a more social justice and individual empowerment-focused application. These two may be valuable heuristics, useful in discussing different philosophical orientations but the schema has weak empirical support.

A common practice in theory-driven evaluation approaches is logic modeling a program's theory at the outset of the evaluation. This modeling provides a baseline understanding of a program's implicit and explicit theories of action and change, and provides evaluators insight

into how to design evaluation questions. By equating evaluation theories with program theories, modeling evaluation practice can provide similar data. Practitioners can model their own evaluation practice and the commonalities in their models might provide a greater understanding of what their practice looks like. This, in turn, can set the stage for further questions related to practice and outcomes.

The remainder of this dissertation describes the three phases of the research: (1) the development of variables and statements that help define T-PE using feedback from three experts on the topic; (2) individual model development by T-PE practitioners; and (3) the series of webinars designed to gather feedback on the most-endorsed model derived from the findings in the second phase. The methodological approach to this study, where findings from each phase informed the development of subsequent stages, does not lend itself to a traditional structure, where findings follow an overall summary of the methodologies employed. Instead, the structure of the document is somewhat non-traditional; each phase—including both the methods used and the findings—is described in its own chapter. Chapter 3 first provides an overview of the entire study, and then describes in detail the first phase of the research.

CHAPTER 3

OVERVIEW OF METHODOLOGY AND PROCEDURES & PHASE ONE (EXPERT PANELISTS)

This study sought to produce empirical knowledge about participatory evaluation by focusing on transformative participatory evaluation (T-PE) as described by Cousins and Whitmore (1998). The overall goal was to further develop participatory evaluation as theorized and as practiced. The study was guided by a desire to determine the key variables in both a theory- and practice-based T-PE and to use these variables to develop a model of T-PE practice that could then be compared to other evaluation theories.

To address these questions, the study had three stages, with each building upon the results of the previous stages. This and the two chapters that follow will cover separately the methods used and results found during each stage. First, however, it is helpful to summarize the entire study in order to provide context for the discussion that follows. Following the summary, the balance of this chapter provides a description of the findings from phase one of the research.

Overview

In an effort to further understand transformative participatory evaluation (T-PE) practice, an expert panel was engaged to create a list of its key variables. A broader sample of T-PE practitioners was then asked to use these variables to model their practice. A most-endorsed model was produced from their models and a subgroup of these practitioners was invited to participate in a series of webinars to describe how that model reflected their practice. This descriptive, mixed method, multi-stage design was used to produce empirical evidence of T-PE practice that informs the development of participatory evaluation theory (Figure 2).

Overall Study Methodology

Mixed Method Theories, Mental Models, & Paradigms

This research is descriptive in that it does not purport to answer any causal questions. Rather, its purpose is to further describe the nature of transformative participatory evaluation as described by Cousins and Whitmore (1998). Therefore, it takes an a-paradigmatic approach to knowledge gains and method choices. Greene (2007) describes a paradigm as a worldview structure defined by the nature and scope of what we know of the world, what we trust as valid knowledge of the world, and how we come to have this knowledge (Greene, 2007; Greene & Caracelli, 1997; Guba, 1998). A-paradigmatic approaches are somewhat pragmatic. Knowledge claims and method choices are not dogmatically ascribed to any one paradigm. Rather, the research process defines reality, truth, and learning. Taking this approach, the strengths and weaknesses of the different methods are capitalized upon to provide the most well supported data for the argument (Tashakkori & Teddlie, 1998).

Within this framework, the current research is predicated on a post-positivist ontology. While multiple realities surely make up our understanding of the world, we can still achieve an understanding of reality that many people would agree upon. Because evaluation practice is decidedly relativist, working to describe something so complex is an ontological struggle. It is constantly adaptive to the contexts at hand and any one evaluator may work within a given theory and yet her practice might look quite different from one evaluation to the next. From a pragmatic sense then, this research makes an effort to describe practice by using methods that offer access to the largest number of practitioners with the hope that the description will have merit to other, similarly focused evaluators.

Further, this research is epistemologically relativist. It balances its truth measures on both subjective and objective beliefs. Tashakkori and Teddlie (1998) say that embracing both subjective and objective points of view is “inevitable” (p. 26) if a researcher uses both qualitative and quantitative tools. In this study, objective measures of central tendency informed choices in the quantitative stage and subjective experiences in the webinars informed beliefs in that stage’s findings.

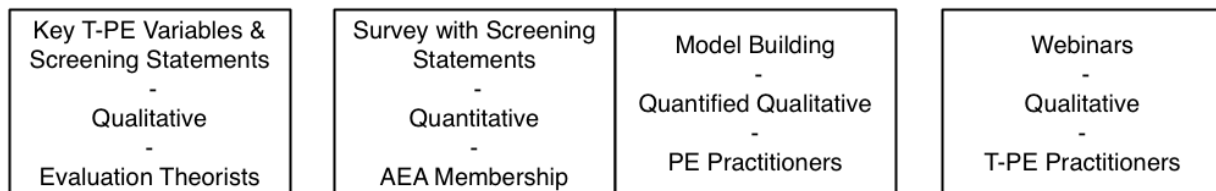
While efforts were made to reduce the effects of the researcher’s values on the research, these effects are inevitable. From an axiological post-positivist perspective, effort was made to control the influence of the researcher’s values on the methods and interpretations by aligning with established research practices and norms. Further, the internal and external validity of the findings were examined so that they are less value-bound. While this research was undertaken to add to the existing knowledge base, the very nature of research means that it is a product of the researcher’s values and hence attached to those values.

This is a sequential mixed method study. Methods are mixed from phase to phase in pursuit of development and of complementarity (Greene, 2007). Developmentally, each phase built upon the knowledge gained in the previous phases. For instance, qualitative T-PE variables were used to create quantitative items on the survey, and qualitative modeling data was quantified to create a most-endorsed model, which was discussed in webinars to produce qualitative insights.

Two different methods were used to provide complementary understanding of T-PE practice (Figure 2). Participants were asked to use these variables to model their actual practice and then later were asked to describe how the model reflected their practice. These two methods provided complementary understandings of evaluation practice, a quantitative construction of practice that

applies values to the endorsed variables, and qualitative interpretations of practice that deepen our understanding of practice.

Figure 2. Mixed Method Design Elements And Participant Samples



Mixed methods were also used to augment each other. For instance, a shortcoming of the reductionist paradigm used to produce the practice model is that it lacks descriptive depth (Greene, 2007; Patton, 2002). This is therefore complemented in a later stage where evaluators were offered an opportunity to describe how the common model reflects their practice. While a quantitative method was used to produce the model, qualitative dialogue provided a fuller understanding of that model. This interaction of methods strengthens the findings because quantitative data often answers a very narrow-scoped question and qualitative data brings with it the nuances of individual personal experience that expand on the answer provided by the quantitative data.

The assertions generated by this study (namely that the resulting model has merit) were judged by the validity of the inferences used to make those assertions. The validity of these inferences rests on the assumption that each method used to collect and interpret the data was well executed. For instance, inferences from qualitative analysis of the webinars are only as valid as the credibility of the analysis (Miles & Huberman, 1994; Tashakkori & Teddlie, 1998). Inferences were further judged from a multiplistic stance that focused on how consistent the

findings were across the study (conceptual consistency), whether or not they would stand up under scrutiny from objective, knowledgeable outsiders (interpretive agreement), and their independence of existing theories and practice (interpretive distinctiveness) (Greene, 2007; Teddlie & Tashakkori, 2003).

We turn now to the first phase of the research, which drew on the expertise of three established evaluation scholars who have published on the topic of participatory and transformative evaluation. This first step in data collection and analysis was designed to develop an informed framework for the deeper exploration of how this approach to evaluation could be modeled.

Phase One: Expert Panelists

The first phase of this study focused on generating two sets of data necessary for developing a further understanding of T-PE. Specifically, a set of key variables in T-PE practice was identified so that evaluators in the next phase could use them to model their practice. To ensure that the created models were attributable to T-PE evaluators, a set of descriptive statements was also developed so that T-PE evaluators could be filtered from other practitioners. The methods employed to develop these two sets of data and the results from this phase of the study comprise the remainder of this chapter.

Phase One Methodology

Participants

Three prominent evaluation theorists were invited as a purposeful panel to develop the T-PE variables and identifying questions: Drs. Elizabeth Whitmore, J. Bradley Cousins, and Donna

Mertens. These three were selected because they have well-established publication records focused on participatory and transformative evaluation.

Elizabeth Whitmore is Professor Emerita in the School of Social Work at Carleton University in Ottawa, Ontario, Canada. She has extensive experience writing and thinking about T-PE and participatory forms of research and evaluation (e.g., Whitmore, 1994, 1998a, 1998b; Whitmore & McKee, 2001; Whitmore et al., 2006; Whitmore, Wilson, & Calhoun, 2011). Her 1988 dissertation explored the empowerment effects of T-PE and she co-authored, with Cousins (1998), the first article to use the term transformative participatory evaluation.

J. Bradley Cousins is Professor of Program Evaluation and Organizational Studies at the Faculty of Education, University of Ottawa, Ontario, Canada. His research on participatory and collaborative evaluation (e.g., Cousins & Chouinard, 2012; Cousins & Earl, 1999; Cousins & Whitmore, 1998; Cousins et al., 1995; Cousins, Whitmore, & Shulha, in press) makes him the second most knowledgeable theorist on T-PE. Though Cousins' research is grounded in practical perspectives stemming from his interest in evaluation use, his research has also identified key variables in participatory evaluation that make him a respected theorist with a studied perspective on T-PE.

Donna Mertens is a professor in the Department of Educational Foundations and Research at Gallaudet University, Washington, DC. She writes extensively about inclusive evaluation (1999, 2005), transformative evaluation (e.g., Mertens, 2009; Mertens & Hopson, 2006) and the transformative paradigm (2007). Mertens' transformative perspective is rooted in issues of diversity, privilege, and power, as well as in evaluation's role in advancing issues of social justice, oppression, discrimination, and power difference. She borrows from literatures as diverse

as queer, feminist and critical theories. Though Mertens recommends participation as a major component of evaluation, and notes that this participation informs transformation, she views participation as encapsulated within transformation. In her view, programmatic and societal transformation are higher order issues that gain more attention than participation (Mertens, 1999). It was anticipated that her strong focus on social justice and democratic pluralism would provide the appropriate transformative balance to Cousins' use-focused perspective.

Combined, these theorists have more than 60 years of evaluation practice and research experience. Their perspectives, while somewhat different, are highly focused on participatory and transformative evaluation. In fact, their work is cited throughout the current study.

Procedures

Each theorist was contacted individually by email and asked to participate in the study. The researcher then met with each at the 2010 American Evaluation Association conference in San Antonio, Texas, where the study was explained in depth and any questions were answered.

They were asked to help develop a set of statements that would identify T-PE evaluators from other evaluators and to help develop a set of key variables of T-PE practice. The researcher's role would be to provide them a preliminary set of statements and variables for their consideration and to facilitate and moderate their work by managing the online environment and contributing responses to their questions and recommendations. The three were also encouraged to provide any missing variables or statements if the researcher's preliminary offerings were incomplete. They all agreed to participate and the remainder of the work was completed online over the next two months, using a wiki and email.

A wiki is an editable web page that allows multiple participants to collaborate on any number of documents. It was not clear at the outset if the research team would want the ability to post supporting documents, and a wiki offered adequate options for document management. Wikis allow each participant to set his or her preferences for alerts and the researcher set his preferences to be informed by email immediately upon an edit so that he could respond promptly.

The initial sets of variables and statements were posted to the wiki on November 20, 2010, and the theorists were invited to comment and edit. (See Appendix A for the wiki front page.) The wiki front page included editing ground rules to allow for a complete record of the development process. Specifically, the participants were asked not to delete anything, but rather to only ~~cross-out~~ words or phrases, adding comments after the edit, followed by the author's initials. The conversation took place over approximately two months. The process had what might be called a meandering quality, where a theorist would log on to the wiki and contribute when he or she had time and the researcher would make comments or edits as necessary soon thereafter. The researcher also sent two reminder emails to encourage participation. Each theorist provided at least two rounds of comments (described below) before agreeing that our product was a "good place to start."

Development of Variables

In keeping with the logic modeling methodology of this research, where assumptions, resources, activities, and outcomes often describe a program, the researcher first developed a preliminary list of key variables and their definitions that addressed the expected principles, activities, and outcomes of T-PE (Appendix B). This list was generated from a close examination

of key writings in transformative and participatory evaluation (Burke, 1998; Cousins et al., 1992, 1996; Cousins & Earl, 1992, 1995; Cousins & Whitmore, 1998; Mertens, 1999, 2009; Patton, 2008; Whitmore, 1998). These were identified through a previous literature review by the author, snowball sampling from the reference lists of those identified in the literature review, key word searches through three databases (Web of Science, PsychInfo, and ERIC) using “transformative participatory evaluation” and a Web of Science “cited reference search” on Cousins and Whitmore’s key article (1998).

An inclusive approach was taken to developing the preliminary list—many variables were included with the expectation that the work with the three theorists would reduce it to a core list. Variables were included if they were identified in the literature as important in distinguishing participatory evaluation from non-participatory and transformative participatory from practical participatory. For instance, Burke’s (1998) seven key principles of participatory evaluation (pp. 44-45) were included in the preliminary list.

The original intention was to develop a broad set of variables representing the qualities that are most important to both practical participatory and transformative participatory evaluation. This list could then be presented to the modelers, who would identify the key variables for their preferred practice. This would ostensibly provide more variance between models by P-PE and T-PE evaluators. After the initial list of variables was defined, however, it became evident that the necessary list would be too long, making the modeling process too cognitively challenging to produce reliable data. At that point, the research design was narrowed to identify only the key T-PE variables. Hence, the original list was designed with an inclusive nature that produced many variables more related to PE or P-PE rather than just T-PE. Most were subsequently deleted

during the editing process with the theorists. (See Appendix B for the preliminary list and Appendix C for the list that followed the theorists' editing process.)

Principles here are defined as worldviews or underlying assumptions that guide one's practice (e.g., multiple method perspective). Some principles are more concrete and may have actions attached to them that imply a principle. For example, "engage intended beneficiaries" is an action that includes the assumption that one values engaging intended beneficiaries. It may be seen as a directive, in that an evaluator is philosophically compelled to engage intended beneficiaries, or as an outcome, as when community trust might engender more involvement from intended beneficiaries.

A beginning set of activities was derived from Daigneault and Jacob's (2009) operationalization of participatory evaluation. They broke the evaluation process into four distinct phases: a) question development, b) data collection and analysis, c) developing judgments, and d) reporting and dissemination. Added to these were key variables that were more activity-based than broad principles. For instance, "educate" and "build capacity" are more action-oriented than those in the principles section and were therefore categorized as activities for this research. Outcomes are those variables that are traditionally thought of as transformative evaluation's intended effects. Included in this list are shared understanding, learning, and credible findings.

The theorists provided 68 comments on the 45 original variables between November 20, 2010 and January 10, 2011. One theorist provided more than half (52.9%) of the comments, with another providing 30.9%, and a third 16.2%. (See Table 1.)

Table 1. Number of Theorist Comments During Variable Development

Theorist	Number of Comments	%
Theorist 1	36	52.9%
Theorist 2	21	30.9%
Theorist 3	11	16.2%
Total	68	100.0%

The comments were manually coded for meaning, categorized into groupings of similar meaning, and the categories were described based on the comments. As was expected from the inclusive nature of the list's development, the most prevalent comments were that the variable was not unique to T-PE and was more applicable to evaluation in general (29.4%) or practical participatory evaluation (14.7%) in particular (Table 2). There was some discussion about variable meanings (20.6%) and word choices (13.2%). Thirteen (19.1%) comments agreed with the variable and definition as presented. Some noted a variable was redundant because it was embedded within another variable (2.9%). Full conversations are not included because the theorists were offered confidentiality in their individual remarks. Where remarks are added their identities have been masked.

Table 2. Comment Types During Variable Development

Note	Count	%
Same as all evaluation	20	29.4%
Meaning discussion	14	20.6%
Agreement	13	19.1%
Same as P-PE	10	14.7%
Word choice	9	13.2%
Redundancy	2	2.9%
Total	68	100.0%

Table 3 presents examples of the editing process where variables were identified as reflecting more than just T-PE.

Table 3. Examples of Editing Process During Variable Development

Variable	Definition	Note
Program Theory Examined	Program theories are critically examined	Theorist 2: doesn't differ from any other type of evaluation, surely. Theorist 1: yes, and in theory, a participatory evaluation could be negotiated to be a black box eval.
Informed Decision-Making	Program decision-making is undertaken with information produced by the evaluation	Theorist 1: true of most evals (at least that's the hope) Theorist 2: right, and more true for P-PE than T-PE MAH: Agreed, more P-PE than T-PE, delete stands.

Based on the theorists' comments of this type, 14 of the original 45 variables were deleted, four were combined to create two more appropriate variables, and two were retained because, in spite of their generality, they held significant value for modeling participatory practice. Another four were deleted because a) discussion on the variable's meaning was inconclusive (22.1% of comments) and b) the cognitive load of modeling practice using a large number of variables was of concern (Gargani, 2003).

Following these revisions, the theorists were asked to revisit the list and make additional comments or recommendations. Other than one word change and a recommendation on verb tenses, no further comments were made; on January 10, 2011, this list was considered complete (see Appendix C).

Though they write about somewhat different perspectives on participatory evaluation, the three theorists agreed that this was a satisfactory set of variables with which to begin a description of T-PE. Their well-documented experience and knowledge provides relative substantive and structural validity (Onwuegbuzie & Johnson, 2006). The iterative process by which the list was developed supports an interpretive agreement (Greene, 2007) that would likely withstand scrutiny by other experts in the field. Importantly, these experts also acknowledged that this was most likely an incomplete set of variables but that it was nevertheless a very solid place to start. With this in mind, a step was added to the second phase of the study (described in Chapter 4) where participants were asked to recommend any variables that were missed.

At the outset, and as part of the original research design, the theorists were asked to create and agree upon a logic model-like representation of T-PE using these identified variables. Two of the three were uncomfortable setting out a model of T-PE for two key reasons. First, they felt that participatory evaluation was an approach to evaluation that is not easily put into a single model. Second, and as an extension, they felt uncomfortable creating a model that might then be reified by the evaluation community as *the* T-PE model and they worried that all practice that did not reflect this model would not fit into a T-PE category. For this reason, the first stage of this research was simplified into the identification of the key variables of a transformative participatory approach.

Variables Development Results

Twenty-six variables (listed in Appendix C) were developed through the asynchronous process with the expert panel as described above. These variables were initially derived from the participatory and transformative evaluation literature and were reduced to this smaller set by the

panel through their work with the researcher. They were further edited for clarity during the piloting of the resulting survey (this pilot process is explained in Chapter 4). Table 4 lists each in its final form with its definition. In the text that follows the table, each variable and its definition is presented in italicized text, followed by a more detailed description of the item.

Table 4. Final 26 Variables and Definitions

Principle	Definition
Community Trust	Evaluator works to build lasting trust by developing working relationships with a broad range of stakeholders.
Negotiable Purpose	The purpose of the evaluation is negotiated with stakeholders.
Multiple Method	Evaluator applies multiple methods as appropriate to the evaluation context.
Perspective	Evaluator ensures representation of diverse perspectives by including concerns, values, and interests of stakeholders.
Diverse Perspectives	Technical decision-making (e.g., survey instrument selection, statistical analyses, data presentation) is negotiated with stakeholders.
Negotiable Decision Making	Scope of stakeholder participation in evaluation is not decided ahead of time. Barriers to and supports necessary for participation are identified and negotiated.
Negotiable Participation	Sampling procedures account for community diversity.
Community Sensitive Sampling	Evaluator engages marginalized program stakeholders (e.g., those who might otherwise lack representation) in meaningful participation.
Engage Marginalized Stakeholders	Evaluator engages intended program beneficiaries in meaningful participation.
Engage Intended Beneficiaries	
Activity	Definition
Build Capacity	Evaluator trains stakeholders in the necessary skills to participate in the evaluation.
Share Control	Evaluator negotiates the giving of control of the evaluation to program stakeholders.
Educate	Evaluator educates stakeholders on the value of evaluation.
Use Local Program Knowledge	Evaluation decisions are made using local program knowledge.
Develop Questions	Evaluator collaborates with stakeholders in defining evaluation purpose and evaluation questions.
Collect & Analyze Data	Evaluator collaborates with stakeholders in data collection and analysis.
Develop Judgments &	Evaluator collaborates with stakeholders in interpreting

Recommendations	findings, and formulating judgments and recommendations from the data.
Report & Disseminate	Evaluator collaborates with stakeholders in reporting and disseminating the findings.
Outcome	Definition
Shared Understanding	All participants develop shared understanding of program functions and processes.
Learning	All participants learn new skills.
Credible Findings	Participants see evaluation findings as credible.
Increase Systematic Inquiry	Increase capacity for participants to engage in and use systematic inquiry.
Increase Self-Critique	Increase participants' capacity for self-critique.
Increase Self-Determination	Increase individual self-determination, emancipation and empowerment.
Increase Social Justice	Enhance social justice.
Increase Social Action	Increase social action.
Outcomes Change	Program outcome expectations change as a result of the process.

Principles

Community Trust: Evaluator works to build trust by developing working relationships with a broad range of stakeholders. Community engagement is generally inherent to participatory evaluation. Though participation can mean only key decision makers, in T-PE the broader “program community” would include all those with any stake in the program. The community is particularly valued because the transformative paradigm values knowledge claims grounded in the community most affected by the program.

Negotiable Purpose: The purpose of the evaluation is negotiated with stakeholders. Program stakeholders are generally defined as those having a stake in the program and, therefore, a stake in the evaluation. They are the individuals or groups most affected by the questions addressed and the subsequent findings from these questions. The Program Evaluation Standard U3 (utility #3) (Yarbrough, et al., 2011) is also named Negotiable Purpose. That definition reads,

“Evaluation purposes should be identified and continually negotiated based on the needs of stakeholders” (p. 29). This description, while focused on more utilitarian reasons for negotiating the purpose, recognizes the need for contextually responsive evaluation designs that address purposes grounded in stakeholder needs. Within the transformative, democratic, and inclusive paradigms, using a broad definition of stakeholders when negotiating the purpose with stakeholders reflects respect for the questions emanating from not only those who have direct decision-making over the evaluation, but also those less heard from (Mertens, 2009).

Multiple Method Perspective: Evaluator applies multiple methods as appropriate to the evaluation context. Rather than having method choices driven only by the evaluation question, a multiple method perspective appropriate to the evaluation context centralizes and empowers context in the evaluation design. Context in a culturally aware evaluation design can be defined in numerous ways, including by ethnicity, gender, socioeconomic status, and program delivery locations. Identifying the many cultures in an evaluation context may bring awareness of potentially unaddressed issues of power and inequity. This variable points to the pragmatism and value of a multiple method perspective. The dialogic characteristic of qualitative methods is essential to transformative axiology.

Diverse Perspectives: Evaluator ensures representation of diverse perspectives by including concerns, values, and interests of stakeholders. Key here is the diversity of viewpoints in an evaluation. More specifically, a T-PE evaluation should ensure that the broadest possible stakeholder base is at least considered in the evaluation design. Multiple realities are specifically valued because they help identify where knowledge and power reside. Representation of these different worldviews may give them voice where they previously were without.

Negotiable Decision-Making: Technical decision-making (e.g., survey instrument selection, statistical analyses, data presentation) is negotiated with stakeholders. As with negotiated purpose, negotiated decision-making throughout the evaluation process empowers stakeholders to feel ownership of the evaluation and its results because decisions are grounded in their lived experiences. The empowerment aspect of decision-making reflects the value that the transformative paradigm places on knowledge ownership and power (Sabo 1999; Whitmore, 1988). Stakeholders taking part in decision-making will learn the skills of problem solving and logic. This capacity building is part of the learning inherent to participatory approaches. When this learning occurs through active decision-making, it provides an opportunity for stakeholders to develop new visions of their abilities and self concepts (Sabo Flores, 2008; Vygotsky, 1978).

Negotiable Participation: Scope of stakeholder participation in evaluation is not decided ahead of time. Barriers to and supports necessary for participation are identified and negotiated. Extending negotiation into stakeholder participation reflects the value placed on contextually grounded decisions about participation. Likewise, the Program Evaluation Standards (Yarbrough, et al., 2011) standard U2, “attention to stakeholders,” recommends creating “conditions for stakeholder engagement that are safe, comfortable, and contribute to authentic participation” (p. 25). This includes negotiating any barriers to participation and developing necessary supports to facilitate it.

Community-Sensitive Sampling: Sampling procedures account for community diversity. Recognition of community diversity is a central tenet of a transformative approach to evaluation. Sampling procedures should honor the diversity of the community and invitations to participate need to be viewed as genuine so that the final sample reflects the actual population. Support for

those with special needs (e.g., hearing challenged or English language learners) is considered to allow everyone to participate equally.

Engage Marginalized Stakeholders: Evaluator engages marginalized program stakeholders (e.g., those who might otherwise lack representation) in meaningful participation. Marginalized stakeholders may not always be direct beneficiaries of a program, but instead they may be interested stakeholders who have historically lacked representation. For example, degree completion is often used as an effectiveness measure of learning communities in community colleges. The learning community in this case might focus only on transferable classes, and students seeking certificate completion are marginalized because they do not directly benefit from that learning community. By attending to those on the certificate track, the broader interests of the student body are considered and given voice, which can potentially help identify any unintended consequences of these students' marginalization.

Engage Intended Beneficiaries: Evaluator engages intended program beneficiaries in meaningful participation. This is a key characteristic of a transformative approach that distances it from a practical approach. While practical participatory evaluators focus on engaging key decision makers (Cousins & Earl, 1992; Cousins & Whitmore, 1998), transformative evaluators prefer to hear the voices of those intended to benefit from the program. Their involvement is grounded in social justice because by bringing them into the evaluation they are given voice. This enables increased systematic inquiry among beneficiaries.

Activities

Build Capacity: Evaluator trains stakeholders in the necessary skills to participate in the evaluation. A central concept of participatory evaluation is that of capacity building and learning

because stakeholders often require some training to be able to participate in the evaluation. This variable does not differ significantly from a practical participatory approach except that by including beneficiaries and marginalized stakeholders learning occurs for traditionally neglected groups.

Share Control: Evaluator negotiates the giving of control of the evaluation to program stakeholders. As much as possible, stakeholders take control of an evaluation as capacity is built. The divestment of control is negotiated so that more control is put in stakeholders' hands as they learn the necessary skills and the evaluator moves to more of an advisor role. This is different from empowerment evaluation in that the evaluator begins fully in control of technical decision-making and slowly distributes it as capacity grows.

Educate: Evaluator educates stakeholders on the value of evaluation. Educating stakeholders on the value of evaluation encourages their participation and the potential use of findings. It also engages them in an undertaking that has the potential to be of great value to the program and their community. This is not solely the domain of a transformative approach, but when marginalized stakeholders and beneficiaries are engaged it answers the transformative call for knowledge and power sharing.

Use Local Program Knowledge: Evaluation decisions are made using local program knowledge. This is a component of participation that goes beyond a token level. It is instead valid, respected, and honored participation that is integral to decision-making in the evaluation process. It could be considered a result of stakeholder involvement or the impetus to involve stakeholders in the process.

Develop Questions: Evaluator collaborates with stakeholders in defining the evaluation's purpose and evaluation questions. Collaboration with stakeholders ensures the questions asked by an evaluation are grounded in the needs of stakeholders. This is the first level of activity articulated by Daigneault and Jacob (2009), where important design questions are answered, such as: “What is the rationale for conducting the evaluation? What is the evaluation focus? What are the informational needs the evaluation is intended to answer?” (p. 339) The involvement of stakeholders in developing these questions is predicated on the understanding that the questions that get asked inform the focus of the evaluation, and the participation of stakeholders, especially those historically with little voice, ensures the evaluation maintains a pluralistic stance.

Collect & Analyze Data: Evaluator collaborates with stakeholders in data collection and analysis. This variable is in keeping with Daigenault and Jacob's (2009) second key decision point. As in developing questions with stakeholder collaboration, involving stakeholders in data collection and analysis ensures the data are grounded in the lived experiences of those the program is intending to reach. Involvement at this stage may also be critical to building community trust. For instance, in an evaluation of a prenatal program for single expectant mothers, Whitmore (1994) recruited women from within the target community to be partners in the evaluation and assist with data collection. Each of these mothers was seen by the community as “one of their own” and they were able to gain entre into the lives of program beneficiaries in ways the “white academic” would never have achieved. Moreover, participation in the analysis phase also provides an opportunity for learning and capacity building. While these skills may be the domain of the evaluator, whose expertise is an important and necessarily respected part of the

evaluation endeavor, the potential empowerment benefits of teaching participant stakeholders some basic analyses would not be ignored by a T-PE evaluator.

Develop Judgments & Recommendations: Evaluator collaborates with stakeholders in interpreting findings and formulating judgments and recommendations from the data. Aligned with Daigneault and Jacob's (2009) third decision point, this variable highlights the notion that when stakeholders engage in merit and worth determinations from collected data, they are taking positions that require critical thinking about a program in which they may be highly invested. It may be argued that this increases the likelihood of critical self-reflection because of their possible ownership in the program. Further, interpretation of the findings by those most knowledgeable about the program lends credibility to the interpretations. Hence, recommendations by these stakeholders are expected to take into account the intricacies of program implementation. Historically, as marginalized stakeholders make recommendations on a program ostensibly serving their needs, it puts them in a position of power that has the potential to improve their social capital (Lin, 1999).

Report & Disseminate: Evaluator collaborates with stakeholders in reporting and disseminating the findings. This variable is in keeping with Daigneault Jacob's (2009) fourth decision point. Program stakeholders have the clearest understanding of how findings might be best used by the target audience and may have a broader perspective on who needs to hear evaluation findings. As such, stakeholder involvement in the reporting and dissemination of the findings has the potential to broaden an evaluation's impact.

Outcomes

Shared Understanding: All participants develop shared understanding of program functions and processes. A shared understanding of program functions and processes contributes to stakeholder buy-in and improves the potential use of and learning from the process. This is an intended outcome for participatory evaluation in general and is included here because of its criticality to mapping any PE practice.

Learning: All participants learn new skills. Critical to any participatory practice is the idea that learning is an intended outcome. For transformative evaluators, this is particularly important because of the potential for social justice embedded in knowledge creation.

Credible Findings: Participants see evaluation findings as credible. Central to an evaluation's utility and propriety (Yarbrough et al., 2011), the credibility of findings is judged from multiple positions. Evaluator practice must be of the highest standard for the work to be seen as credible. In participatory forms of evaluation, credibility is also seen from the position of stakeholders. If their participation is genuine, respected, honored, and supported, there is a greater chance that the findings will be credible. This variable is not only important to participatory evaluation, but to all evaluation practice. It was retained in this list because of its importance to mapping evaluation practice and the potential usefulness of a greater understanding of how it is described and valued by T-PE evaluators.

Increase Systematic Inquiry: Increase capacity for participants to engage in and use systematic inquiry. As an explicit outcome of T-PE, systematic inquiry is directly related to self-efficacy and evaluative thinking. Engaging stakeholders in evaluation's systematic inquiry provides a learning opportunity with the potential to increase application of this learning in the future. Successful attempts at evaluation inquiry under the guidance of a trained evaluator can

give stakeholders a concrete image of themselves as accomplished evaluators, and within this proximal zone of development (Sabo Flores, 2008; Vygotsky, 1978), build their self-efficacy towards future applications of systematic inquiry and evaluation.

Increase Self-Critique: Increase participants' capacity for self-critique. In the transformative paradigm, Mertens (2009) recommends the researcher pay particular attention to him or herself. This is because an immersed evaluator must be clear about how personal biases and perspectives might inform decisions and interpretations. Stakeholders in the process must equally attend to self-awareness. Because they are situated within the program's community, they play a dual role of observer and participant. Clarity about these roles and the necessary objectivity is a capacity that might need to be learned. Mertens (2009) discusses evaluators using autoethnographic methods to create reflexivity in their practice. These methods might also be used with participant stakeholders.

Increase Self-Determination: Increase individual self-determination, emancipation and empowerment. In a few instances, both theorists and pilot participants inquired about whether or not these three were similar enough to be grouped under one heading. The researcher chose to keep them together because emancipation and empowerment can be considered different facets of self-determination and not distinct enough to warrant another two variables. This choice was somewhat informed by the need to keep the number of variables low, but more so because together they are related to individual empowerment, a key outcome of a transformative participatory approach (Brisolara, 1998; Cousins & Whitmore, 1998).

Increase Social Justice: Enhance social justice. Social justice and social action are major tenets of a transformative paradigm. A focus on social justice means, among other things,

knowledge is shared and those previously ignored gain voice and an even treatment in the allocation of resources (Mertens, 2009).

Increase Social Action: Increase social action. Social action might be considered the activity that leads to social justice. It was kept separate from social justice because it was considered more of an activity expected to result more directly from the evaluation activities, whereas social justice is a longer-term impact that would result from social action.

Outcomes Change: Program outcome expectations change as a result of the process. The necessity to continually negotiate the purpose and focus of an evaluation is a recognition that as a program evolves its uses and needs often evolve as well (Chambers, 2007). A stagnant evaluation that is not nimble enough to change in order to account for changes in outcome expectations is not useful. Transformative participatory evaluations allow room for the process to inform the choices made in question development, analyses, and judgment development.

Development of T-PE Identifying Statements

As with the list of variables, the researcher generated a set of statements and the theorists vetted them. The researcher created a preliminary set of 21 statements that borrowed from previous participatory research (e.g., Cousins et al., 1992) and drew from Cousins & Earl's (1992, 1995) three-part framework of selection, depth, and control. A fourth philosophical/political dimension was added to align with a transformative approach, with the expectation that it would be instrumental in parsing those in the practical stream from those in the transformative stream (Cousins & Whitmore, 1998; Mertens, 2009; Weaver & Cousins, 2004). (See Appendix D for the full list of statements.) An inclusive approach was also used here to ensure no dimension was under-represented.

The first set of statements was loaded onto the wiki on November 20, 2010, and the theorists were invited to comment. Similar instructions as described above were included here as well (see Appendix A). Theorists provided 30 comments; one provided 53.3%, another 33.3%; and the third provided 13.3% (Table 5).

Table 5. Number of Theorist Comments During Statement Development

Theorist	Total	%
Theorist 1	16	53.3%
Theorist 2	10	33.3%
Theorist 3	4	13.3%
Total	30	100.0%

Comments mostly expressed concerns that the statements did not quite mean T-PE (26.7%) or that they were more P-PE than T-PE (20%). There was also some dialogue among the theorists where, for example, one would state they did not understand how a specific term was used and another theorist would respond with an interpretation (13.3%). (See Table 6)

Table 6. Comment Types During Statement Development

Meaning	Total	%
Inaccurate	8	26.7%
More P-PE	6	20.0%
Dialogue Among Theorists	4	13.3%
Word Choice	4	13.3%
General Comment	3	10.0%
Item Redundancy	2	6.7%
Asking Clarification	1	3.3%
Double-barreled	1	3.3%
Item Too Vague	1	3.3%
TOTAL	30	100.0%

Table 7 presents an example of a dialogue from the edit process that helped inform the appropriate changes to a proposed statement.

Table 7. Example of Editing Process During Statement Development

Statement	Comments	Interpretation	Response
Whenever possible, evaluators should share control of evaluation projects equally with practitioners.	<p>Theorist 1: what does equally mean?</p> <p>Theorist 2: ok it is technical decision-making I think.</p> <p>MH: Good question. I think I can strike equally and still get at the topic of control. I also added “Whenever possible” to this question.</p>	Word choice	Edited

Based on these comments, five of the original 21 statements were deleted because they were more P-PE than T-PE, and another six were removed because they were too vague or they were deemed to be inaccurate depictions of T-PE. In three cases, even though the theorists stated the statements were more general PE than T-PE, they were retained to identify PE modelers (discussed in Chapter 4). Of the remaining 10 statements, only three had not been commented on or edited in the first round. The other seven were slightly edited based upon the recommendations of the theorists. This set of 10 was loaded back onto the wiki on December 3, 2010, and the theorists were asked to again comment on this set.

One theorist provided seven comments directly on these ten statements. The other two did not comment directly on the items. The comments resulted in the deletion of two and the edit of three more so that a set of eight statements was ready for the pilot (see Appendix E). The other two theorists each provided a general comment at that point saying this was a good place to start.

As with the T-PE variables, though the expert panel comprised very knowledgeable theorists, there is no way of entirely describing T-PE so that a comprehensive set of statements could be developed. Given the breadth of experience of this expert panel and of our understanding of practice and the evaluation theories that describe practice, it is reasonable to conclude that the eight statements have relatively adequate content validity and interpretive agreement.

The relativist nature of evaluation practice, and participatory practice in particular, means that every evaluation experience is different and unpredictable. For instance, during the statement development phase, one theorist wrote:

I believe many people who do participatory evaluation do it when the shoe fits but involve themselves in others' approaches, even conventional approaches, depending on the context and information needs driving the evaluation. I just finished an evaluation of an international training program; there was nothing participatory about it. Yet when I do participatory evaluation it tends to be [of one particular type]. Are there T-PE types with a similar modus operandi?

Statements like this during the editing process informed the instructions used in the survey that asked participants to respond to the questions about how their “ideal” practice was reflected (described in the next chapter). It was expected that their ideal practice would be more reflective of their own practice principles, rather than being based on how most of the actual practice plays out.

As voiced in the above comment, the complex nature of evaluation practice may also make an evaluator hesitant to commit in a survey to one theory or principle over another. As another theorist commented, “I find many of these questions hard to answer because it always depends on the context, how I would answer them.” Therefore, to give participants the option to “hedge”

their response, these answer frames were chosen: “strongly disagree,” “slightly disagree,” “slightly agree,” and “strongly agree.” This four-point scale allowed participants to be grouped as either agreeing or disagreeing by requiring them to select the positive or negative side of the center (no mid-point). It also gave them the ability to “hedge” their commitment, yet still be grouped. (Dillman, 2007; Isaac & Michael, 1995).

Identifying Statements Results

As with the set of variables, the researcher generated a preliminary set of statements and the panelists reduced these to a smaller set through online interactions (Table 8). They were also edited for clarity during the survey’s pilot phase described in the next chapter. The statements reflected four components of participatory evaluation: control of the evaluation, depth of participation, selection of stakeholders, and philosophical preference for individual and program transformation and social justice.

Table 8. Eight Identifying Statements

Dimension	Statement
Depth	I always try to involve non-evaluator participants in my evaluations. I prefer not to take on an evaluation unless it has a strong participatory component.
Selection	Program beneficiaries should participate in carrying out evaluation. People representing all important perspectives should be involved in any evaluation.
Control	Evaluators should share technical decision-making with non-evaluator participants.
Transformative Philosophy	Evaluators should help train all legitimate groups to do evaluation. Evaluation should focus on bringing about individual empowerment, emancipation and self-determination. Evaluation should focus on bringing about social justice.

The results from this first phase of the study were used to gather input from evaluation practitioners. Specifically, participants in the second phase of the study were asked how strongly they agreed with each of the eight statements listed above. Those whose levels of agreement indicated they were aligned with T-PE thinking, were then invited to use the list of 26 variables to model their practice. This phase of the study is described in more detail in the next chapter.

CHAPTER 4

PHASE TWO (SURVEY AND MODELING)

The purpose of the second phase of the study was to obtain a graphical representation or a logic model-type reflection of transformative participatory evaluation. Using the statements developed in the previous phase of the research, a survey was developed that would make clearer the differences between T-PE evaluators and other participatory practitioners. The survey was administered to the American Evaluation Association membership and those who responded positively to three of the participatory statements were invited to model their practice to provide information on the importance of these variables in practice and on how the variables interact with each other. Following the modeling, each participant was asked to list any variables they believed were missing from the group. They were also asked if they would be willing to participate in a webinar (described in Chapter 5) to discuss the product of the modeling phase, a most-endorsed version based on all of the submitted models. The survey, model, and proffered variables are covered in this chapter.

Phase Two Methodology

Participants

This portion of the study used a web-based survey to identify transformative participatory evaluators in the American Evaluation Association membership. The survey was followed by a web-based modeling process to obtain graphical representations of the evaluation practices of a subset of these evaluators. The instruments were linked so that those who responded favorably to three key participatory statements on the survey were invited directly into the modeling.

The American Evaluation Association (AEA) is the largest association of evaluators in the world with a current membership of approximately 6,700 individuals, representing all 50 states and more than 60 countries. The AEA membership consists of a broad spectrum of evaluators and evaluation-interested people working in a variety of contexts. There are 47 sub-groups within the organization, called Topical Interest Groups (TIGs), serving the focused interests of the membership. From Advocacy and Policy Change to Theories of Evaluation, the TIGs organize the membership's varied interests and provide support for its many events, including providing peer review for proposed presentations for the association's annual conference. The AEA's official website states that "AEA works to improve evaluation practices and methods, increase evaluation use, promote evaluation as a profession, and support the contribution of evaluation to the generation of theory and knowledge about effective human action." ("American Evaluation Association", n.d.) AEA also supports evaluators by providing training opportunities through workshops, webinars and institutes, organizing an annual conference to encourage knowledge sharing, and maintaining an active presence in the Washington, DC, policy arena.

The AEA membership was selected for participation in the research because the membership is the largest sample of evaluators available in one place, and because the AEA membership practitioners represent a broad range of approaches that was expected to provide a large sample of T-PE evaluators. Following an application process required of anyone wishing to conduct research with the association's membership base, the AEA Executive Director provided the researcher with a current copy of the membership list with contact information.

The initial AEA membership list included 6,632 names and email addresses. There were 10 duplicate email addresses, one with non-recognizable characters, and four associated with more

than one person. Two of the members had completed the whole survey during the pilot stage (described below) so they were therefore eliminated from the sample. The final sample included 6,615 potential subjects. In response to the first invitation, 546 respondents opted out (446 used the survey's built-in opt-out feature, and 118 emailed the researcher to opt out), 84 addresses bounced, and 43 email addresses resulted in "out of office" replies for the remainder of the survey period. Another 208 opted out in response to a follow-up reminder. Combined, the opt-out rate was 11.5%. Overall, 1,323 individuals began the survey, but 59 dropped out before providing more than cursory data, and another 36 did not provide complete data. The final sample providing complete data was 1,228, yielding a response rate for the entire sample of 18.56%; removing the bounced addresses, the response rate rises to 18.80%. (See Table 9.)

One might consider an 18.8% survey response rate low. If the intention were to develop a model generalizable to all T-PE evaluators then it would be difficult to support such a claim with only ~18% of the target population even venturing into the study. The purpose of the study was not to develop such a model but to develop a deeper description of T-PE practice. Therefore, the survey's purposes were to identify a group of T-PE evaluators that fit the profile developed in the first phase and to invite them to the modeling phase. With 1,228 respondents completing the survey, 561 responding positively to the three PE questions, and 240 completing a model, the process satisfactorily identified an adequate number of practitioners that fit the profile this research was designed to study.

Table 9. Survey Response Rates

	N
Total Population	6,632
Duplicates/Excluded	17
Invited	6,615

Opted Out	754
Bounced	84
Incomplete or No Data	4,549
Complete Data	1,228
Response Rate	18.56%

Survey Sample

Almost all the survey respondents (Table 10) were active practitioners (94%), with about two-thirds (66%) conducting six or fewer evaluations a year and just over a quarter (28%) conducting seven or more a year. Almost half (48.4%) have been in the evaluation field for 10 years or fewer and they are evenly split on primary or secondary identity as evaluators (44.6% and 44.2%, respectively). Most (43%) see themselves as having intermediate knowledge and experience and have either master's or doctorate degrees (43.8% and 44%, respectively). There is a great diversity of disciplines represented in the degrees held by this sample, with education and psychology standing above the rest.

As practitioners, participants said they prefer a broad range of theoretical orientations. They were given these instructions for this question:

Is your preferred theoretical orientation similar to any of these? I know that many evaluators say that they design evaluations that are context specific and none of these orientations covers every evaluation. But, I also know that you probably have a perspective you 'prefer.'

Table 10 provides an overview of the evaluation characteristics of the respondents.

Utilization-focused evaluation (UFE) was most endorsed (24.3%) and participatory evaluation gained about half as many (11.5%) selections. The respondents typically either do a mixture of internal and external evaluations (34.4%) or external (32.1%). Most do program evaluations (88.4%) in a variety of contexts, with nonprofits (36.7%), health (35.9%), and K–12 education

(34.2%) holding the top three spots. Their settings are mostly academic (38.6%), non-profits (25.6%) or private business/consulting (22.7%), and they general work in North America (75.1%).

Table 10. Survey Participants' Characteristics

Evaluations Per Year	N	%
1–3	528	39.9%
4–6	345	26.1%
7 or more	370	28.0%
None	75	5.7%
Blank	5	0.4%
Total	1,323	100%
Years in Evaluation		
Less than two	87	6.6%
2–5	280	21.2%
6–10	273	20.6%
11–15	227	17.2%
16–20	113	8.5%
More than 20	185	14.0%
Blank	158	11.9%
Total	1,323	100%
Evaluation Identity		
Primary	590	44.6%
Secondary	585	44.2%
Not my professional identity	62	4.7%
Blank	86	6.5%
Total	1,323	100%
Evaluation Knowledge and Experience		
A relative beginner	190	14.4%
At an intermediate level	569	43.0%
At an advanced level	403	30.5%
Blank	161	12.2%
Total	1,323	100%
Highest Education Level Completed		
High school/some college	1	0.1%
Associate's degree	2	0.2%
Bachelor's degree	70	5.3%
Master's degree	580	43.8%
Doctoral degree	582	44.0%
Blank	88	6.7%
Total	1,323	100%

Field of Your Highest Degree		
Education	229	17.3%
Psychology	179	13.5%
Evaluation/Research methods	118	8.9%
Public health	103	7.8%
Public policy	76	5.7%
Sociology	62	4.7%
Business	38	2.9%
Economics	29	2.2%
Social welfare	29	2.2%
Anthropology	20	1.5%
Advanced quantitative methods	9	0.7%
Nursing/Medicine	9	0.7%
School administration	6	0.5%
Advanced qualitative methods	1	0.1%
Art/Music	1	0.1%
Other		
Applied social science	156	48.0%
Social science	78	24.0%
Natural science	30	9.2%
Humanities	27	8.3%
Applied science	14	4.3%
Formal science	9	2.6%
Interdisciplinary	6	1.7%
(blank)	5	0.3%
Subtotal	325	100%
Decline to answer	2	0.2%
(blank)	87	6.6%
Total	1,323	100%
Preferred Theoretical Orientation		
Utilization-focused	321	24.3%
Participatory evaluation	152	11.5%
Evaluation research	97	7.3%
Theory-driven	95	7.2%
Developmental evaluation	57	4.3%
Empowerment evaluation	37	2.8%
CIPP Model	30	2.3%
Stakeholder evaluation	26	2.0%
Social justice-driven	22	1.7%
Fourth generation evaluation	5	0.4%
Connoisseurship evaluation	1	0.1%
My theoretical orientation is not listed here	41	3.1%
I do not have a preferred theoretical orientation	166	12.5%
I do not know enough about these to select one	112	8.5%
(blank)	161	12.2%

	Total	1,323	100%
Role as an Evaluator			
External		424	32.1%
Internal		277	20.9%
Mix of internal & external		455	34.4%
(blank)		167	12.6%
	Total	1,323	100%
Primary Type(s) of Evaluations Performed			
Program evaluations		1,169	88.4%
Performance auditing/monitoring/reviewing		458	34.6%
Policy evaluations		330	24.9%
Curricula evaluations		266	20.1%
Evaluation of research		261	19.7%
Student/Trainee evaluations		174	13.2%
Personnel evaluations		162	12.2%
Consumer evaluations		84	6.4%
Product evaluations		57	4.3%
I do not do evaluations		4	0.3%
	Total	2,965	
Primary Context(s) of Evaluations Performed			
Nonprofits		486	36.7%
Health/Public health		475	35.9%
K–12 education		452	34.2%
Higher education		399	30.2%
Youth development		332	25.1%
Adult education		318	24.0%
Government		288	21.8%
Child care/Early childhood education		281	21.2%
Advocacy and policy change		258	19.5%
Human services		256	19.4%
Public policy/Public administration		251	19.0%
Evaluation methods		241	18.2%
STEM		227	17.2%
Educational technologies		212	16.0%
Special needs populations		196	14.8%
Organizational behavior		193	14.6%
Workforce/Economic development		188	14.2%
Alcohol or drug abuse		178	13.5%
Foundations		170	12.9%
Social justice		156	11.8%
International/Cross-cultural		143	10.8%
Environmental programs		141	10.7%
Social work		139	10.5%
Human development		129	9.8%
Indigenous peoples		105	7.9%

Business and industry	93	7.0%
Law or criminal justice	80	6.1%
Medicine	78	5.9%
Disaster/Emergency management	76	5.7%
Gender rights	71	5.4%
Human resources	71	5.4%
Information systems	67	5.7%
Media	47	3.6%
LGBT	41	3.1%
Total	6,838	
Primary Setting(s) of Evaluations Performed		
College or university	511	38.6%
Nonprofit foundation/organization	339	25.6%
Private business or consulting	300	22.7%
Federal government agency	108	8.2%
State/Provincial government agency	104	7.9%
School system	91	6.9%
Local government agency	40	3.0%
Total	1,493	
Primary Location(s) of Evaluations Performed		
North America	994	75.1%
Asia	104	7.9%
Africa	102	7.7%
Europe	56	4.2%
South America	48	3.6%
Central America	35	2.7%
Australia/New Zealand	31	2.3%
Total	1,370	

Model Sample

During the development of the identifying statements (described in the previous chapter), three were mentioned as being more generally applicable to participatory evaluation and not exclusively the domain of T-PE (Table 11). Therefore, to avoid asking every evaluation practitioner to model their practice, these three were kept in the group of eight statements so they could be used here to separate participatory from non-participatory evaluators. If respondents somewhat or strongly agreed with all three of these statements—identifying themselves as participatory evaluators—they were invited to participate in the modeling phase.

Table 11. Three Statements Used to Identify Participatory Evaluators

Statement
Evaluators should share technical decision-making with stakeholders.
I always try to involve stakeholders in my evaluations.
I prefer not to take on an evaluation unless it has a strong participatory component.

Of those who participated in the survey (n=1,323), 42.4% (n=561) agreed with these three statements. Of these, 78.8% (n=442) agreed to continue and 42.8% (n=240) completed a model. The modeling phase was linked directly from the end of the survey.

Responses to the full set of eight T-PE statements were used to place each modeling phase participant into one of three groups. Specifically, as shown in Table 12, of the 240 who completed models, 142 responded positively to all eight T-PE statements and were therefore labeled T-PE evaluators. The small group who disagreed with the key principles of social justice and empowerment but agreed with the other six statements comprised a comparison group and were labeled as practical participatory evaluators (P-PE) (n=16). The others were labeled as participatory evaluators (PE) (n=82).

Table 12. Participants' Participatory Evaluation Categories

Grouping	n	%
T-PE	142	59.2%
P-PE	16	6.7%
PE	82	34.2%
Total	240	100%

Table 13 offers more detail on each of the three groups. For example, the P-PE group was a more active and more experienced group of evaluators than the T-PE or PE evaluators (43.8% conducting seven or more evaluations a year vs. 33.1% and 31.7%, respectively) and all groups

about evenly considered evaluation as their primary identity (52.1% to 57.3%). The P-PE evaluators also considered themselves more experienced and more knowledgeable than either of the other two groups (62.5% described themselves as “advanced” vs. 33.8% and 40.2%) but were slightly less educated (43.8% with doctorates vs. 50.7% and 47.6%).

More of the P-PE evaluators considered their theoretical orientations to be aligned with utilization focused evaluation (56.3%) than the T-PE (29.6%) or the PE (31.7%) groups. Fewer T-PE evaluators were external evaluators (28.9% vs. 43.8% and 40.2%). Most in each of these subgroups reported doing program evaluations (between 97% and 100%) in health (between 25% and 48%) and non-profit (between 44% and 56%) contexts, and the P-PE group reported working most often in special needs or business and industry (43.8% each).

The majority of the T-PE and PE groups worked in colleges or universities (from 41% to 43%), non-profit organizations (30%) or private business or consulting (from 25% to 32%). The P-PE group worked mostly in private business or consulting (60%) and colleges or universities (25%). Most of the modelers worked in North America (between 81% and 92% of each group), with 19% of the P-PE group doing work in Asia.

Table 13. Modeling Participants’ Characteristics

Evaluations Per Year	Modeling					
	T-PE		P-PE		PE	
	Total	%	Total	%	Total	%
1–3	57	40.1%	5	31.3%	36	43.9%
4–6	38	26.8%	4	25.0%	20	24.4%
7 or more	47	33.1%	7	43.8%	26	31.7%
Total	142	100%	16	100%	82	100%
Years in Evaluation						
Less than two	8	5.6%	0	0%	5	6.1%
2–5	37	26.1%	2	12.5%	18	22.0%
6–10	36	25.4%	3	18.8%	18	22.0%

11–15	30	21.1%	4	25.0%	16	19.5%
16–20	11	7.7%	4	25.0%	9	11.0%
More than 20	20	14.1%	3	18.8%	15	18.3%
Blank	0	0%	0	0%	1	1.2%
Total	142	100%	16	100%	82	100%
Evaluation Identity						
Primary	74	52.1%	9	56.3%	47	57.3%
Secondary	64	45.1%	6	37.5%	33	40.2%
Not my professional identity	4	2.8%	1	6.3%	2	2.4%
Total	142	100%	16	100%	82	100%
Evaluation Knowledge and Experience						
A relative beginner	14	9.9%	0	0%	11	13.4%
At an intermediate level	80	56.3%	6	37.5%	38	46.3%
At an advanced level	48	33.8%	10	62.5%	33	40.2%
Total	142	100%	16	100%	82	100%
Highest Education Level Completed						
Bachelor's degree	7	4.9%	1	6.3%	4	4.9%
Master's degree	63	44.4%	8	50.0%	38	46.3%
Doctoral degree	72	50.7%	7	43.8%	39	47.6%
Blank	0	0%	0	0%	1	1.2%
Total	142	100%	16	100%	82	100%
Field of your Highest Degree						
Education	28	19.7%	3	18.8%	11	13.4%
Psychology	23	16.2%	1	6.3%	12	14.6%
Evaluation/Research methods	14	9.9%	1	6.3%	8	9.8%
Public health	11	7.7%	1	6.3%	7	8.5%
Public policy	5	3.5%	2	12.5%	3	3.7%
Sociology	4	2.8%	0	0%	4	4.9%
Business	3	2.1%	3	18.8%	3	3.7%
Economics	4	2.8%	0	0%	0	0%
Social welfare	4	2.8%	0	0%	0	0%
Anthropology	6	4.2%	0	0%	1	1.2%
Advanced quantitative methods	1	0.7%	0	0%	0	0%
Nursing/Medicine	0	0%	0	0%	3	3.7%
School administration	4	2.8%	0	0%	0	0%
Other	0	0%	0	0%	0	0%
Applied social science	15	10.6%	3	18.8%	13	15.9%
Social science	9	6.3%	2	12.5%	7	8.5%
Natural science	1	0.7%	0	0%	4	4.9%
Humanities	4	2.8%	0	0%	2	2.4%
Applied science	2	1.4%	0	0%	2	2.4%
Interdisciplinary	3	2.1%	0	0%	1	1.2%
Total	142	100%	16	100%	82	100%
Preferred Theoretical Orientation						
Utilization-focused	42	29.6%	9	56.3%	26	31.7%

Participatory evaluation	34	23.9%	2	12.5%	10	12.2%
Evaluation research	3	2.1%	0	0%	7	8.5%
Theory-driven	9	6.3%	1	6.3%	8	9.8%
Developmental evaluation	8	5.6%	0	0%	3	3.7%
Empowerment evaluation	7	4.9%	0	0%	4	4.9%
CIPP Model	5	3.5%	1	6.3%	1	1.2%
Stakeholder evaluation	1	0.7%	0	0%	1	1.2%
Social justice-driven	5	3.5%	0	0%	0	0%
Fourth generation evaluation	1	0.7%	0	0%	0	0%
My theoretical orientation is not listed here	4	2.8%	0	0%	2	2.4%
I do not have a preferred theoretical orientation	13	9.2%	3	18.8%	11	13.4%
I do not know enough about these to select one	10	7.0%	0	0%	9	11.0%
Total	142	100%	16	100%	82	100%
Role as an Evaluator						
External	41	28.9%	7	43.8%	33	40.2%
Internal	45	31.7%	5	31.3%	15	18.3%
Mix of internal & external	55	38.7%	3	18.8%	33	40.2%
Blank	1	0.7%	1	6.3%	1	1.2%
Total	142	100%	16	100%	82	100%
Primary Type(s) of Evaluations Performed						
Program evaluations	141	99.3%	16	100%	80	97.6%
Performance auditing/monitoring/reviewing	53	37.3%	8	50%	27	32.9%
Policy evaluations	41	28.9%	5	31.3%	21	25.6%
Curricula evaluations	33	23.2%	4	25%	14	17.1%
Evaluation of research	28	19.7%	3	18.8%	20	24.4%
Student/Trainee evaluations	19	13.4%	0	0%	13	15.9%
Personnel evaluations	17	12%	4	25%	8	9.8%
Consumer evaluations	13	9.2%	1	6.3%	3	3.7%
Product evaluations	5	3.5%	0	0%	1	1.2%
Total	350		41		187	
Primary Context(s) of Evaluations Performed						
Nonprofits	63	44.4%	7	43.8%	46	56.1%
Health/Public health	65	45.8%	4	25.0%	39	47.6%
K–12 education	55	38.7%	4	25.0%	32	39.0%
Higher education	41	28.9%	5	31.3%	30	36.6%
Youth development	50	35.2%	3	18.8%	27	32.9%
Adult education	38	26.8%	6	37.5%	22	26.8%
Government	33	23.2%	5	31.3%	25	30.5%
Child care/Early childhood education	34	23.9%	2	12.5%	20	24.4%
Advocacy and policy change	43	30.3%	2	12.5%	21	25.6%

Human services	36	25.4%	4	25.0%	21	25.6%
Public policy/Public administration	30	21.1%	5	31.3%	18	22.0%
Evaluation methods	27	19.0%	4	25.0%	21	25.6%
STEM	22	15.5%	2	12.5%	14	17.1%
Educational technologies	26	18.3%	5	31.3%	20	24.4%
Special needs populations	35	24.6%	7	43.8%	16	19.5%
Organizational behavior	25	17.6%	6	37.5%	23	28.0%
Workforce/Economic development	16	11.3%	4	25.0%	14	17.1%
Alcohol or drug abuse	20	14.1%	2	12.5%	16	19.5%
Foundations	25	17.6%	4	25.0%	17	20.7%
Social justice	39	27.5%	1	6.3%	7	8.5%
International/Cross-cultural	25	17.6%	2	12.5%	7	8.5%
Environmental programs	12	8.5%	3	18.8%	12	14.6%
Social work	22	15.5%	1	6.3%	10	12.2%
Human development	10	7.0%	2	12.5%	9	11.0%
Indigenous peoples	15	10.6%	1	6.3%	8	9.8%
Business and industry	7	4.9%	7	43.8%	8	9.8%
Law/Criminal justice	15	10.6%	0	0.0%	8	9.8%
Medicine	9	6.3%	2	12.5%	7	8.5%
Disaster/Emergency management	9	6.3%	4	25.0%	0	0.0%
Gender rights	12	8.5%	1	6.3%	4	4.9%
Human resources	7	4.9%	4	25.0%	4	4.9%
Information systems	9	6.3%	2	12.5%	7	8.5%
Media	5	3.5%	2	12.5%	3	3.7%
LGBT	9	6.3%	1	6.3%	1	1.2%
Total	889		114		537	
Primary Setting(s) of Evaluations Performed						
College or university	58	40.8%	4	25%	29	35.4%
Nonprofit foundation/organization	42	29.6%	1	6.3%	19	23.2%
Private business or consulting	35	24.7%	10	62.5%	23	28.1%
Federal government agency	7	4.9%	2	12.5%	4	4.9%
State/Provincial government agency	14	9.9%	1	6.3%	8	9.8%
School system	14	9.9%	1	6.3%	7	8.5%
Local government agency	6	4.2%	0	0%	4	4.9%
Total	176		19		94	
Primary Location(s) of Evaluations Performed						
North America	117	82.4%	13	81.3%	74	90.2%
Asia	16	11.3%	3	18.8%	5	6.1%
Africa	10	7.0%	2	12.5%	7	8.5%
Europe	3	2.1%	2	12.5%	4	4.9%
South America	4	2.8%	1	6.3%	4	4.9%
Central America	3	2.1%	1	6.3%	4	4.9%
Australia/New Zealand	7	4.9%	0	0%	4	4.9%
Total	160		22		102	

Although 561 participants were invited to complete models, only 240 successfully did so. Approximately one fifth ($n=112$) of those who did not complete models but were invited to do so were categorized as T-PE evaluators. These 112 non-completers selected participatory as a preferred theoretical orientation slightly more than the survey sample, 29% selected participatory vs. 23.9% of those who did complete.

After participants completed their models they were asked if there was any variable and its associated definition that they thought was missing from the list. Respondents to this question are therefore a subgroup of the survey population (See Table 14). They were slightly more likely to be active evaluators than the broader survey population—just over a third (35.8%) of those who responded reported working on seven or more evaluations a year, compared to 28.0% of the broader survey population. They were also more experienced in evaluation: 75% had practiced six or more years versus 60.3% in the full survey population; 50.8% called themselves intermediate and 40% called themselves advanced versus 43.0% and 30.5% in the overall survey population, respectively. And finally, more considered evaluation their primary professional identity (59.21% of those who commented vs. 44.6% of the whole survey population), rather than a secondary identity (39.2% vs. 44.2%, respectively).

Those who provided potential variables were also slightly more educated than the overall survey population; 49.2% (vs. 43.8% of the overall sample) had completed master's degrees and 48.3% (vs. 44.0%) had completed doctoral degrees. Their top five disciplinary areas were very similar to those in the survey population, with education most frequently selected (17.5% compared to 17.3% overall), followed by psychology (14.2% vs. 13.5%), evaluation/research methods (9.2% vs. 8.9%), public health (6.7% vs. 7.8%), and public policy (5.8% vs. 5.7%).

More frequently than the overall survey population, this subsample selected utilization focused (34.2% vs. 24.3%), participatory (17.5% vs. 11.5%), empowerment (6.7% vs. 2.8%) and CIPP model (5.0% vs. 2.3%) evaluation as their primary approaches. They less often had no preferred theoretical orientation (8.3% vs. 12.5%) or did not know enough about these to select one (5.0% vs. 8.5%). About equal numbers considered themselves external evaluators (34.2% vs. 32.1%), but slightly more were internal evaluators (23.3% vs. 20.9%) or a mix of internal and external (40.8% vs. 34.4%).

Almost all of the participants who created models said they do program evaluations (99.2%, compared to 88.4% of the overall sample) and more of them reported doing policy evaluation (31.7%) than the general survey population (24.9%). The practice context of this subsample is more diverse than that of the broader survey population. The contexts they work in most often are nonprofits (50.0% compared to 36.7% in the general sample), health and public health (47.5% vs. 35.9%), youth development (39.2% vs. 25.1%), K–12 education (36.7% vs. 34.2%), and government (30.0% vs. 21.8%). A greater proportion of the subsample work in private business or consulting (30.8% vs. 22.7%) and in North America (87.5% vs. 75.1%).

Table 14. Added Variable Participants' Characteristics

Evaluations Per Year	Total	%
1–3	48	40.0%
4–6	29	24.2%
7 or more	43	35.8%
Total	120	100%
Years in Evaluation		
Less than two	6	5.0%
2–5	23	19.2%
6–10	30	25.0%
11–15	25	20.8%
16–20	12	10.0%
More than 20	23	19.2%

Blank	1	0.8%
Total	120	100%
Evaluation Identity		
Primary	71	59.2%
Secondary	47	39.2%
Not my professional identity	2	1.7%
Total	120	100%
Evaluation Knowledge and Experience		
A relative beginner	11	9.2%
At an intermediate level	61	50.8%
At an advanced level	48	40.0%
Total	120	100%
Highest Education Level Completed		
High school/some college	0	0%
Associate's degree	0	0%
Bachelor's degree	3	2.5%
Master's degree	59	49.2%
Doctoral degree	58	48.3%
Total	120	100%
Field of your Highest Degree		
Education	21	17.5%
Psychology	17	14.2%
Evaluation/Research methods	11	9.2%
Public health	8	6.7%
Public policy	7	5.8%
Sociology	5	4.2%
Business	5	4.2%
Economics	2	1.7%
Social welfare	3	2.5%
Anthropology	3	2.5%
Advanced quantitative methods		
Nursing/Medicine	1	0.8%
School administration	1	0.8%
Advanced qualitative methods		
Art/Music		
(blank)	1	0.8%
Other	35	29.2%
Applied social science	19	54.3%
Social science	8	22.9%
Natural science	2	5.7%
Humanities	2	5.7%
Applied science	1	2.9%
Formal science		
Interdisciplinary	2	5.7%
(blank)	1	2.9%

	Total	120	100%
Preferred Theoretical Orientation			
Utilization focused	41	34.2%	
Participatory evaluation	21	17.5%	
Evaluation research	7	5.8%	
Theory driven	10	8.3%	
Developmental evaluation	4	3.3%	
Empowerment evaluation	8	6.7%	
CIPP Model	6	5.0%	
Stakeholder evaluation			
Social justice driven	3	2.5%	
Fourth generation evaluation			
Connoisseurship evaluation			
My theoretical orientation is not listed here	4	3.3%	
I do not have a preferred theoretical orientation	10	8.3%	
I do not know enough about these to select one	6	5.0%	
Total	120	100%	
Primary Role as Evaluator			
External	41	34.2%	
Internal	28	23.3%	
Mix of internal & external	49	40.8%	
Total	120	100%	
Primary Type(s) of Evaluations Performed			
Program evaluations	119	99.2%	
Performance auditing/monitoring/ reviewing	42	35.0%	
Policy evaluations	38	31.7%	
Curricula evaluations	23	19.2%	
Evaluation of research	21	17.5%	
Student/Trainee evaluations	15	12.5%	
Personnel evaluations	15	12.5%	
Consumer evaluations	10	8.3%	
Product evaluations	2	1.7%	
I do not do evaluations			
Total	285		
Primary Context(s) of Evaluations Performed			
Nonprofits	60	50.0%	
Health/Public health	57	47.5%	
K–12 education	44	36.7%	
Higher education	30	25.0%	
Youth development	47	39.2%	
Adult education	30	25.0%	
Government	36	30.0%	
Child care/Early childhood education	27	22.5%	
Advocacy and Policy Change	33	27.5%	
Human services	35	29.2%	

Public policy/Public administration	34	28.3%
Evaluation methods	31	25.8%
STEM	17	14.2%
Educational technologies	21	17.5%
Special needs populations	33	27.5%
Organizational behavior	28	33.3%
Workforce/Economic development	13	10.8%
Alcohol or Drug Abuse	23	19.2%
Foundations	27	22.5%
Social justice	25	20.8%
International/Cross-cultural	15	12.5%
Environmental programs	16	13.3%
Social work	22	18.3%
Human development	10	8.3%
Indigenous peoples	13	10.8%
Business and industry	8	6.7%
Law/Criminal justice	12	10.0%
Medicine	8	6.7%
Disaster/Emergency management	5	4.2%
Gender rights	9	7.5%
Human resources	9	7.5%
Information systems	8	6.7%
Media	6	5.0%
LGBT	7	5.8%
Total	799	
Primary Setting(s) of Evaluations Performed		
College or university	44	36.7%
Nonprofit foundation/organization	31	25.8%
Private business or consulting	37	30.8%
Federal government agency	6	5.0%
State/Provincial government agency	13	10.8%
School system	8	6.7%
Local government agency	6	5.0%
Total	145	
Primary Location(s) of Evaluations Performed		
North America	105	87.5%
Asia	10	8.3%
Africa	10	8.3%
Europe	5	4.2%
South America	5	4.2%
Central America	5	4.2%
Australia/New Zealand	6	5.0%
Total	146	

Procedures: Pilot Survey

The survey instrument was piloted twice in spring 2011. The first pilot tested the technical process and garnered feedback on the variables and statements. A convenience sample was obtained by posting a call for volunteers to the EVALTALK listserv, the American Evaluation Association's Collaborative, Participatory & Empowerment Topical Interest Group (CP&E TIG) webpage, and to the researcher's Facebook page on March 1, 2011 (Appendix F). By March 9, 30 people had responded and they were sent invitations to the Lime Service link for the pilot. Twenty completed the survey within two weeks and the survey was closed.

The pilot version of the survey had four opportunities for open-ended feedback, evenly spaced through the survey. Also, comment boxes were placed next to each of the principles, and respondents were asked to provide feedback on whether they were understandable and if they could be improved. This garnered a total of 387 comments. As shown in Table 15, most of these were approval statements (41.6%) indicating agreement with the offered principle and definition. Another 21.5% offered clear definitional edits, 16.5% provided questions about the meanings of principles or definitions, and 14.2% provided word choice recommendations for clarity (e.g., "stakeholder" instead of "non-evaluator"). The remaining 6% consisted of comments responding to a survey question, offering their own reflections, or comments about the technical workings of the survey.

Table 15. Survey Pilot Comments

Comments	Total	%
Approval Statements	161	41.60%
Definition Edits	83	21.45%
Meaning	64	16.54%
Word choice	55	14.21%
Other	24	6.20%

Total	387	100.0%
-------	-----	--------

Each variable definition, most variable names, and three of the statements were edited based on pilot input. Edits were relatively minor and only made for clarification of meanings. (For specific edits to the variables and statements across all phases, see Appendix P)

During the first pilot, participants experienced some problems accessing the modeling software. Ultimately it was determined that the modeling software was not prepared to handle the broad array of possible web browsers. The modeling software coding and the survey design were edited to accommodate the technical problems and they were tested before the second pilot.

A second pilot was conducted in early May 2011 to test the technical functioning and to gain any final insights from participants. As a convenience sample, the first pilot group was invited to return to retake it and a convenience sample of the researcher's classmates at Claremont Graduate University were also invited (n=20; 9 returnees and 11 new recruits). In this version, one open text box was offered at the end for comments. Though six comments were provided, they added no new understanding. The technical process worked smoothly for all participants. The instrument was ready for administration by mid-May 2011.

Procedures: Survey Administration

The final survey (see Appendix G) was written using Lime Survey's open-source Linux-based program and was hosted online using Lime Service's hosting service. Personalized invitations were sent to all 6,615 email addresses in AEA's membership list on May 24, 2011. The invitation included an explanation of the research and a summary of informed consent (see

Appendix I). A more detailed informed consent was included as the first page of the survey. Potential participants were also offered an incentive: those who completed the survey could be entered into a drawing for a \$200 Amazon.com gift certificate. One week after the initial email, a personalized follow-up email reminder was sent to individuals who had not yet completed the survey (Appendix J). An opt-out option was included in each invitation and reminder. The survey was closed on June 13, 2011.

On the final screen of the survey, those participants who were categorized as participatory based on their responses to the eight identifying statements were invited to model their practice. They were provided with a link that took them directly to the modeling interface where they were presented with an informed consent form for the modeling phase. One week after the first survey reminder, a separate reminder was sent to any individuals who had been invited to move from the survey to the modeling phase but who had not yet modeled their practice (Appendix K). The modeling software also closed on June 13, 2011.

In response to the reminder to complete the modeling, three modeling participants informed the researcher that they had completed models but their data did not make it into the system. Though the modeling process was tested and found to be technically sound, it is possible that some technical issue occurred. It is also possible that some operator error occurred. No technical issues could be replicated through user testing before and after the survey was implemented.

Instruments: Survey

The survey's primary purpose was to identify and engage T-PE practitioners in the subsequent phases of the research. Therefore, one question was placed early in the instrument to screen non-practitioners from being burdened by unrelated questions. If they were not practicing

evaluators, they were moved very quickly to the end and thanked for their participation.

Likewise, participants were also asked early in the survey to respond to the eight statements that would identify them as T-PE evaluators.

The survey was also used to elicit important corollary information regarding participants' practice, experience, training, and philosophical orientation toward evaluation. Responses to these questions were used to further understand the different subgroups identified above.

A section of the survey was also developed to introduce participants to the 26 variables developed in this study's first phase. Those who agreed with the three statements regarding participatory evaluation were asked to rank the 26 variables by importance in their preferred practice. Having the participants rank the importance of each variable increased their familiarity with and understanding of them. This was done to increase the validity of the models.

Because it would be too burdensome and unreliable to rank 26 items in one group (Streiner & Norman, 2008), the variables were presented in the three groupings of *principles*, *activities*, and *outcomes* described earlier. Participants were given these instructions for the ranking:

On this and the next two pages, I list a number of variables that are theorized to be important to participatory evaluation. I would like to know how important they are in your practice. To narrow the effort, I have categorized these variables and I ask you to only rank the top few in each category.

From the top box, select in order of importance, from highest to lowest, the four principles or activities you see as most important in your participatory evaluation practice.

Once they had ranked the variables, participants were asked to continue on to a modeling phase where they would use these same variables to model their practice.

Instruments: Modeling Procedure

The 26 T-PE variables produced in an earlier phase of this research were loaded into a web-hosted model building software package designed and managed by Dr. John Gargani (2003). The software is a web browser-based platform where multiple stakeholders can independently draw program theory models and save them as products. Gargani's intent in designing this software was to be able to show the disparate perspectives on any given program's theory. The current study employed this software as a way to bring practitioners from around the world together to model their evaluation practices so the researcher could use those models to create a single most-endorsed model of practice.

The first page of the modeling software was a login that asked participants for the email address through which they were first recruited. After login they were presented an informed consent page, followed by a page that provided detailed instructions and a video that walked participants through the model building process. A note on the instruction page asked participants to consider the list of variables presented for their modeling and to identify any other important variable that was not already listed, because they would be given an opportunity on the final screen to provide their input. The next screen was the model-building layout. (See Appendix L for each page of the modeling interface.)

Participants were asked to model their practice by answering a question provided across the top of the modeling page: "How do you ensure stakeholder involvement and what outcomes do you intend to create?" On the left side of the page, the variables were available for participants to drag into the white space on the right. When a mouse hovered over a variable (rollover) its definition was shown. To anchor the models, the phrase "Stakeholder Involvement" was

included in the open white space at the outset for every participant. Participants were to draw arrows from and to any variables in the white space. When they were finished, they used the “Done” button in the upper right hand corner to save their models before moving on.

The next page asked if any variables were missing from the list, and a final page asked if they would participate in a webinar to discuss a model derived from the submissions of all participants and if they would like to opt into the drawing for a \$200 Amazon gift certificate.

Analyses

The T-PE and P-PE models produced in this phase were statistically compared using differential item functioning (DIF) analysis. DIF was used in this context to test whether groups were more or less likely to include specific links in their models. If the members of one group tended to include a particular link in their practice models more than the members of another group, that link would then help discriminate how the two groups conceive of their practice.

The parameters for the DIF analysis were estimated by using an expanded Rasch model. A Rasch model is the simplest item response theory (IRT) model, in the sense that it takes into account only two variables, fewer than other IRT models (Emberson & Reise, 2000). The DIF analysis adds an interaction term to the two terms found in the traditional Rasch Model. The statistical model is presented in Equation 1:

$$\eta_{DIF} = \theta_j - \delta_i + \lambda_{gi} \quad \text{Equation (1)}$$

Here, η is the log of the odds (logit) that person j will include link i in his or her practice model; θ_j is the level of model complexity preferred by person j , where complexity is operationalized as the number of links in a model. Respondents with higher θ_j estimates tend to

construct models of their practice that are more complex (i.e., logic models with more links). Further, δ_i is the relative difficulty of including link i in a model, or alternatively the relative likelihood that a link will be excluded from a model. Links with lower δ_i estimates tend to be included in the practice models of more respondents (i.e, less difficult and more common among modelers). And λ_{gi} is a group-by-item interaction term, where membership in any group (g) interacts with the difficulty of endorsing and item (i). Items with positive, 0, or negative λ_{gi} estimates are found more, equally, or less often, respectively, in T-PE practice models than the P-PE practice models.

To determine whether λ_{gi} is statistically different from 0 (in which case the two groups included the link in their models with the same frequency), Wald tests were performed for each estimated λ_{gi} . A Wald test is a Z test that is typically used in DIF analysis. The cutoff for statistical significance (alpha) was set to 0.05. No adjustment was made for multiple inferences.

Phase Two Results

This study's main purpose was to develop a model of T-PE practice from a sample of T-PE practitioners. Most of the results are therefore directly related to the modeling process. First, the survey question responses are reported, followed by the relative importance of the variables. Finally, the modeling and the variables offered by participants following the modeling are discussed.

T-PE Questions Results

Survey respondents were asked to indicate how strongly they agreed or disagreed with the eight statements described above. They were provided with these instructions and working definitions:

In answering these questions, please think about how you prefer to practice evaluation. I know that answers to these questions are almost always context dependent, and "it depends" might be your answer choice. But, I would like you to think of your ideal evaluation situation.

The term "stakeholder" is used here to mean anyone, other than the evaluator, with a vested interest in the entity (evaluand) being evaluated.

"Participants" are those stakeholders who take an active role in the evaluation.

"Participation" is any active role and may vary widely in breadth and depth.

As expected, these questions drew generally positive responses from participants. More than two thirds agreed or strongly agreed (78.7%) with the whole set of statements, compared to just over a fifth (21.2%) who disagreed or strongly disagreed with all eight (see Table 16). This is somewhat reflective of previous research with regards to participation. More specifically, the broad support (95.3%) for the statement "I always try to involve stakeholders in my evaluations" echoes findings in prior studies (e.g., Cousins et al., 1992; Fleischer & Christie, 2009; Preskill & Caracelli, 1997).

Table 16. Participant Response to T-PE Statements

Statement	Dimension	Strongly Agree		Somewhat Agree		Somewhat Disagree		Strongly Disagree		Total
		N	%	N	%	N	%	N	%	N
Intended program beneficiaries should participate in carrying out evaluation.	Selection	407	34.3%	533	44.9%	195	16.4%	53	4.5%	1,188
People representing all important perspectives should be involved in any evaluation.	Selection	747	62.9%	359	30.2%	67	5.6%	14	1.2%	1,187
*I always try to involve stakeholders in my evaluations.	Depth	795	67.0%	336	28.3%	43	3.6%	7	0.6%	1,181
*I prefer not to take on an evaluation unless it has a strong participatory component.	Depth	197	16.6%	441	37.2%	431	36.3%	112	9.4%	1,181
*Evaluators should share technical decision-making with stakeholders.	Control	575	48.4%	482	40.6%	115	9.7%	16	1.3%	1,188
Evaluators should help train all legitimate groups to do evaluation.	Control	416	35.0%	532	44.8%	196	16.5%	43	3.6%	1,187
Evaluation should focus on bringing about individual empowerment emancipation or self-determination.	Social Justice	265	22.3%	567	47.8%	284	23.9%	71	6.0%	1,187
Evaluation should focus on bringing about social justice.	Social Justice	272	22.9%	547	46.1%	275	23.2%	93	7.8%	1,187
Total		3,674	38.7%	3,797	40.0%	1,606	16.9%	409	4.3%	9,486

*Agreement with this statement identifies participatory evaluators

Convergent construct validity of these questions was explored using responses to three additional statements that were also included in the survey because they were expected to negatively correlate with the eight T-PE statements. All three negatively correlate with a computed mean of the T-PE items.

Table 17. T-PE Items Correlation to Convergent Items (n=1157)

Statement	<i>r</i>	<i>p</i>
Only key decision-makers should participate in carrying out evaluations.	-.240	<.001
Evaluators should maintain technical decision-making of evaluation projects.	-.259	<.001
I prefer to involve stakeholders in very limited ways.	-.432	<.001

Responses to these statements were expected to negatively correlate because: a) the choice of engaging only key decision-makers in carrying out an evaluation is more aligned with the definition of practical participatory evaluation and reflects a more utilization-focused evaluation stance (Cousins & Whitmore, 1998); b) it was made clear during the statement development process that T-PE evaluators negotiate the divestment of decision-making control as capacity is built; and c) T-PE evaluators generally have a very broad definition of stakeholder involvement and do not limit that scope *a priori*.

The internal reliability of the eight items in Table 16, measured by the coefficient alpha, is moderately strong (Cronbach's $\alpha=.736$) (DeVellis, 2003). When any of the items is removed the internal consistency is reduced to unacceptable levels (Table 18). In particular, when removing either of the two items that address the core philosophical strength of T-PE of involving beneficiaries and bringing about empowerment and emancipation for the beneficiaries, the alpha

drops below .7. Given that this is a relatively short set of items, and their purpose is to identify different groups of practitioners, this is sufficient to consider these internally consistent (Crano & Brewer, 2002; Steiner & Norman, 2008).

Table 18. T-PE Items Coefficient Alpha

Statement	α if item removed
Intended program beneficiaries should participate in carrying out evaluation.	.687
People representing all important perspectives should be involved in any evaluation.	.707
I always try to involve stakeholders in my evaluations.	.715
I prefer not to take on an evaluation unless it has a strong participatory component.	.702
Evaluators should share technical decision-making with stakeholders.	.697
Evaluators should help train all legitimate groups to do evaluation.	.719
Evaluation should focus on bringing about individual empowerment emancipation or self-determination.	.727
Evaluation should focus on bringing about social justice.	.714

Comparing groups aligned with different theoretical orientations (known groups method) may also provide some construct validity by showing how well they identify a unique group of evaluators. If participants agreed with all eight statements, they were labeled T-PE. If they disagreed with two statements regarding social justice and empowerment but agreed with the remaining six, they were labeled P-PE. If they agreed with the three participatory statements and only some of the other five, they were labeled PE. If they did not agree with any one of the three PE statements, they were labeled non-PE.

In the survey population, those who were categorized as P-PE should have chosen a utilization focused evaluation theoretical approach more often than those in the T-PE group. To test this hypothesis, the theoretical orientation selections were dummy coded so as to compare those who selected a particular orientation across groups. Those who selected utilization focused evaluation were coded “1” and those who selected another were coded with “0.” The same

process was undertaken for those who selected participatory evaluation, “I do not have a preferred theoretical orientation,” and “I do not know enough about these to select one.”

There was no significant difference among the four groups (T-PE, P-PE, PE, non-PE) in the number who selected utilization focused evaluation over any other, or among those who selected “I do not know enough about these.” There was a significant difference in the distribution across groupings for people who selected participatory evaluation ($F=34.801$, $df=3$, $MS=3.496$, $p<.001$) and those who selected “I do not have a theoretical orientation” ($F=3.104$, $df=3$, $MS=.361$, $p=.026$) (Table 19).

Table 19. Differences in Theoretical Preference by PE Groupings

	Sum of Squares	df	Mean Square	F	Sig.
Utilization focused	.908	3	.303	1.575	.194
Participatory	10.408	3	3.469	34.801	.000
I do not have one	1.084	3	.361	3.104	.026
I do not know enough	.334	3	.111	1.355	.255

Post hoc analyses on these two significant findings (Table 20), controlling for familywise error rate using Bonferroni methods (Howell, 2002), showed significant differences for only a few comparisons. Of those who selected participatory as their preferred theoretical orientation, significantly more were categorized as T-PE than non-PE (mean diff=.2167, $p<.001$) or PE (mean diff=.1203, $p<.001$). Significantly more of those who selected “I do not have a theoretical orientation” were categorized as non-PE than T-PE (mean diff=.0662, $p=.026$). No differences surfaced between those categorized as P-PE and T-PE.

Table 20. Comparisons Between PE Groupings on Two Theoretical Preferences

Dependent	(I)	(J)	Mean	Std. Error	Sig.	95% CI
-----------	-----	-----	------	------------	------	--------

						Lower Bound	Upper Bound
PE	Non-PE	PE	-.09638*	.02565	.001	-.1642	-.0286
		P-PE	-.12019	.05023	.101	-.2529	.0125
		T-PE	-.21668*	.02140	.000	-.2732	-.1601
	PE	Non-PE	.09638*	.02565	.001	.0286	.1642
		P-PE	-.02381	.05369	1.000	-.1657	.1181
		T-PE	-.12030*	.02859	.000	-.1958	-.0448
	P-PE	Non-PE	.12019	.05023	.101	-.0125	.2529
		PE	.02381	.05369	1.000	-.1181	.1657
		T-PE	-.09649	.05179	.376	-.2334	.0404
	T-PE	Non-PE	.21668*	.02140	.000	.1601	.2732
		PE	.12030*	.02859	.000	.0448	.1958
		P-PE	.09649	.05179	.376	-.0404	.2334
I do not have a theoretical orientation	Non-PE	PE	-.00524	.02772	1.000	-.0785	.0680
		P-PE	.03388	.05428	1.000	-.1096	.1773
		T-PE	.06624*	.02313	.026	.0051	.1274
	PE	Non-PE	.00524	.02772	1.000	-.0680	.0785
		P-PE	.03912	.05801	1.000	-.1142	.1924
		T-PE	.07148	.03089	.125	-.0102	.1531
	P-PE	Non-PE	-.03388	.05428	1.000	-.1773	.1096
		PE	-.03912	.05801	1.000	-.1924	.1142
		T-PE	.03236	.05596	1.000	-.1155	.1802
	T-PE	Non-PE	-.06624*	.02313	.026	-.1274	-.0051
		PE	-.07148	.03089	.125	-.1531	.0102
		P-PE	-.03236	.05596	1.000	-.1802	.1155

*. The mean difference is significant at the 0.05 level.

Many more of the P-PE modelers selected UFE than the rest of the sample. In the survey sample, more than twice as many selected UFE (24.3%) as PE (11.5%). The T-PE modeling sample was more balanced (29.6% UFE vs. 23.9% PE). This balance was not the case for the P-PE modelers; four times as many selected UFE (56.3%) over PE (12.5%). Participants in the webinar were about as balanced as the overall survey population (34.2% UFE vs. 17.5% PE).

Those grouped as P-PE were weighted heavily as preferring a UFE theoretical orientation. Practical participatory evaluation, with its practical focus on engaging stakeholders for the purpose of use, is much more closely aligned with UFE than it is with transformative evaluation. The selections by the P-PE group suggest this grouping was appropriate for those modelers.

Variables Results

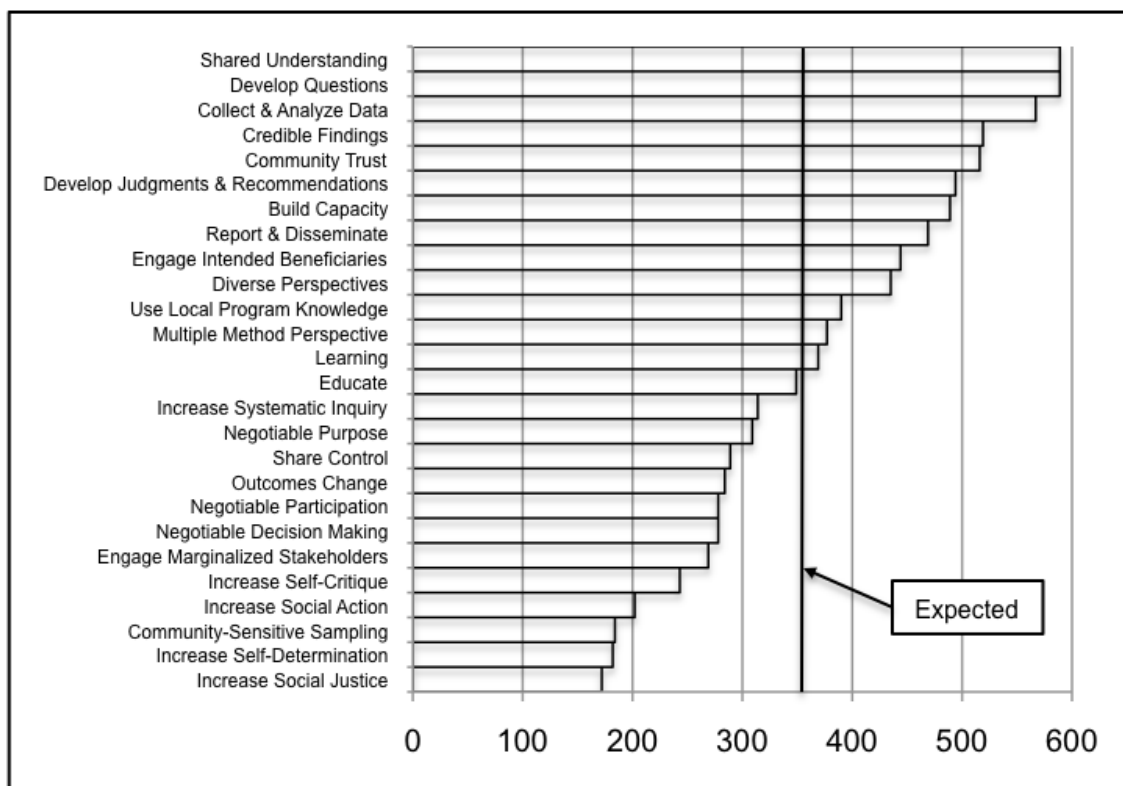
In the modeling process, participants pulled variables from the list on the left side of the screen and included them in the model building area to depict their practice. The inclusion of a variable in a model indicated a participant's endorsement of its importance. By drawing an arrow from one variable to another, the participant implied a relationship. For every arrow there was a beginning variable and an end variable, and each end of an arrow indicated an endorsement of that variable's importance in the participant's practice. Arrows could be drawn to and from as many variables as a participant desired. For example, one could include 20 variables in a model and draw an arrow from "Stakeholder Involvement" to each one and between and among the entire set. Arrows could have a single direction or be bi-directional.

The data obtained from this model-building exercise are the "from" variable and the "to" variable for each arrow drawn. For each of the 27 variables (26 listed, 1 constant), an arrow could be drawn to or from each variable, for a possible 702 arrows. Greater variable usage indicates greater importance and more frequently endorsed relationships between two variables elevate the importance of that theoretical relationship. Together these would ostensibly decide the variables and links for a most-endorsed model.

Variable Endorsement By All Modelers

When all the participants' models were combined, the distribution was not even across all variables ($n=9,600$, $X^2=1,163.51$, $df=25$, $p<.001$). Figure 3 below shows the number of times a variable was endorsed by having an arrow drawn from or to it. All variables had an equal chance of being used in the model. If they had all been used equally, the number of links for each would be 369.3 (the average number of actual uses).

Figure 3. Endorsement of Principles, Activities, and Outcomes

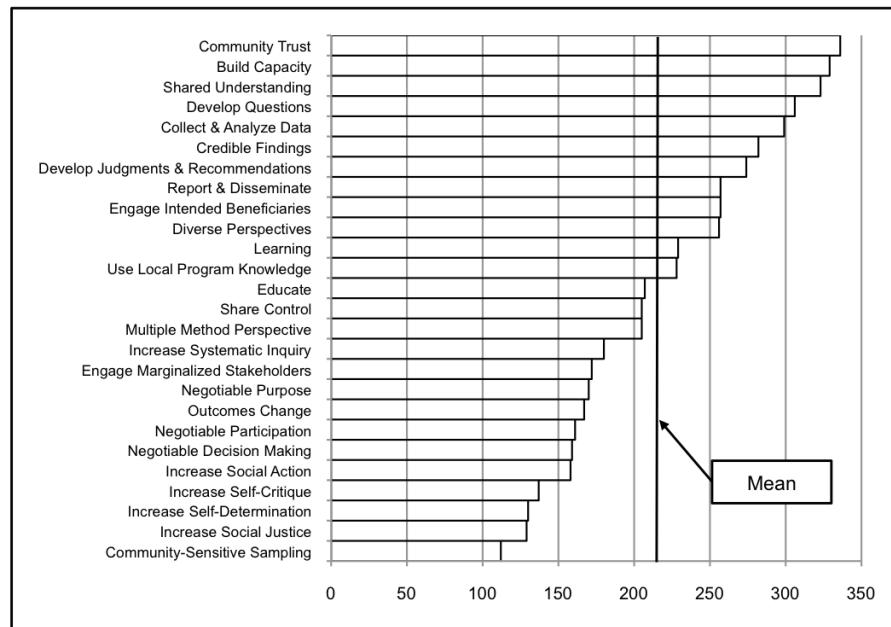


At the high end of usage, Shared Understanding was used as frequently as Develop Questions. At the low end of usage, Increase Social Justice, Increase Self-Determination and Community-Sensitive Sampling were used much less than expected. Mixed Method Perspective, Learning, and Educate were all endorsed around the expected amount.

Variable Endorsement by T-PE Evaluators

The sample of T-PE evaluators also did not endorse the variables evenly ($n=5,668$, $X^2=533.17$, $df=25$, $p<.001$). Figure 4 below shows that the order of endorsement is different for this sample.

Figure 4. T-PE Evaluator Endorsement of Principles, Activities, and Outcomes



Community Trust, Build Capacity, and Shared Understanding were the top 3 variables used by this group. Learning, Use Local Program Knowledge, Educate, Share Control, and Multiple Method Perspective all were endorsed close to the mean number of endorsements (Mean=218). Community-Sensitive Sampling was the lowest endorsed, with Increase Social Justice and Increase Self-Determination also at the bottom of the list.

Relationships Between Variables

The relationship between variables in the model is another important consideration in understanding the modeling data. Participants placed variables into their model and drew arrows

from one to another, and these arrows could be assumed to depict some sort of causal chain. The exact meaning of the linkages was not measured, but if an arrow was drawn from Stakeholder Involvement to Shared Understanding, it could reasonably be interpreted that Stakeholder Involvement is necessary to have Shared Understanding or that Shared Understanding is a result of Stakeholder Involvement.

The number of arrows drawn into or out of each variable also supports our understanding of principles and outcomes. Those variables with more arrows drawn *from* them may be important principles or activities that cause some outcome. Conversely, more arrows drawn *into* a variable support categorizing the variable as an outcome.

Figure 5 summarizes the direction of all the arrows drawn in the 240 models that were produced. Each row represents 100% of the arrows drawn to or from a variable. To the left of each row's center is the percentage of arrows drawn out of the variable and to the right is the percentage of arrows drawn into a variable. A marker to the left side of the center bar means more arrows were drawn into that variable than out (i.e., greater percentage to the right of center), and a marker to the right side of center indicates more arrows were drawn out of the variable than in (i.e., greater percentage to the left of center). Significant differences between the numbers of in arrows vs. out arrows are noted.

Figure 5. All Modelers' Relative Endorsement of Principles, Activities, and Outcomes



* $p < .05$; ** $p < .01$; *** $p < .001$

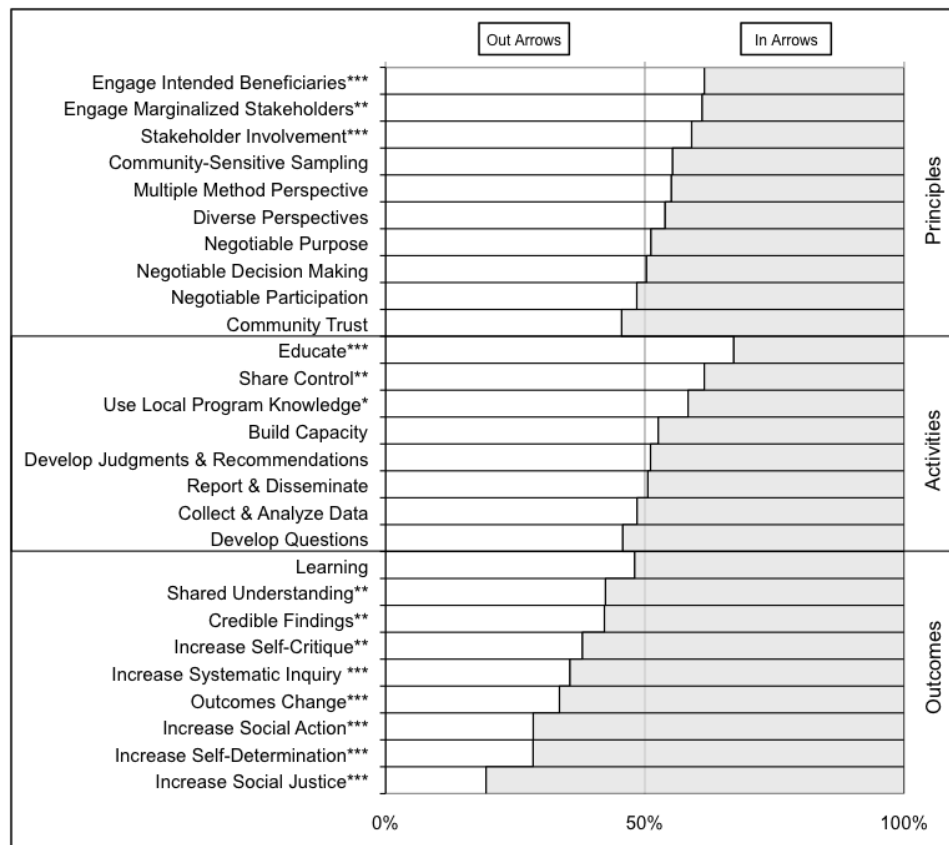
At the top of the graph, all of the variables categorized as *principles* in an earlier phase of this research, with the exception of Community Trust, have more arrows going from them. Negotiable Participation and Negotiable Decision-Making about evenly had arrows into them and arrows going out of them, suggesting they may be more like *activities* than was perceived in the earlier phase of this research. In the middle of the figure, many of the individual *activities* are about equally as likely to have as many arrows going in as going out. There are four that are different: Use Local Program Knowledge, Educate, and Share Control all have more arrows going out of them than in, suggesting they may be more accurately viewed as *principles* rather than activities. Develop Questions exhibits the opposite trend; the fact that more arrows are

going into than out of this variable suggests that it might be more of an *outcome* than an *activity*. Finally, all of the variables that had previously been identified as outcomes were treated as expected by this group, with more arrows going into them than out.

T-PE Evaluators' Relationships Between Variables

Examining this same question for the T-PE evaluators the distribution is quite different (Figure 6). Many of those variables previously identified as principles had about even number of arrows going into than out of them.

Figure 6. T-PE Relative Endorsement of Principles, Activities, and Outcomes



*p<.05; **p<.01; ***p<.001

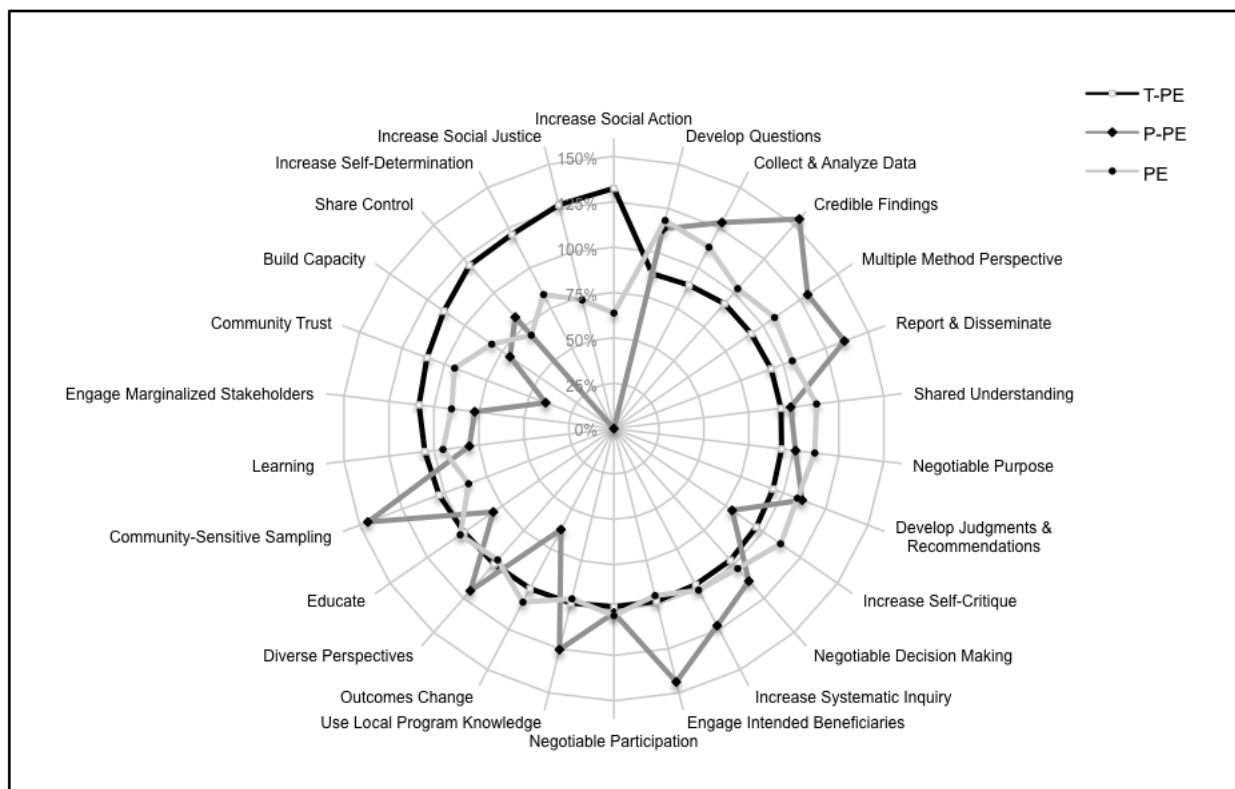
Community-Sensitive Sampling, Multiple Method Perspective, Diverse Perspectives, Negotiable Purpose, Negotiable Decision Making, Negotiable Participation, and Community Trust, previously identified as principles, each had about even number of arrows drawn out of them as into them, indicating they might be as much catalysts as outcomes.

As with the broader sample, Educate, Share Control, and Use Local Program Knowledge, each previously identified as activities, had more arrows going out than in. Other than Learning, all the variables previously identified as outcomes had significantly more arrows drawn into them than out.

Variable Usage by Practitioner Groups

An alternative way of interpreting the modeling data is to examine any differences in endorsement by the groups identified by the eight T-PE statements. Figure 7 presents the distribution of the variables by how close each group's endorsement comes to its *expected* value, represented as a percentage. If each variable was evenly endorsed (i.e., they were linked equally across groups) the points would all be at 100%. This radar chart shows each of the groups and how close to equal their endorsements fall on each variable. It is sorted lowest to highest by the T-PE group's values, beginning at 12 o'clock and moving clockwise.

Figure 7. Principles, Activities, and Outcomes Endorsement Percentage of Expected Value



As would be expected from the philosophical importance of empowerment and social justice, Increase Social Action, Increase Social Justice, Increase Self-Determination, and Share Control, at about the 11 and 12 o'clock points of the radar chart, appear high above the expected value for the T-PE group. The first three of these do not show up in a P-PE model at all, and are therefore at zero in the radar chart. Instead, the P-PE group's endorsement of Credible Findings, Engage Intended Beneficiaries, and Community-Sensitive Sampling are higher than might be expected; Community Trust, Increase Self-Critique, and Outcomes Change are lower than might be expected.

Participant Proffered Variables Results

After participants completed their models, they were asked if there was any one variable and its definition they thought was missing from the list. Of the 240 modeling participants, 134 comments made by 120 modelers were interpretable for coding. The comments were first interpreted for meaning, grouped by these meanings, and the groupings were defined based on the original postings. Topics with fewer than four comments (3% overall) were grouped together as miscellaneous (Table 21).

Table 21. Proffered Additional Variables

Comment Topic	#	% N=134
Program Improvement	11	8.21%
Contextual Consideration	10	7.46%
Use	10	7.46%
Learning Loop	7	5.22%
Stakeholder Engagement	6	4.48%
Stakeholder Support	6	4.48%
Evaluation Capacity	5	3.73%
Process Use	5	3.73%
Program Theory Development	5	3.73%
Evaluation Culture	4	2.99%
Formative Feedback	4	2.99%
Improve Evaluation Practice	4	2.99%
Miscellaneous	57	42.53%

The most frequently mentioned variable related to program improvement as an outcome. For example, participants suggested, “*Positive results/outcome (i.e. the evaluation leads to better program outcomes)*” and “*Improved programming that results from a CQI or systemic loop of data collection, reflection, program management/delivery changes, repeat.*”

The next most frequently mentioned topic related to contextual considerations. Many of these comments referred to issues around the organization, its funding, or its reporting requirements.

These were identified as contextual in nature; 7.5% of the 134 comments could be considered contextual. For example, participants suggested: *“A key principle guiding this work is attention to the adequacy of human and financial resources to ensure that other conditions have been addressed in the work”* and *“Requirements from external accreditation agencies.”*

Use was also a topic that was frequently mentioned. As with issues related to context, about 7.5% of the 134 comments were related to some sort of use. Most often these were more specifically related to how findings were used rather than the process, which were categorized as “process use.” Respondents noted the importance of *“Use of evaluation results for program improvement or program evolution (you have similar items, but this is how I phrase it)”* and *“Evaluation use—increased use of evaluation findings as a result of stakeholder input/participation.”*

Some of these topics are more applicable to participatory evaluation and evaluation practice more generally and do not aid the specificity of a T-PE model. For instance, topics related to program improvement can be embedded in *use* and are a key outcome of all evaluations; Awareness of context is a necessity in any evaluation; and “process use” is considered a sensitizing concept for any evaluation involving stakeholders (Patton, 2002). Other topics, such as those related to engaging stakeholders, were already included in the developed list.

In sum, some suggested variables were valuable improvements to the list produced in the first phase of this research. Others were relevant but reflected participation more generally. In the end, two variables and their definitions are submitted as additions to the developed list:

- Learning Loop (outcome): *Learning generated by the evaluation is incorporated into the organization and improves future stages of the evaluation.*
- Stakeholder Support (principle): *Commitment to participatory evaluation by program leaders and other key stakeholders.*

Model Results

At the completion of the survey, those who had responded favorably to the participatory evaluation statements were invited to model their practice; 240 evaluators did so, and the results included a diverse set of representations of practice. The number of arrows in any given model was extremely varied. All three groups produced models with varied complexity (range: 2–82 links) that did not normally distribute. The distribution was positively skewed for all three groups, with the PE group showing the most skew (skew=1.875). The P-PE group's models were more compact (range: 5–48 links; SD=10.82). Histograms of each subgroup's arrow usage (Figures 8–10) follow Table 22, which summarizes these findings.

Table 22. Arrows in Models by Participant Grouping Categories

Grouping	#	Range	Mean	Median	Mode	S.D.	Skew	Kurtosis
T-PE	3,327	2–59	23.43	21	17	12.28	0.70	0.28
P-PE	365	5–48	22.81	22	16	10.82	0.57	0.60
PE	1,889	7–82	23.04	19	11	14.85	1.875	4.05
Total	11,162							

Figure 8. Histogram of Arrows in T-PE Models (n=140)

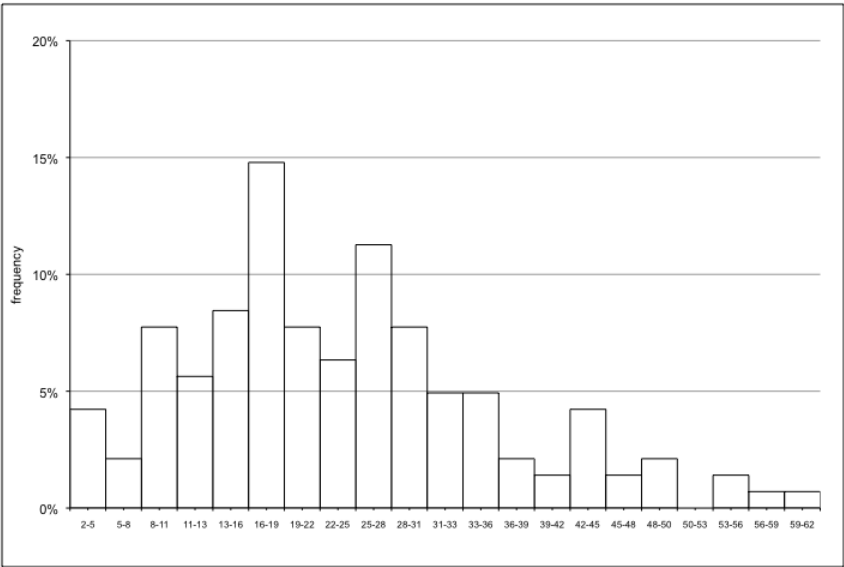


Figure 9. Histogram of Arrows in P-PE Models (n=16)

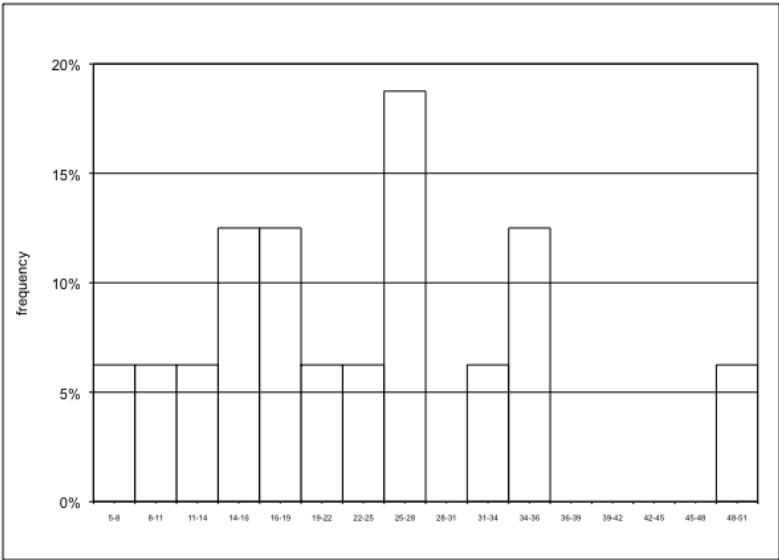
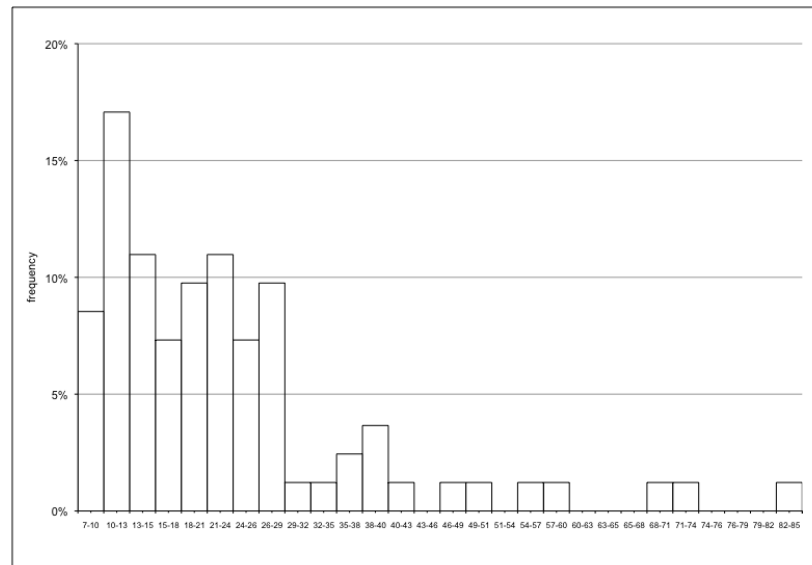


Figure 10. Histogram of Arrows in PE Models (n=82)



T-PE Model

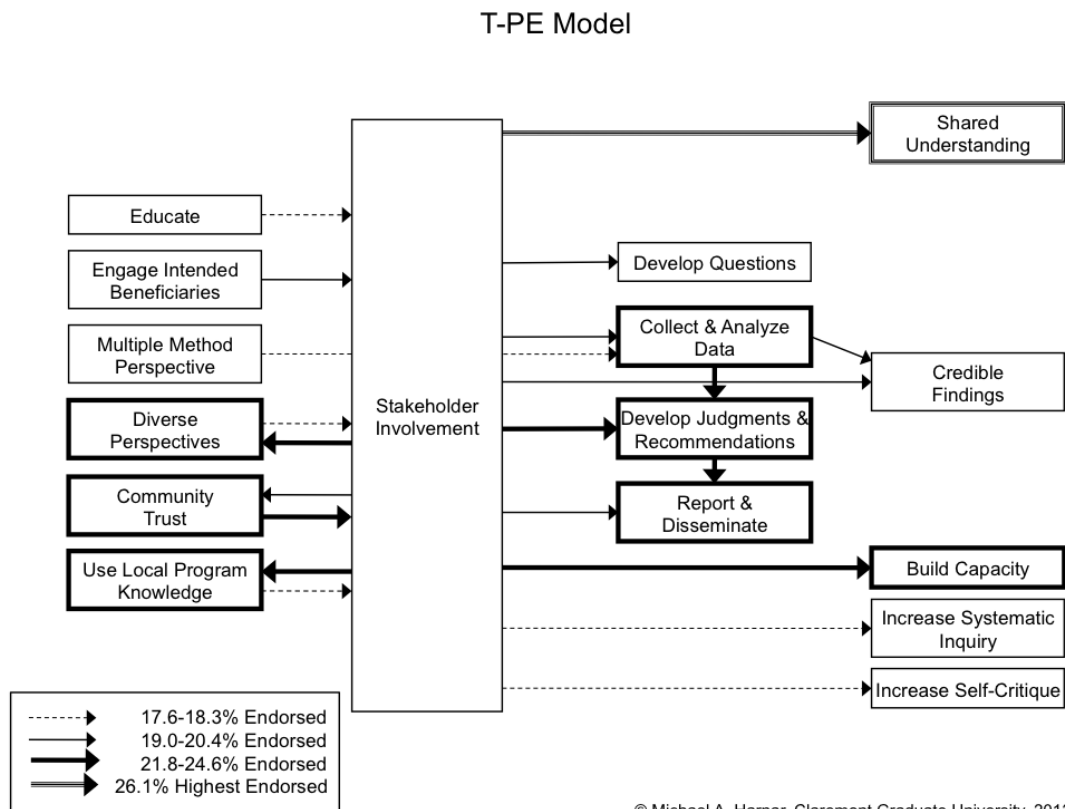
Each participant who created a model dragged each variable into the model and decided what arrow needed to go from that variable to either another variable they already placed into the model or to the constant “Stakeholder Involvement” variable. The participant could choose from any of the 26 variables and could connect them using one- or two-way arrows (e.g., for recursive relationships). A model could therefore be as complex as 702 links or as simple as no links.

Because of the extreme variance in participants’ model complexity, distilling a manageable T-PE model from the input of 142 modelers was not straightforward. Some participants created models with two links, and one used 59 links. A model with 59 links is very difficult to interpret, given the complexity it represents; a model with two links is also difficult to interpret because it lacks detail. The goals of this study are to a) produce a model that is interpretable by the T-PE evaluators who participated in defining its content, so as to have their interpretations be

meaningful; b) create a product that practitioners might find useful; and c) develop a model that provides some discernment from other evaluation theories.

The average number of links in a given model might then be thought of as the average model complexity desired by the participant. In the case of T-PE, the mean was 23 links. The mean number of links, what might be considered a model of *average* complexity, represents only an average and the distribution is flat and positively skewed ($SD=13.11$). In this case, the middle of the distribution (median=21) is a better representation of average complexity. With this in mind, a model with the top 21 endorsed links and the principles attached to those links is presented in Figure 11.

Figure 11. Transformative Participatory Evaluation Model



While the selection of links and variables is defined by their endorsement, their relative placement in the model is more of an art. In this case, their location reflects the categories assigned earlier as well as the arrow distribution. On the left side of the model, more arrows were drawn from the variables than to them; on the right side, more arrows were drawn into the variables than from them.

More than half (n=12) of these links were drawn *from* Stakeholder Involvement, and almost a quarter (n=5) were drawn *to* Stakeholder Involvement. That means 81% of the top 21 links included Stakeholder Involvement. This is not surprising, given that Stakeholder Involvement was the starting variable in everyone's model. Though unlikely, one could have entered variables into the model and not link at all to Stakeholder Involvement.

Because of the variance in model complexity among the study participants, this product represents only a fraction of the T-PE group. The most endorsed link, Stakeholder Involvement to Shared Understanding, was selected by just over a quarter (26.06%) of the T-PE modelers and the least endorsed links in this group, Stakeholder Involvement to Increase Systematic Inquiry and Stakeholder Involvement to Increase Self-Critique, were endorsed by only 17.61% of the 142 T-PE modelers (Table 24)

Table 23. Percent of T-PE Modelers Selecting Top 21 Links

Rank	%	From	To
1	26.06%	Stakeholder Involvement	Shared Understanding
2	24.65%	Stakeholder Involvement	Diverse Perspectives
3	23.24%	Develop Judgments	Report & Disseminate
4	23.24%	Stakeholder Involvement	Build Capacity
5	21.83%	Stakeholder Involvement	Use Local Program Knowledge
6	21.83%	Collect & Analyze Data	Develop Judgments

7	21.13%	Stakeholder Involvement	Develop Judgments
8	21.13%	Community Trust	Stakeholder Involvement
9	20.42%	Stakeholder Involvement	Develop Questions
10	20.42%	Stakeholder Involvement	Community Trust
11	19.72%	Stakeholder Involvement	Credible Findings
12	19.72%	Stakeholder Involvement	Report & Disseminate
13	19.72%	Stakeholder Involvement	Collect & Analyze
14	19.01%	Collect & Analyze Data	Credible Findings
15	19.01%	Engage Intended Beneficiaries	Stakeholder Involvement
16	18.31%	Educate	Stakeholder Involvement
17	18.31%	Diverse Perspectives	Stakeholder Involvement
18	18.31%	Use Local Program Knowledge	Stakeholder Involvement
19	17.61%	Multiple Methods	Collect & Analyze
20	17.61%	Stakeholder Involvement	Increase Systematic Inquiry
21	17.61%	Stakeholder Involvement	Increase Self-Critique

Comparative Analysis

Though a model by those who purport to practice T-PE and who agree with all eight items is of most interest, the validity of their model and variable ranking is strengthened if it is contextualized. Therefore, all respondents who were at all participatory in their preferred practice were asked to create models and these were compared to the T-PE model. For convergence validity support of a T-PE model, some key differences should appear between the rankings and model of a comparison group and those created by T-PE evaluators.

The T-PE model was also compared to other theory modeling research produced to date. Specifically, at the 2010 American Evaluation Association meeting in San Antonio, Texas, students from the University of California, Los Angeles (UCLA) presented their interpretations of four evaluation theories through modeling. They interpreted the writings of key scholars in

each theoretical field and produced visual representations. These models are also compared to the T-PE model below.

Practical Participatory Evaluation Model

A comparison group of non-T-PE evaluators was identified. These respondents endorsed the three PE statements but disagreed with two others on key philosophical underpinnings of T-PE. This yielded a group of 64 participatory evaluators who did not agree that evaluation should “focus on bringing about individual empowerment, emancipation, and self-determination” and that evaluation should “focus on bringing about social justice.” Of these 64 survey participants, 16 created models (25%). This comparison group is labeled P-PE (practical participatory) because of their stronger focus on utilization rather than empowerment and social justice. Their modeling is compared to the T-PE group’s.

A visual P-PE model with 22 links (the median number of P-PE links) would potentially misrepresent a P-PE model because the small sample size did not provide a distribution that allowed for a clear cutoff at 22 links (See Table 25). The link at the 22nd rank is only five links into the “three endorsers” category that is 21 variables deep. Who is to say which five of the 21 links should be in the model. Therefore, a model with 17 links—i.e., the links endorsed by at least 25% of the modelers—would be the most appropriate model. There would be limited validity in any qualitative comparisons between such a model and the T-PE model because it would have fewer links.

Table 24. Ranking of P-PE Modelers Top 38 Links

Ranking	Link Number	# of endorsers	Link Number	# of endorsers	Ranking
1	L1718	7	L0110	3	20

2	L1516	6	L0115	3	21
3	L1015	5	L0119	3	22
4	L1617	5	L0319	3	23
5	L1821	5	L0416	3	24
6	L0114	4	L0506	3	25
7	L0121	4	L0515	3	26
8	L0415	4	L1005	3	27
9	L0421	4	L1415	3	28
10	L0521	4	L1522	3	29
11	L1003	4	L1612	3	30
12	L1601	4	L1812	3	31
13	L1621	4	L1820	3	32
14	L1721	4	L1823	3	33
15	L1801	4	L1910	3	34
16	L1901	4	L1912	3	35
17	L2117	4	L2012	3	36
18	L0105	3	L2118	3	37
19	L0108	3	L2127	3	38

The P-PE model and the T-PE model were compared quantitatively using differential item functioning (DIF) and Rasch modeling. Those links endorsed by 10% or more of the total sample (56 links) were used to create a Rasch model. The output of the equation (Equation 1) is the likelihood that a link will be included in any particular model. Table 26 and Table 27 present the links that the T-PE modelers were more likely to endorse and the links that the T-PE modelers were less likely to endorse, respectively, compared to the P-PE modelers.

Six links were more likely to be endorsed by T-PE modelers than P-PE modelers (Table 25): Diverse Perspectives to Stakeholder Involvement, Educate to Stakeholder Involvement, Stakeholder Involvement to Increase Systematic Inquiry, Stakeholder Involvement to Build Capacity, Engage Marginalized Stakeholders to Stakeholder Involvement, and Develop Questions to Multiple Method Perspective. Five of the links used by T-PE modelers were not

used at all by P-PE modelers: Community Trust to Stakeholder Involvement, Stakeholder Involvement to Community Trust, Stakeholder Involvement to Increase Self-Critique, Increase Social Action to Increase Social Justice, and Build Capacity to Stakeholder Involvement.

Table 25. Variables More Likely to be Endorsed by T-PE Modelers than P-PE Modelers

From	To	T-PE Proportion	P-PE Proportion	Coef.	Standard Error	Z	P
Community Trust	Stakeholder Involvement	0.211	0	*			
Stakeholder Involvement	Community Trust	0.204	0	*			
Stakeholder Involvement	Increase Self-Critique	0.176	0	*			
Increase Social Action	Increase Social Justice	0.148	0	*			
Build Capacity	Stakeholder Involvement	0.148	0	*			
Diverse Perspectives	Stakeholder Involvement	0.183	0.063	0.624	0.220	2.836	0.0046
Educate	Stakeholder Involvement	0.183	0.063	0.624	0.220	2.836	0.0046
Stakeholder Involvement	Increase Systematic Inquiry	0.176	0.063	0.599	0.223	2.686	0.0072
Stakeholder Involvement	Build Capacity	0.232	0.125	0.392	0.200	1.960	0.0500
Engage Marginalized Stakeholders	Stakeholder Involvement	0.162	0.063	0.545	0.230	2.370	0.0178
Develop Questions	Multiple Method Perspective	0.155	0.063	0.517	0.234	2.209	0.0271

The T-PE group was less likely than the P-PE group to endorse nine links: Report & Disseminate to Credible Findings, Develop Questions to Collect & Analyze Data, Develop Judgments & Recommendations to Report & Disseminate, Multiple Method Perspective to Develop Questions, Shared Understanding to Stakeholder Involvement, Collect & Analyze Data to Stakeholder Involvement, Develop Judgments & Recommendations to Credible Findings,

Report & Disseminate to Stakeholder Involvement, and Credible Findings to Outcomes Change.

(See Table 26)

Table 26. Variables less likely to be endorsed by T-PE modelers over P-PE modelers

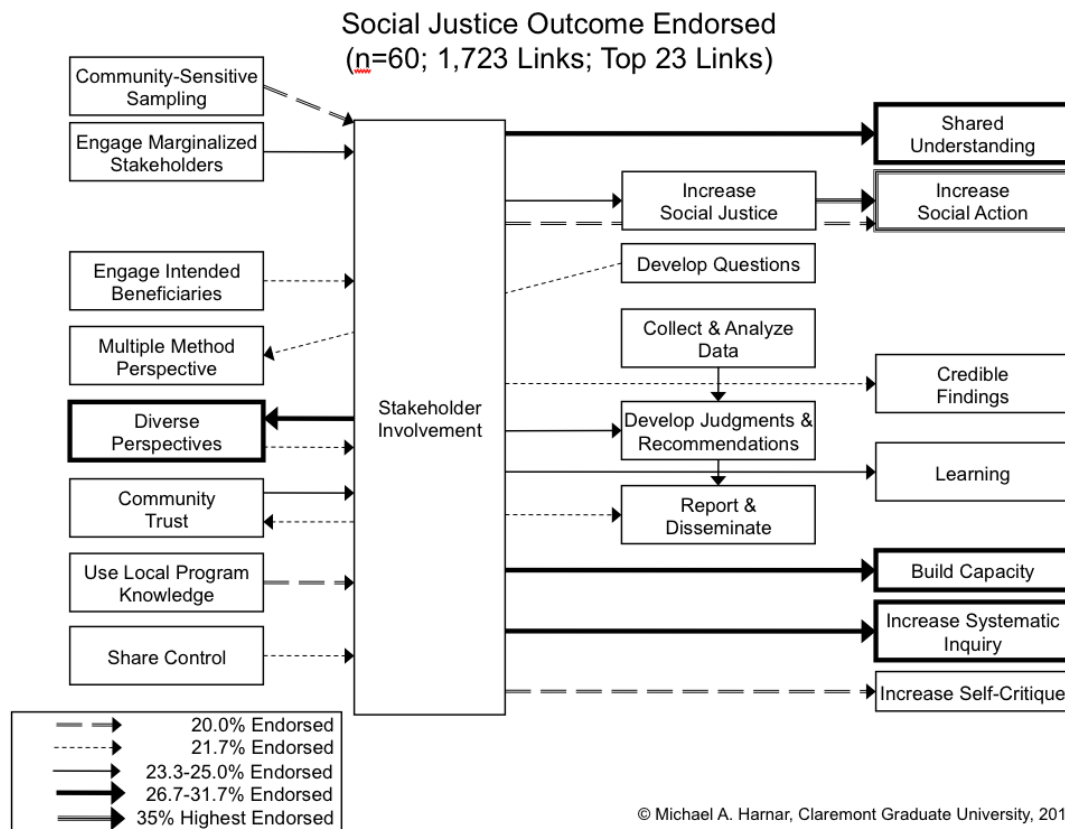
From	To	T-PE Proportion	P-PE Proportion	Coef.	Standard Error	Z	P
Report & Disseminate	Credible Findings	0.099	0.313	-0.772	0.257	-3.004	0.0027
Develop Questions	Collect & Analyze Data	0.162	0.375	-0.623	0.216	-2.88	0.0039
Develop Judgments & Recommendations	Report & Disseminate	0.232	0.438	-0.524	0.193	-2.715	0.0066
Multiple Method Perspective	Develop Questions	0.085	0.250	-0.692	0.274	-2.526	0.0116
Shared Understanding	Stakeholder Involvement	0.099	0.250	-0.603	0.260	-2.319	0.0204
Collect & Analyze Data	Stakeholder Involvement	0.120	0.250	-0.489	0.243	-2.012	0.0442
Develop Judgments & Recommendations	Credible Findings	0.120	0.250	-0.489	0.243	-2.012	0.0442
Report & Disseminate	Stakeholder Involvement	0.120	0.250	-0.489	0.243	-2.012	0.0442
Credible Findings	Outcomes Change	0.070	0.188	-0.597	0.298	-2.003	0.0451

Social Justice Model

The data provided by the survey and modeling participants offers almost infinite possibilities for analysis. One question that arose during this research, however, related to the importance of social justice in T-PE. While all the identified subgroups ranked social justice of low importance and used it sparingly in their models, some still used it as an outcome. This raised the question: “What would a model look like if all the modelers felt social justice was an important outcome?”

There were 60 modelers who used Increase Social Justice as an outcome (i.e., had arrows going into it). Figure 12 represents their combined model.

Figure 12. Social Justice As Outcome Model



Both Increase Social Action and Increase Social Justice now appear in this model. Social Justice is linked *from* Stakeholder Involvement and *to* Increase Social Action. The latter is the most endorsed link for these modelers. A direct line could be drawn from Stakeholder Involvement through Increase Social Justice to Increase Social Action.

Stakeholder Involvement is still strongly linked to Shared Understanding. In this model, the links between Stakeholder Involvement and Build Capacity and Increase Systematic Inquiry rise

into the upper 20% of endorsed links. A link now appears between Develop Questions and Multiple Method Perspective. Community Sensitive Sampling and Learning were missing from the T-PE model, but they do appear here.

The results from the second phase of the research offer a good deal of insight into how evaluators conceptualize T-PE, but it is limited by its quantitative nature. In order to gather qualitative insight from study participants about a T-PE model, a subset was invited to participate in online discussions about the combined, most-endorsed model. The findings from this third and final phase of the study are described in Chapter 5.

CHAPTER 5

PHASE THREE (WEBINARS)

A graphical representation of transformative participatory evaluation was produced in the second phase of the study, but the model's value is somewhat limited because it is a reductionist representation of the work of the 142 T-PE practitioners; in fact, the creation was a "most-endorsed" model that, because of the diversity of models created, represented no more than 26.1% of the T-PE evaluators—the most endorsed link in the model was only selected by 26.1% of the 142 T-PE evaluators. Also, because it was moved from the lived experiences of evaluators to a two-dimensional space, the nuances of each variable were lost. With these issues in mind, the third and final phase of the research was designed to gain a fuller understanding of this created model. A purposeful sample of respondents who had contributed to the model was invited to participate in a series of webinars to discuss it further. Eight webinars were held and 21 evaluators provided their insights. This process and their resulting contributions are discussed here.

Phase Three Methods

Participants

A most-endorsed model of T-PE practice, with moderate yet interpretable complexity, incorporated the top 21 of the 541 links endorsed by respondents in the phase of the study discussed in Chapter 4. This model was contributed to by 131 of the 142 T-PE modelers, and their contributions ranged from 14 links to one link. Practice by evaluators with only one link in this model (representing 4% of the model's links) would have little in common with it; those who endorsed at least 25% of the 21 links may be able to provide more valuable insight. Of the 55 modelers with five or more links (23.8% of 21 links) in this most-endorsed model, 41 agreed

to participate in one of a series of webinars (representing 28.9% of 142 T-PE modelers).

Ultimately, 21 individuals took part. Table 27 provides an overview of participation in each of the webinars.

Table 27. Webinar Dates and Participants

Date	Participants
10/17/11	5
10/18/11	3
10/19/11	4
11/18/11	3
11/21/11	1
11/22/11	2
12/8/11	1
12/9/11	2
Total	21

Table 28 provides descriptive data for the sample of webinar participants. They were all active evaluators, with most (38.1%) conducting seven or more evaluations a year. They were somewhat more experienced than the total group of modelers and the total survey respondents (71.4% conduct four or more a year vs. 59.2% & 54.1% respectively). The majority of them (76.2%) had been practicing six or more years (vs. 70.8% of modelers and 60.3% survey participants). Most (66.7%) said evaluation is their primary professional identity (vs. 65.4% of modelers and 44.6% of survey participants) and reported having an intermediate level of evaluation knowledge and experience (vs. 51.7% of modelers and 43% of survey participants).

They are a well-educated group and more often than all modelers or survey participants had master's degrees (61.9% vs. 45.4% & 43.8%, respectively). At the same time, fewer webinar participants had doctorate degrees (38.1% vs. 49.2% & 44.0%, respectively). The top two fields of expertise for all three groups were education and psychology, though the webinar participants

were more concentrated in these areas (23.8% each) than the modelers (17.5% & 15.0%, respectively) or the survey participants (17.3% & 13.5%, respectively). The next most common degree varied for each group: public policy and social welfare were most common for the webinar participants (9.5% each); among the modelers and overall survey population, evaluation/research methods (9.6% & 8.9%, respectively) and public health (7.9% & 7.8%, respectively) were most common.

When asked if any of the listed theoretical orientations matched their own, most of the webinar participants (42.9%) chose utilization focused, followed by participatory (28.6%). This is in contrast to modelers and survey participants who chose utilization focused (32.1% vs. 24.3% respectively) and participatory (19.2% vs. 11.5% respectively) orientations. With respect to their role in evaluations, webinar participants were about evenly distributed across the categories of external (33.3%), internal (38.1%), and a mix of internal and external (28.6%). This was similar to the modelers (33.8%, 27.1%, & 37.9%) and the survey participants (32.1%, 20.9%, & 34.4%). All the webinar participants (100%) reported conducting program evaluations (vs. 98.8% of modelers, 88.4% of survey participants), and more often than the larger groups said they do performance auditing and reviews (42.9% vs. 36.6% of modelers, 34.6% of survey respondents) and policy evaluations (33.3% vs. 27.9% of modelers, 24.9% of survey participants).

A bit over half of the webinar participants work in health or public health fields (57.1%) and almost half (47.6% each) work with special needs populations and youth development. The top three contexts for modelers were health/public health (45.0%), non-profits (48.3%), and K–12

education (37.9%); for survey participants, non-profits (36.7%), health/public health (35.9%), and K–12 Education (34.2%) were the most common

The webinar participants were generally from college or university settings (42.9%) or nonprofit/foundations (38.1%), and were typically from North America (90.5%). Similarly, modelers were most often from colleges or universities (37.9%), with a large percentage from non-profits (25.8%) and private business/consulting (28.3%), again in North America (85.0%). And survey participants were most commonly from colleges or universities (38.6%), non-profit organizations (25.6%) or private business/consulting (22.7%), once again from North America (75.1%).

Table 28. Webinar Participants' Characteristics

Evaluations Per Year			
1–3	6	28.6%	
4–6	7	33.3%	
7 or more	8	38.1%	
Total	21	100%	
Years in Evaluation			
2–5	5	23.8%	
6–10	9	42.9%	
11–15	3	14.3%	
16–20	2	9.5%	
More than 20	2	9.5%	
Total	21	100%	
Evaluation Identity			
Primary	14	66.7%	
Secondary	7	33.3%	
Total	21	100%	
Evaluation Knowledge and Experience			
At an intermediate level	14	66.7%	
At an advanced level	7	33.3%	
Total	21	100%	
Highest Education Level Completed			
Master's degree	13	61.9%	
Doctoral degree	8	38.1%	

	Total	21	100%
Field of your Highest Degree			
Education		5	23.8%
Psychology		5	23.8%
Evaluation/Research methods		1	4.8%
Public health		1	4.8%
Public policy		2	9.5%
Social welfare		2	9.5%
Anthropology		1	4.8%
Other			
Applied social science		3	14.3%
Interdisciplinary		1	4.8%
	Total	21	100%
Preferred Theoretical Orientation			
Utilization focused		9	42.9%
Participatory evaluation		6	28.6%
Theory driven		2	9.5%
Social justice driven		2	9.5%
I do not have a preferred theoretical orientation		1	4.8%
I do not know enough about these to select one		1	4.8%
	Total	21	100%
Primary Role as an Evaluator			
External		7	33.3%
Internal		8	38.1%
Mix of internal & external		6	28.6%
	Total	21	100%
Primary Type(s) of Evaluations Conducted			
Program evaluations		21	100%
Performance auditing/monitoring/reviewing		9	42.9%
Policy evaluations		7	33.3%
Curricula evaluations		5	23.8%
Evaluation of research		5	23.8%
Student trainee evaluations		3	14.3%
Personnel evaluations		3	14.3%
Consumer evaluations		1	4.8%
	Total	54	
Primary Evaluation Context(s)			
Nonprofits		9	42.9%
Health/Public health		12	57.1%
K–12 education		9	42.9%
Higher education		8	38.1%
Youth development		10	47.6%
Adult education		6	28.6%
Government		8	38.1%
Child care/Early childhood education		6	28.6%

Advocacy and policy change	5	23.8%
Human services	9	42.9%
Public policy/Public administration	6	28.6%
Evaluation methods	2	9.5%
STEM	2	9.5%
Educational technologies	5	23.8%
Special needs populations	10	47.6%
Organizational behavior	2	9.5%
Workforce/Economic development	3	14.3%
Alcohol or Drug Abuse	3	14.3%
Foundations	6	28.6%
Social justice	8	38.1%
International/Cross-cultural	6	28.6%
Environmental programs	1	4.8%
Social work	5	23.8%
Human development	2	9.5%
Indigenous peoples	2	9.5%
Business and industry	2	9.5%
Law/Criminal justice	5	23.8%
Medicine	1	4.8%
Disaster/Emergency management	1	4.8%
Gender rights	1	4.8%
Media	1	4.8%
LGBT	1	4.8%
Total	157	
Primary Setting(s) of Evaluations Conducted		
College or university	9	42.9%
Nonprofit/Foundation organization	8	38.1%
Private business or consulting	1	4.8%
Federal government agency	2	9.5%
State/Provincial government agency	4	19.1%
School system	2	9.5%
Local government agency	2	9.5%
Total	28	
Primary Location(s)		
North America	19	90.5%
Asia	1	4.8%
Africa	2	9.5%
Europe	1	4.8%
South America	1	4.8%
Central America	1	4.8%
Australia/New Zealand	2	9.5%
Total	27	

Procedures

All 41 potential participants were sent an email inviting them to take part in a webinar. Because participation was lower than expected, three rounds of webinars (with three webinars each) were scheduled. Those who participated in the first round did not receive invitations to subsequent rounds. The email invitation asked them to visit a Lime Survey webpage that listed three dates and times; they were asked to select the one they could attend (Appendix M). Another \$200 Amazon.com gift certificate incentive was offered to these participants if they completed a webinar. The webinars were held between October 17, 2011 and December 9, 2011. Just over half of the 41 invitees participated (n=21).

Each webinar was guided by the same set of questions and they were consistent in their design. At least a day before they were scheduled, participants received a copy of the model and were provided with a list of questions to consider (see Appendix N for a copy of the webinar protocol). The webinars were hosted on the Internet using www.GoToWebinar.com to engage practitioners in diverse locations in a live conversation. The service included a recording function and the recordings were sent to www.VerbalInk.com for transcription.

The transcripts were analyzed using standard qualitative procedures that included deductive analysis of how practice was reflected in the model and how the variables and links were described (Miles & Huberman, 1994; Patton, 2002). A sensitizing concept of social justice was added to the analysis (Patton, 2002, 2007) because it was absent from the model, and the researcher was particularly interested in how it was represented in participants' practice. Themes were identified and analyzed to produce overarching themes of model-to-practice congruence. Quotes were subsequently organized by themes, and the themes were described using these

quotations. Thick descriptions, including examples for most themes, are provided in this chapter so that the analysis process and inferences are transparent (Tashakkori & Teddlie, 1998).

Instrument

Conversations during the webinar were guided by a semi-structured protocol (included in Appendix N) intended to identify key similarities between the most-endorsed model and participants' practice. The intention was to understand how the common model explained these evaluators' practice and to determine if any important differences existed.

Webinar Results

The guiding purpose of the webinars was to gain some qualitative perspective on a quantified model of transformative participatory evaluation practice. The conversations were semi-structured with a focus on understanding how participants believed their practice was or was not reflected in the model. They were also asked to identify any unusual or ill-placed variables. In the process, participants described many of the variables in some detail.

The analysis, therefore, took a deductive stance where mentions of specific variables were coded and analyzed for themes. The five variables mentioned more than 10 times across the eight webinars are covered in the discussion that follows. They are covered in decreasing order of mentions. The discussion also addresses issues of social justice, and the role of the evaluator in evaluation practice. Within the analysis of these themes, the model begins to gain deeper meaning and the examples provide evidence of the model's links in practice.

Community Trust

The principle of Community Trust is located on the left side of the model. In the first phase of this research, the expert panelists defined Community Trust as *Evaluator works to build trust by developing working relationships with a broad range of stakeholders*. The participants in the webinars described this principle generally across three categories: 1) as an outcome of stakeholder involvement (36%); 2) by its importance to the evaluation process (26%); and 3) as an ongoing part of the evaluation cycle, from credible findings to further stakeholder involvement and into future evaluations (19%). Another six comments did not fit into any of the above categories and did not coalesce around any particular theme.

One of the participants described community trust as outcome of stakeholder engagement in this way:

“If the stakeholders are involved in determination of the questions and collaborative determination of the types and means of data collected and [they are] looking at the raw data with the evaluator to develop their own judgments and their own recommendations, stakeholders are part of the reporting and dissemination of the results. The fear and anger are transformed into common concern and responsible – response for intervention. The evaluation becomes part of the story not the end of the story.”

Slightly more than one fourth of the comments (26%) referred to the importance and vulnerability of community trust in gaining broad stakeholder buy-in and access to important data. For example, one person described the context of their evaluations and the centrality of trust to these endeavors:

“...the issues we took on were really edgy. We worked in the area of sweatshops. We worked on teenage pregnancy. We worked on women in prostitution. We worked with grandparents raising grandchildren. We were working on Medicare benefits. It was—there were a lot of issues—I would say 90 percent of the issues that we took on were ones where there was a whole lot of potential for nothing to happen as a result of the research if it wasn’t done with a great deal of

involvement on the part of both the community and the government. So all of these were critically important.”

About one in five (20%) of the comments had to do with community trust not as a principle but as an outcome of credible findings, and its role as an integral component of the cyclical nature of longitudinal evaluation. This quote exemplifies the category of time and cycle:

“I think especially the community trust and use of program knowledge, like I really see community trust coming out of having stakeholders involved throughout the process, especially with regard to developing recommendations, report and dissemination time, so— And that credible findings leads back to more community trust, that it’s less a relationship of...researchers taking from the community, interpreting, making recommendations that they really think are useful, and more [a] relationship of kind of involving affected persons in making recommendations, almost like a circular path.”

Shared Understanding

The link between Stakeholder Involvement and Shared Understanding was the highest endorsed in the T-PE model. More than a third (37%) of the T-PE modelers included this link in their model. After Stakeholder Involvement and Community Trust, it was the most discussed topic in the webinars, with 29 comments. This was certainly by design, as the researcher was particularly interested in understanding this most frequently endorsed link.

The definition of Shared Understanding developed in the first phase of this research is *All participants develop shared understanding of program functions and processes*. This definition was challenged somewhat by those who participated in the webinars. While a majority of comments (31%) did reflect a shared understanding of the program and its processes, others (14%) preferred a broader shared understanding that, in one case, meant “who they are and their role” in the process. Still others (7%) focused on the effect that stakeholder involvement had on gaining a shared understanding and the idea that active participation, rather than passive “sitting

at the table,” produced a shared understanding of the program, the evaluation, and the decisions that needed to be made. For example:

“It wasn’t just because we had the Department of Labor and the unions and the religious groups and whatever at the table to talk about ‘what are we gonna do about sweatshops’ that we had a shared understanding. It was because all of them were at the table and we all talked about developing questions. We talked about methodology. We talked about all these things. We analyzed the data together. We worked on implementing changes together in the systems and through that we came to a shared understanding.”

Another interesting note about shared understanding entered the conversation in three different webinars. That is, a shared understanding was necessary for findings to be credible. One person put it quite simply:

“I think credible findings are perhaps not as valued to us as is shared understandings. Credible findings are absolutely context-based. True credible findings are shared understandings.”

The other facet of shared understanding, mentioned in a few comments, is that of context (n=3). Comments in this category included the difficulty of gaining shared understanding when a program is large and/or multi-site, or when sites are spread across the state or nation.

Credible Findings

On the far right side of the model, Credible Findings is linked most directly to the center activities column. When asked to discuss the outcomes in the model and how they reflect their practice, participants talked about the links between credible findings, stakeholder involvement, and shared understandings.

About a third of the comments (24%) intimated that if stakeholders were deeply involved in the evaluation, the findings were more likely to be seen as credible. Others linked findings’

credibility to whether or not the users would see them as credible (n=5). As one participant explained:

“I would think that it would be more something that resonates as, ‘Yes, that’s correct,’ not ‘Yes, that was done rigorously, [with] rigorous research methods.’ [They would think] ‘That’s the correct interpretation of the work that you did and that’s what makes sense in my cultural context, in my personal context.’ I think it’s a different...it has to be community-related, culturally-related, culturally believable rather than having been done in an academically rigorous manner, because that’s where the reviewers are looking at your methods as much as your findings. Whereas the community members are really looking at your findings to determine if those are credible or did you make them up.”

Another quarter (n=5) of the comments were related to the evaluator’s limited ability to interpret findings without the participation of stakeholders. Evaluators’ interpretations were only credible if those most effected by the findings informed the interpretation.

“...They definitely validated our findings, so I think that gets at the credible findings aspect. So they’ve had years of experience in the field working with young women so they could validate or... It was kind of a way of triangulating data with their experiences and if there was anything that conflicted with their practices that they’ve conducted, then we talked about it.”

Build Capacity

Of the 13 comments related to capacity building, most (69%) were related to the skills necessary to participate in an evaluation or the skills that would be required to further improve a program. Two comments linked capacity building to social justice and the notion that the evaluator is improving the lives of those involved by providing skills to stakeholders. Another two described capacity building as an outcome of participation in general (more specifically, going through all the steps of an evaluation). One participant described the value of building capacity of stakeholders in the data analysis phase, so that they can provide input throughout the evaluation:

“I do a little bit of preliminary kinda cleaning and analysis and then typically set up either webinars or conference calls with our local program coordinators to talk through the preliminary findings. So a lot of this is—involves not only educating the coordinators about the value of the evaluation, but also just a little bit of basic statistics and that type of thing, so some capacity building as well. And then I really ask them to comment on the findings based on their individual experiences with their participants. They certainly have all the one-on-one and qualitative interaction with them, and so hearing from them really helps me—well, helps us both tease out some of the findings that we see.”

Participants mentioned the principles of Diverse Perspectives a few times but their input did not add to the definitions produced in the first phase of the research. Nonetheless, one participant offered an example of how diverse perspectives help inform the model, pointing to the “different systems” that need to coordinate, from primary care clinicians, to educational practitioners, to mental health providers in nonprofits, to state-level stakeholders such as the Department of Education and the Department of Human Services. This participant continued:

“We really, in this project, have people coming from very different theoretical backgrounds but all working towards this common goal, that they want to get kids services and support as early as possible. Everybody has that goal, but how to go about doing that really requires a shared understanding of how the system works, how families move through the system, and how that coordination can be improved. And what’s interesting is that these evaluators who’ve kind of been able to be a link for these different systems—because we get to ask everyone questions, we get to talk to everybody, and then kind of tell all of these different system stakeholders what we’ve heard from everyone—so we’ve been in a really unique position where we get to hear from everyone and tell them all what we’ve heard. It really has moved this system of coordination forward. They’re making practical changes in how they work with these children and families at all of these systems’ levels.”

An important take-away from this comment is the notion of shared understanding.

Throughout the webinars, discussions around diverse perspectives consistently linked them to shared understanding.

Engaging Intended Beneficiaries

Though engaging intended beneficiaries is a key concept of T-PE (one of the eight statements addresses this principle) it was only mentioned five times across the eight webinars. In discussing involving intended beneficiaries, all but one participant mentioned it as something they do. Efforts toward this end were for the reasons aligned with participatory evaluation in general; the engagement grounds their data in the lived experiences of the people they are serving and leads to program improvement.

Social Justice

Because the outcome Increase Social Justice/Enhance Social Justice did not make it into the top 21 endorsed links, its absence in the model was explicitly raised as a topic in each webinar. Of the 25 comments on the topic, most (44%) described it as an overarching principle, or as undergirding the whole model, and not necessarily as an “outcome.” One person described social justice as the universe within which the T-PE model fit:

“...the populations in which you would use this type of a model...and the issues that you would be studying or addressing with this kind of a model... It’s sort of like the world that this model is housed in. ...It would almost be like this model would be inside of a circle and the whole circle is social justice. Something like that.”

Another eight (32%) comments were related to how a program or evaluation’s context might influence how social justice is incorporated. For instance, one person with an advocacy preference articulated the importance of engaging stakeholders unencumbered by an advocacy position:

“I’ve found that depending on the needs, values, and goals of the program evaluated, exposition of social justice issues may or may not increase community trust in the evaluation....I think that if I entered an evaluation with the desire to advocate, the opportunity for the stakeholders and me to learn and see and discover more may be reduced. I have been sobbing more than once at my

computer as results have revealed such gaps and oppression, but I have found that if I bring the results back in a manner that cannot be ignored or dismissed easily, they actually become much more motivating than if I were to pronounce an interpretation.”

Six (24%) comments discussed social justice in terms of how it might be achieved. For instance, two people talked about empowering partners who have advocacy missions (e.g., proactively providing data) and the others discussed their program outcomes or involving stakeholders as effecting social justice.

Impact of/on Evaluators

Though only eight comments touched on this topic, because evaluation is a very context-dependent practice, and participatory evaluation might be considered a “high touch” field, it is important to consider the effects of the evaluator on the evaluation, as well as the effects of the evaluation on the evaluator.

This theme produced two categories of comments, those related to what the evaluator learns from the process (75%), and the idea that the evaluator becomes part of the community and in turn becomes a stakeholder in the evaluation (25%). This comment sums up the first, more common of these two themes:

“When you get a program you don’t know a lot about it, and you can read the documents, but it’s not the same as you are when you’re involved in a process with the stakeholders and you learn from them what really matters in the program to them. And then you just get educated for the program. ...In my practice what I really find out sometimes is you...sit with the stakeholders and try to start defining the goals of the project. For a lot of them, even for them, their goals are not really clear, so then through this process you both get more educated about the program and what the program wants to achieve.”

Summary of Webinar Comments

The webinar conversations begin to add life to what is otherwise a static model. Participants' comments shed light on some of the more nuanced aspects of the model, as they interpreted various connections between variables. For example, community trust is both an outcome and an input in the cycle of evaluation. It is critical to engaging stakeholders and to gaining entry into communities that might otherwise be skeptical of evaluation. Community in T-PE includes the evaluator and this inclusion is key to the building of trust.

Building capacity for stakeholders to participate in the evaluation process is important to developing a shared understanding. It is not just that they are sitting at the table, but sitting at the table means they are engaged deeply in the evaluation process so that they understand the methodological choices and can help interpret the findings.

A shared understanding is vital to the credibility of findings. When stakeholders are involved and understand the process, they are more likely to see the findings as credible. The credibility of findings also builds on community trust in a cyclical way so that the more they are provided with credible findings, the more trust is built into the process.

Social justice, while an important tenet of transformative evaluation (Mertens, 2009), is not an easy principle or outcome to map. Participants understood social justice to be the core reason for their work and not a specific outcome they could put into a model. Those who did talk about social justice described it as an underlying value of the process and not bounded by any practice model.

The webinars were undertaken to acquire more grounded understanding of the two-dimensional T-PE model produced in the prior phase. Insights from practitioners regarding

community trust, shared understanding, social justice, and credible findings add complexity to the model and deepen the model's interpretation in the final chapters of this dissertation.

CHAPTER 6

RESULTS SUMMARY

To aid the reader in interpreting the findings, this results summary covers each phase of the study in brief. This will prepare the reader for the discussion in Chapter 7.

Phase One Results

The first phase of the research was undertaken to identify the key variables expected to be important to transformative participatory evaluators and to develop a set of statements that would distinguish T-PE evaluators from others.

With the help of an expert panel of evaluation theorists and using an online wiki, a list of 26 principles, activities, and outcomes considered key to T-PE practice was developed (Appendix C, Table 29). These variables were edited based on input from the pilot of the survey in phase two, and they were then used in the survey and modeling activities. While the preliminary set of variables presented to the panel was purposefully large, the final set was kept small to ease the cognitive burden in the modeling process.

Table 29. Final 26 Transformative Participatory Evaluation Variables

Principle
Community Trust
Negotiable Purpose
Multiple Method Perspective
Diverse Perspectives
Negotiable Decision Making
Negotiable Participation
Community Sensitive Sampling
Engage Marginalized Stakeholders
Engage Intended Beneficiaries
Activity
Build Capacity
Share Control

Educate
Use Local Program Knowledge
Develop Questions
Collect & Analyze Data
Develop Judgments & Recommendations
Report & Disseminate
<hr/> Outcome <hr/>
Shared Understanding
Learning
Credible Findings
Increase Systematic Inquiry
Increase Self-Critique
Increase Self-Determination
Increase Social Justice
Increase Social Action
Outcomes Change <hr/>

Also during this phase, again using the online wiki, eight statements were developed to help filter T-PE evaluators from others. These statements address the four key dimensions of philosophical perspective on participation, selection of participant stakeholders, depth of their involvement, and control over technical decision-making. The statements were also edited based on input gathered during the survey pilot in phase two, and the final set is provided in Appendix H and Table 30.

Table 30. Transformative Participatory Evaluation Identifying Statements

Statement
Intended program beneficiaries should participate in carrying out evaluation.
People representing all important perspectives should be involved in any evaluation.
I always try to involve stakeholders in my evaluations.
I prefer not to take on an evaluation unless it has a strong participatory component.
Evaluators should share technical decision-making with stakeholders.
Evaluators should help train all legitimate groups to do evaluation.
Evaluation should focus on bringing about individual empowerment emancipation or self-determination.
Evaluation should focus on bringing about social justice.

Phase Two Results

The second phase of the study involved two major aspects: a survey of the American Evaluation Association membership to identify T-PE evaluators and a modeling activity where a participatory subsample of the AEA membership was invited to graphically illustrate both how their practice “ensures stakeholder participation” and what outcomes they intend.

The survey had an 18.5% response rate, with 1,228 AEA members participating. Most (78.7%) somewhat or strongly agreed with all eight statements, and about a fifth (21.2%) somewhat or strongly disagreed with them. Participants who agreed with the three participatory statements that were embedded within the set of eight were invited to the next step in the research, where they were asked to model their practice. Just under half of the full sample (42%, n=563) agreed with all three statements and just under half of those (43%, n=240) completed models.

The eight statements crafted in the phase one of the study were used to group modelers into three participatory evaluation categories. Transformative participatory evaluators were those who somewhat or strongly agreed with all of the eight statements (n=142). Those who agreed with three key participatory statements but disagreed with two statements about evaluation’s mandate to increase social justice and affect individual empowerment and emancipation were considered part of a comparison group and labeled practical participatory (P-PE) (n=16). Those who responded positively to the three participatory statements but did not fit in the above two categories were labeled participatory evaluators (PE) (n=82). (See Table 31.)

Table 31. Modelers' Participatory Evaluation Categories

Grouping	n	%
T-PE	142	59.2%
P-PE	16	6.7%
PE	82	34.2%
Total	240	

In the modeling stage, participants were provided an online white space and a list of the 26 variables. They were asked to drag relevant variables into the space and draw arrows between them to depict their practice. Endorsement of variables (i.e., how many times they were placed in a model and arrows were drawn to them) was examined as a relative value measure.

Endorsement by T-PE evaluators was uneven across the variables, with Community Trust being the most endorsed outcome, followed by Build Capacity and Shared Understanding. T-PE modelers least often endorsed Community-Sensitive Sampling and the outcomes of Increase Social Justice and Increase Self-Determination (Figure 4).

The categorization schema of principles, activities, and outcomes developed in the first phase was also examined by looking at the relative number of arrows drawn to and from variables by both the overall sample and the T-PE evaluators. It might be expected that principles would have more arrows drawn from them than to them, outcomes would have more arrows drawn to them than from them, and the activities might have an equal number of arrows going in and out of them. This schema generally held for the outcomes, but not the principles and activities. While this schema may have been helpful in introducing the variables to participants, the modeling data do not support it very well.

The modeling process provided a most-endorsed model of transformative participatory evaluation (Figure 11) that laid out the most important causal links in practice. The top 21 most endorsed links were used to display T-PE practice. Stakeholder Involvement linked to Shared Understanding was the most endorsed link, with 26.1% of the 142 modelers endorsing it. Most (81%) of the top 21 links in the T-PE model were connected to Stakeholder Involvement.

Though the PE groupings described above did not provide an adequate sample size to create a comparable qualitative non-T-PE model, endorsements of links provided adequate data for quantitative comparisons through differential item functioning analyses. Of those links endorsed by 10% or more of all modelers, 20 were significantly more or less likely to be endorsed by one or another group. Compared to P-PE modelers, eleven links were more likely to be endorsed, and nine were less likely to be endorsed by T-PE modelers (Table 32). Stakeholder Involvement was represented more heavily in the links that were more likely to be endorsed by T-PE modelers. Activities, such as Develop Questions, Develop Judgments & Recommendations, and Report & Disseminate, were less likely to be endorsed by T-PE than P-PE modelers. None of the 16 P-PE modelers included Increase Social Justice, Increase Self-Determination, or Increase Social Action in their models.

Table 32. Variables More or Less Likely to be Endorsed by T-PE Modelers than P-PE Modelers

More or Less Likely	From	To
More Likely	Community Trust	Stakeholder Involvement
	Stakeholder Involvement	Community Trust
	Stakeholder Involvement	Increase Self-Critique
	Increase Social Action	Increase Social Justice
	Build Capacity	Stakeholder Involvement

More or Less Likely	From	To
	Diverse Perspectives	Stakeholder Involvement
	Educate	Stakeholder Involvement
	Stakeholder Involvement	Increase Systematic Inquiry
	Stakeholder Involvement	Build Capacity
	Engage Marginalized Stakeholders	Stakeholder Involvement
	Develop Questions	Multiple Method Perspective
	Report & Disseminate	Credible Findings
	Develop Questions	Collect & Analyze Data
	Develop Judgments & Recommendations	Report & Disseminate
	Multiple Method Perspective	Develop Questions
Less Likely	Shared Understanding	Stakeholder Involvement
	Collect & Analyze Data	Stakeholder Involvement
	Develop Judgments & Recommendations	Credible Findings
	Report & Disseminate	Stakeholder Involvement
	Credible Findings	Outcomes Change

Though some modelers used Increase Social Justice as an outcome, it was absent from the T-PE model created by this research. In response, an exploratory model was created with the top links of those who used Social Justice as an outcome (n=60). This model included Social Action as an outcome linked from Social Justice. The link from Stakeholder Involvement to Shared Understanding was one of the most highly endorsed for this group as well.

After they finished modeling their practice, participants were asked if the list of 26 variables was missing any important variable. Modelers described a Stakeholder Support principle where T-PE evaluators need the commitment of key program leaders to stakeholder participation and to implement any changes the evaluation finds. A meta-evaluation of two empowerment evaluations showed that successful implementation of the Getting To Outcomes (Chinman et al., 2004, 2005) model required a commitment to expending resources on evaluation (Whitmore, Guit, Mertens, et al., 2006). This included the time not only to participate in the evaluation but

also the time needed to learn the necessary skills to implement the findings. Without support, participation (and use) is unlikely. Learning Loop, defined as *Learning achieved during the evaluation is incorporated into the experience of participants so that they use that learning going forward*, is an important consideration when undertaking transformative participation and was also offered. This scaffolds learning and increases the potential of change at the individual, program, and project level.

Phase Three Results

This final phase of the study was undertaken to gain some qualitative understanding of the model produced in the previous phase. The T-PE model created from the 21 most endorsed links reflected the contributions of only just over half ($n=75$) of the modelers. Only 41 of those had at least five links in the model and had agreed to participate in a webinar. All 41 were invited and, in the end, 21 participated.

Webinars provided qualitative insights into the model. For some, this model was a perfect reflection of their practice; for others it was too ideal to fit their real-world experiences. Webinar participants described Shared Understanding as a lynchpin to participation. This shared understanding had to encompass more than just the program and its processes (which support the evaluator); it also had to include the roles of the evaluator and stakeholders in the process.

Community Trust was also highly endorsed and was described by webinar participants as the bedrock to participation. Because participatory evaluation involves the whole community surrounding a program, the ability to collect data from hard-to-access populations is made even more difficult if trust is not present. The evaluator is expected to become part of the community as part of the process of building trust. Moreover, Community Trust is part of the cyclical nature

of evaluation where it builds during the evaluation and adds to stakeholder involvement in subsequent evaluations.

Credible Findings were closely linked to Stakeholder Involvement and Shared Understanding. Specifically, the credibility of findings was increased through the interpretations of stakeholders, especially beneficiaries. This type of shared understanding increased community trust.

Given the presence of two filtering statements on social justice and empowerment in phase one, the absence of these concepts in the T-PE model was counterintuitive. When explaining the absence of social justice in the model, some webinar participants noted it was a difficult outcome to enter into a model, describing it as more of an overarching principle that describes the context within which they work.

CHAPTER 7

DISCUSSION OF THE FINDINGS

There are currently two interesting trends in the evaluation field: Most evaluation practitioners value participation by stakeholders in the process (Fleischer & Christie, 2009; Preskill & Caracelli, 1997) while they are also expected to heed calls for more experimental designs in their work (Gersten & Hitchcock, 2009). There is no inherent incompatibility between randomized controlled trials (RCTs) and participatory evaluation, but the values underlying the two are quite different. Whereas an RCT intends to establish replicable causation through objectivity, participatory evaluation values learning and is apt to be used in situations where a program is more interested in involvement than proving irrefutable causation (Cousins & Laithwood, 1986). Given the ongoing debate that pervades the evaluation literature over what constitutes credible findings (e.g., Donaldson, Christie, & Mark, 2009), it is critical for practitioners to have empirical support for a variety of evaluation approaches so as to better frame their own practice in the evaluation constellation (Mark, 2008).

The contextually driven, applied science nature of evaluation practice means that theory develops in a cyclical process common to most applied disciplines. Observations in the field are synthesized into theories that are then tested in the field and classified into categorization schemas. These schemas provide for a common language and lay the groundwork for more observations to substantiate them.

Participatory evaluation has been classified as having at least two forms. One with an intention towards usefulness and practicality and the other more inclined towards social justice and empowerment. The former benefits from the extensive use and utilization-focused evaluation

research; the latter may be actively researched under different monikers outside North America, but has decidedly less research focus in the North American evaluation literature. Individual observations (e.g., case examples of evaluations) provide some understanding of practice, and efforts are afoot to synthesize these (e.g., Cousins & Chuiorad, in press). Beyond this, empirical research is absent. Practitioners working with a transformative approach may look to Mertens (e.g., 2009) for philosophical guidance. But for the nuances of participatory practice with a transformative approach, there is little guidance.

The evaluation discipline is still a nascent field with few well-described theories (Alkin, 1991, 2004a; Smith, 1979, 1993). And yet evaluation theory is “what we do” (Shadish, 1998). Among the calls for more research on evaluation theories and practice, Mark’s (2008) taxonomy for the study of evaluation has advanced *evaluation context*, *activities*, *consequences*, and *professional issues* as a viable framework. Drawing from this taxonomy, the current study has attempted to further describe evaluation activities through a classification study of participatory evaluation theory and practice.

This research adds to the empirical knowledge of participatory evaluation by focusing on one type, transformative participatory evaluation (Cousins & Whitmore, 1998). To that end, answers to the following research questions were sought: What are the key principles, activities, and outcomes theorized to be key to T-PE? What are the key variables to practitioners? What does a model of T-PE practice look like? And how does this model compare to other evaluation theories?

Developing a model of T-PE involved three phases. In the first, an expert panel helped develop a set of variables key to T-PE and a set of statements that would identify T-PE

evaluators from other evaluators. In the next phase, a survey identified participatory evaluators and invited them to model their practice using innovative online modeling software. A most-endorsed model was created and in the third and final phase of the research, those whose model was most like the created model were invited to participate in webinars to discuss it.

Interpretation of Findings

Research Question 1: What are the key variables of T-PE?

To answer the first research question, a panel of evaluation theorists collaborated with the researcher to develop what became a list of 26 variables central to the practice of T-PE. The researcher and the theorists agreed that this was probably an incomplete list, but was nevertheless “a good place to start.” Given this recognition of incompleteness, an additional step was added to the second phase to garner additional variables from those who had just completed the modeling. From their input, Stakeholder Support was added to the list of principles and Learning Loop was added to the list of outcomes (Appendix O).

Research Questions 2 and 3: What are the most important variables in evaluators’ practice and how do they model it? How does evaluators’ practice relate to the model?

The second and third questions of this research are intertwined. Question two asks what T-PE evaluators see as the most important variables in their practice and how they model their practice. The third question asks how they explain their practice in relation to the model. For these questions, a sample of T-PE evaluators needed to be identified. With this in mind, the initial panel of theorists engaged in developing filtering statements in the first phase. Eight core statements emerged, and these were used in the survey in the second phase to filter T-PE evaluators from other participatory evaluators.

The statements have acceptable internal reliability but weak construct validity. Though the evaluation theorists who created them are very familiar with this form of evaluation, the discrimination strength the statements provided to distinguish T-PE evaluators and other subgroups of PE evaluators is questionable. The importance of this identification schema cannot be understated. The T-PE model developed by this research is only as valid as the grouping mechanism that identified these practitioners from others.

Comparing the preferred theoretical orientation of the subgroups provided some support. Many more of the P-PE modelers selected UFE than was the case for the rest of the sample. With its strong emphasis on the practical use of participation, practical participatory evaluation is much more closely aligned with UFE than it is with transformative evaluation. This greater percentage of UFE-oriented modelers in the P-PE group lends some support to the questions' construct validity.

More support is provided by the quantitative comparison of the resultant T-PE and P-PE models. Through Rasch modeling and differential item functioning analyses, links endorsements were compared and 20 links showed significant endorsement differences between the groups. The T-PE group's links were more stakeholder- and community trust-based and the P-PE group's were more activity- and outcome-focused. This is congruent with the underlying philosophy of engagement of diverse perspectives of T-PE and therefore supports both the validity of the models and the statements' ability to discern T-PE evaluators from P-PE evaluators.

An additional way of considering the validity of this grouping mechanism is to compare participants' responses to the practice-oriented survey questions. The P-PE evaluators are more

active (43.8% working on 7 or more evaluations per year vs. 33.1% T-PE and 31.7% PE); report being more experienced (43.8% having 16 or more years in evaluation vs. 21.8% T-PE and 29.3% PE); and consider their knowledge and experience “at an advanced level” (62.5% vs. 33.8% T-PE and 40.2% PE). Also, more are external (43.8% vs. 28.9% T-PE and 40.2% PE).

Once groups were identified, the question of the relative importance of variables in their respective practices was addressed through online practice modeling. There was much variation in the complexity of models produced. Some used as few as two links to describe their practice and one used 82 links. This is somewhat supportive of the diversity of practice described by Cousins and others when describing participatory evaluation (e.g., Cousins et al., in press). No two contexts or situations are alike and the evaluations that arise are never identical. Regardless of the perspective that an evaluator brings to the process, the client’s willingness and interest in engaging in the evaluation will define the degree to which it is participatory. The current study asked participants to model how they ensure stakeholder involvement and the outcomes they intend; the variance in model complexity mirrors the inherent complexity of participatory evaluation.

A most-endorsed model of T-PE was created using the top 21 links endorsed by the group of 142 T-PE evaluators (Figure 11). The most frequently endorsed link was between Stakeholder Involvement and Shared Understanding. The webinar participants highlighted the connection between these variables as requiring legitimate stakeholder participation in many of the evaluation activities. This also informed the credibility of evaluation findings, which is dependent upon stakeholders understanding them and their beliefs in their validity. This was

accomplished in practice by having the stakeholders at the table and deeply engaged in as much of the evaluation decision-making as practical, given resource and time constraints.

Social justice was not reflected in the T-PE model. This was unexpected given the filtering statements related to social justice. Webinar participants explained its absence as characteristic of the constructs of social justice and social action, which they said are too vague and undefined to be labeled outcomes. Some webinar participants also commented that setting social action or social justice as an outcome for an evaluation might put them in an advocacy role and they felt that would reduce the credibility of their objective evaluation voice. One preferred to have the data speak for itself. When asked if social justice or social action should be considered long term impacts, participants pushed back and said that these were more like the principles that guided their choice of work more than expected outcomes.

This finding suggests the modelers were not able to clearly articulate a social justice perspective when modeling practice. If we continue using logic modeling and program theory methodology to discuss evaluation theory and attempt to identify key evaluation practice variables within the categories of principles, activities, and outcomes, then we must also develop better ways to include principles in the modeling. The method used in this study only provided data on variables that had arrows drawn from or to it. Principles like social justice or social action, which underlie the choices of taking on a particular evaluation, by their very nature, might not link directly to activities or outcomes. Their absence in this model may thus be an artifact of the method used, rather than a fully accurate reflection of the evaluators' perspectives.

Research Question 4: How does the T-PE model compare to other evaluation theories?

The final research question addressed the need to situate this T-PE model within the broader evaluation theory landscape. A comparison group was identified and labeled as practical participatory (P-PE) based on their disagreement with two statements related to evaluation's effect on social justice and individual empowerment and emancipation. As shown in the radar chart (Figure 7), the P-PE group highly endorsed Community-Sensitive Sampling and Engage Intended Beneficiaries, and less often endorsed Community Trust. This lack of attention to trust is counterintuitive for any evaluator engaged in community sensitive sampling and engaging beneficiaries, because it seems central to these undertakings. What is perhaps most striking about this finding, though, is that the data suggest that those who did not believe evaluation should affect social justice and individual empowerment less often endorsed Community Trust. While community trust is a key tenet of a transformative approach that is also grounded in a social justice imperative (Mertens, 2009), this lends some support to the ability of the eight statements to create different subgroups. But, it should be taken quite tentatively because of the small sample size.

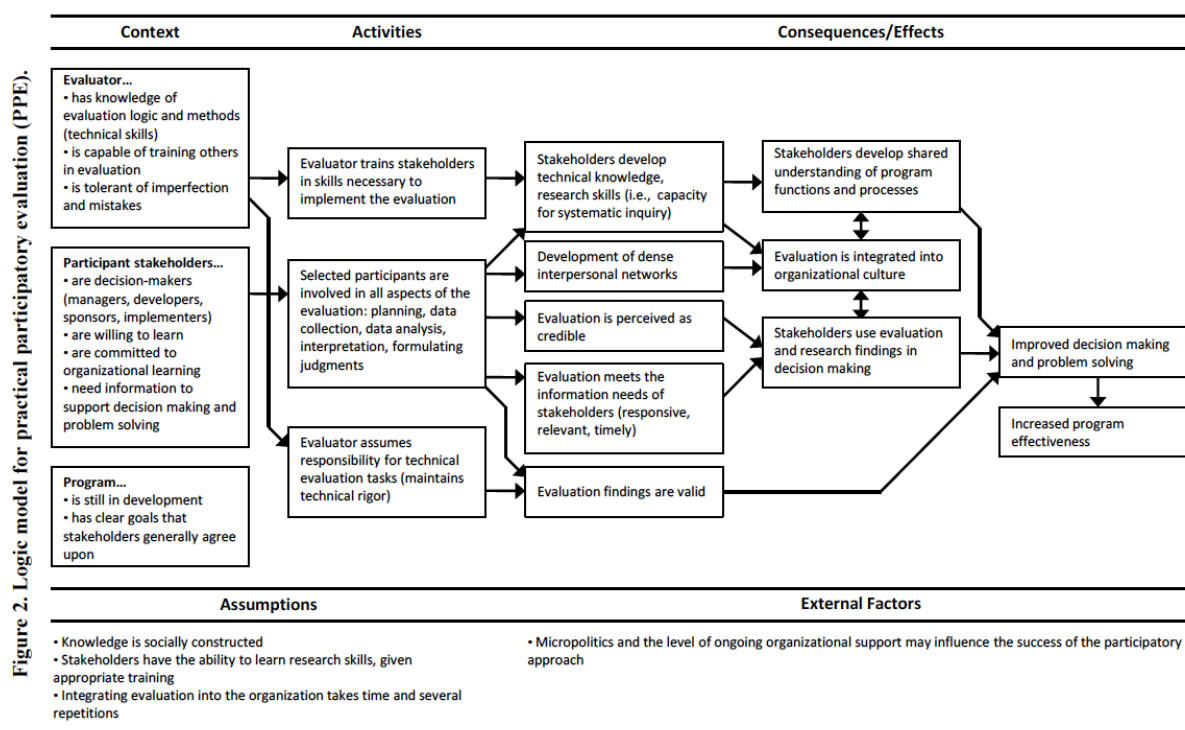
The grouping that distinguished T-PE evaluators from others was done with the intent of identifying a comparison group whose model might look different from the T-PE model in order to provide some context. Unfortunately, the grouping created a relatively small P-PE comparison group whose model had too little variance from the other model to make a reliable qualitative comparison. Nevertheless, quantitative comparisons between the models showed significant differences on 20 links. The P-PE links endorsements suggest a more utilitarian slant and the T-PE links suggested more of a focus on involving stakeholders and building trust within the community. The data suggest that those who disagreed that social justice and empowerment

outcomes are important for evaluation may see their process as more practical and less transformative.

Connections to Existing Knowledge

Evaluation theory development is not static. The UCLA group mentioned earlier has continued their efforts at modeling evaluation theories and a volume of *Evaluation and Program Planning* focused on evaluation theory modeling is currently in press. While their earliest work (Alkin, 2007) developed models of empowerment, practical participatory, transformative and utilization-focused evaluation, their later work, presented at the 2010 American Evaluation Association conference in San Antonio, Texas, focused on just three theories: value-engaged, practical participatory, and emergent realist evaluation. They used their models to compare the theories through Mark's (2008) evaluation research framework. They focused on creating visual representations of theory (Wallace & Hansen, 2010), visualizing context through theory decomposition (Vo, 2010), comparing activities (Dillman, 2010) and consequences (Luskin, 2010), and the utility of visual representations (Ho, 2010). Of their products, the model that provides the best comparison to this research's T-PE model is one prepared for their *Evaluation and Program Planning* article, that of practical participatory evaluation. It is included below in Figure 13.

Figure 13. Practical Participatory Evaluation Model by Hansen, Alkin, and Wallace (in press)



Structurally, this model looks quite similar to the one presented in Chapter 4 (Figure 11). They both use a left-to-right format with activities on the left and outcomes on the right. Arrows depicting causal mechanisms connect the variables. The addition of the left-hand column of “Context” and the lower box of “Assumptions” and “External Factors” provide a valuable framework when describing this form of evaluation.

Though it was not possible to incorporate these aspects with the modeling software employed in the current study (variables were only included when an arrow was drawn to or from it), the webinars provided additional perspective that could be incorporated as values. For instance, the fact that social justice may be the value underlying a choice to involve stakeholders could be

added to the model as an assumption with the caveat that the webinars provide limited support (i.e., comments were from 20% of webinar participants and only 10% of those represented in the model). What might also be added to the final model are the statements that identified those who built it. Inclusion of these additional descriptors would provide valuable context.

A major difference between these models is the seemingly weighted focus on the right side of the model presented in Figure 13. Here, there are three columns of Consequences/Effects, comprising 10 different outcomes, in comparison to one column of Activities with three boxes. The T-PE model produced in this research was lighter on outcomes and heavier on activities and principles, and it included only five outcomes, six principles and four activities. These latter two groupings are similar in scope to both the Activities and the Context columns included in Figure 13.

This difference in weights is possibly due to the limited number of outcomes made available to the T-PE modelers in this study. The variables offered by modelers included Use, Program Improvement, and Improving Evaluation Practice. Had these and other variables been included at the outset, more of them might have been included on the right side of the T-PE model.

There is some continuity between the two models: 15 T-PE model variables are represented in the model in Figure 13. These variables are central to participatory evaluation and should be included in any representation of participatory practice. Each is listed in Table 33 with its corresponding representation in the T-PE model.

Table 33. Comparison of UCLA P-PE Model to Current T-PE Model

UCLA P-PE Model	Current T-PE Model
Evaluator trains stakeholders in skills necessary to implement the evaluation	Educate; Build Capacity
Stakeholders develop technical knowledge, research skills (i.e., capacity for evaluation and systematic inquiry)	Build Capacity; Increase Systematic Inquiry
Selected participants are involved in all aspects of the evaluation: planning, data collection, data analysis, interpretation, formulating judgments	Use Local Program Knowledge; Develop Questions; Collect & Analyze Data; Develop Judgments & Recommendations
Stakeholders develop shared understanding of program functions and processes	Shared Understanding
Evaluation is perceived as credible	Shared Understanding; Credible Findings
Evaluation findings are valid	Shared Understanding; Credible Findings

Two variables included in their P-PE model are specific to practical participatory evaluation and were not offered to the T-PE modelers in this study: “Participant stakeholders are decision makers (managers, developers, sponsors, implementers)” and “Evaluator assumes responsibility for technical evaluation tasks (maintains technical rigor).” Typically, T-PE evaluators do not limit their selection of participating stakeholders to primary decision makers and they work to divest responsibility for decision-making to participating stakeholders as capacity is developed.

The other variables included in their P-PE model but missing from the T-PE model are similar across any participatory evaluation and were also excluded from this research in an effort to limit the cognitive complexity of the modeling process. Specifically, “Evaluation meets the information needs of stakeholders,” “Stakeholders use evaluation and research findings in

decision-making,” “Improved decision-making and problem-solving,” and “Increase program effectiveness” are all part and parcel to participatory evaluation and were rejected in the variable identification stage in the first phase of the study.

Interestingly, the variables that are included in the current T-PE model but are absent from the UCLA P-PE model are all related directly to the philosophical premise of expanded involvement and diverse perspectives. Table 34 provides an overview of these variables.

Table 34. Variables Unique to Transformative Participatory Evaluation

Variable	Definition
Diverse Perspectives	Evaluator ensures representation of diverse perspectives by including concerns, values, and interests of stakeholders.
Engage Intended Beneficiaries	Evaluator engages intended program beneficiaries in meaningful participation.
Increase Self-Critique	Increase participants’ capacity for self-critique.
Community Trust	Evaluator works to build trust by developing working relationships with a broad range of stakeholders.
Multiple Method Perspective	Evaluator applies multiple methods as appropriate to the evaluation context.

This difference supports the internal validity of the current model because it is on the dimensions of selection and philosophical differences that T-PE differs from P-PE (Cousins & Whitmore, 1998; Mertens, 2009). More precisely, T-PE is predicated on the belief that the more perspectives in an evaluation, the better, and that involvement of intended beneficiaries provides an empowerment opportunity that is missing in an evaluation that only involves key decision makers. Through this involvement in the process of developing questions and producing and analyzing data to answer those questions in a safe and supportive environment, participants should gain the capacity of self-critique.

Community trust is an important principle in transformative evaluation, and more specifically in T-PE, because it is through participation that community trust is built. By engaging a full range of stakeholders in an evaluation, the evaluator is showing respect for the values of the community; in time this respect translates into trust. In relation to Multiple Method Perspective, an evaluator in most any evaluation should choose methods appropriate to the context. This may be one of those variables that is a norm for any evaluation and need not be included in T-PE more specifically.

Though there is overlap between the current model and the P-PE model described above, and the current research used a similar process to develop key T-PE variables (literature review & expert panel), this study nevertheless makes a unique contribution to the evaluation literature. Specifically, its value lies in the contributions made by practitioners to the definitions and to the decisions about the inclusion of, and the relationships between variables in the final model.

Cousins et al., (in press) have advanced the notion of developing overarching principles of participatory evaluation. Their paper, forthcoming in the *American Journal of Evaluation*, points out three arguments for developing principles related to the approach, mostly in response to current research. They cite Daigneault and Jacob's (2009) conceptualization of participatory evaluation as defined within the tripartite frame of control, and depth, selection as full and complete. They feel there is no need to reduce or expand that framework because it fully conceptualizes the process elements of PE.

Their second argument is directly related to the current research. As mentioned earlier in this document, they feel that participatory and collaborative forms of evaluation are very complex, contextually bound, interpretivist practices that can never be fully operationalized with a two-

dimensional model. After seeing the work of Alkin (2007), and no doubt as a result of their early efforts in this area of inquiry, they became concerned that the “rigidity, and pre-ordinate nature” of modeling risks under-representing context and culture by such a “mechanistic representation” of dynamic processes (p. 6).

Finally, these scholars see efforts by the Collaborative, Participatory, and Empowerment Evaluation Topical Interest Group of the American Evaluation Association to parse the approaches in their TIG name into separate categories as sorely lacking. The dimensions the TIG presenters have used to parse PE from EE or CE are unnecessarily slim and practically impossible to witness in practice. Further, they are concerned that the creation of clear boundaries for these approaches may cause early career practitioners to feel they are prepared for an evaluation context if they know only one of these forms. It is extremely unlikely that an evaluator can choose a specific evaluation theory to implement before the negotiation process occurs between context, stakeholders, and evaluation needs. Setting apart these different approaches to participatory forms of evaluation may create unrealistic categories of practice.

Their line of inquiry, while somewhat predicated on the limited value of reductionist thinking about participatory practice, has the potential to add significantly to our understanding of practice. Developing key principles is a component of the current research and may prove useful to Cousins et al.,’s activities. Much like the schemas used to describe evaluation practice, the modeling produced in the current study is, by its nature, reductionist and limited. Its value is in the bridge it provides between theory and practice. Cousins et al., (in press) understate that value but Gargani (in press) supports it.

Coryn et al., (2011) conducted a systematic review of recent practice (1990–2009) of theory-driven evaluation (TDE). They developed a set of five core principles to describe it, each with no fewer than two subprinciples (Coryn et al., p. 205). They engaged a sample of theory-driven evaluation scholars and writers to ensure they reflected the key tenets of the theory-driven approach.

Coryn and colleagues posit TDE's five principles as: Theory Formulation; Theory-guided Question Formulation and Prioritization; Theory-guided Planning, Design, and Execution; Theory-guided Construct Measurement; and Identification of Breakdowns and Side Effects, Effectiveness or Efficacy, and Causal Explanations. Each of the subprinciples expands the related principle into discrete actions. Much like the principles resulting from the current research, these ideas might be used to better define theory-driven evaluation.

Like the UCLA group's work, Coryn and colleagues' work is based only on published literature and the scholars' insights and has not been interpreted through practice observation. These core principles can be used preliminarily in describing this form of evaluation and they are therefore very helpful in moving the evaluation research agenda forward. A next step would be to move beyond the relatively limited perspective of published articles that self-describe an approach as TDE and to compare them to actual practice.

These efforts at defining the key principles of various evaluation theories provide the field with the schemas that are part and parcel to theory development. With these developed, the next step is to test them in the field and refine them as necessary. The same can be said about the principles in the current research. Like those of Coryn and colleagues (2011), they begin to

describe the dimensions of T-PE, but the current research has taken the process into the next step by asking for endorsement from the practice community.

Another area where the current research relates to existing research is that of the intersection of T-PE and transformative evaluation. Since the beginning of this research, Dr. Mertens has continued her look at the transformative paradigm and its relationship to evaluation practice. In her most recent publication (Mertens & Wilson, 2012) she spends more time focusing on the participation element and how T-PE can be considered an application of the transformative perspective on evaluation.

The current study has attempted to connect participation to the transformative through the identification of T-PE evaluators using eight developed statements, and with the principle and outcome variables provided to modelers. Though the statements specifically included two transformative paradigm-oriented statements, transformative-oriented principles and outcomes that were expected to be more present for this group were not well represented in the T-PE model.

Implications

Theoretical Implications

Cousins et al., (1992, 1996) developed the process dimensions of participatory research and evaluation (control, depth, selection) and Cousins and Whitmore (1998) used them to describe two forms of participatory evaluation. This research advances our understanding of that schema by expanding the dimensions to include principles and outcomes. It also begins the process of developing a key set of variables and demonstrates how practitioners value these variables in their practice.

Discussing participatory evaluation practice through the three-dimensional lens of principles, activities, and outcomes expands our ability to describe and define our work. For instance, transformative participatory evaluation's process characteristics (activities) have been well described (e.g., Cousins & Earl, 1996; Cousins & Whitmore, 1998) and its philosophical foundations (principles) are covered by Mertens (e.g., 2009; Mertens & Wilson, 2012). But, the links between principles, activities, and outcomes have not been focused on. This research begins that exploration by showing that for these evaluators, a shared understanding is the most important outcome of their stakeholder involvement and that this shared understanding is closely linked to credible findings and the building of community trust. Furthermore, and counter to expectations, it places social justice and self-determination at the bottom of their valued principles, though these may exist at a level not measured by the modeling approach used in this study.

The current research offers a heretofore-unexplored method of empirically building evaluation theory. By bringing this methodology to the scientific process, this research has opened a new avenue of theory building. Practitioners and others who are distant from the research base can participate in this process by describing their practice using a familiar tool—that of program theory or logic model building. The resulting product is more accessible to practitioners than traditional research findings, and can serve as a bridge between theory and practice. In short, this process provides practitioners access to theory.

In 2006, Patton participated in a panel on process use at AEA's annual conference in Portland, Oregon, and subsequently wrote a chapter in the *New Directions for Evaluation* volume on process use (number 116, 2007). Before the panel, Patton had devoted 26 pages of the

third edition of his book, *Utilization-Focused Evaluation* (1997), to process use (Chapter 5 - *Intended Process Uses*). His fourth edition's chapter 5 (2008) contributed 42 pages to the topic. Clearly, as Alkin has noted (1991, 2004a), pushing theorists to reflect on their theories causes them to reconsider and further define their thoughts.

The model produced from this research, although endorsed by only a quarter of its participants, still provides a valuable starting point for theorists to reconsider their ideas about the key principles and outcomes of T-PE. By engaging this model in their work, the discipline's academics might further develop the definition of participatory evaluation broadly and transformative participatory evaluation more specifically.

Practice Implications

Twice during this study, in one form or another the question was raised "Do these types of evaluators exist or are you just identifying a group of practitioners who believe in this form of practice?" As noted earlier in this study, there are published reflective case examples of T-PE evaluation practice (e.g., Whitmore, 1994). Cousins and Chouinard (2012) identify more. T-PE is being practiced. While Cousins and Whitmore (1998) posited the two streams of practice from experience and an interpretation of the literature, this research goes two steps further by identifying evaluators who actually prefer to practice T-PE and developing a representation of their preferred practice. One implication to practice is that those practicing T-PE are no longer a group identified by a handful of published reflections but by their agreement with the eight statements developed in this study and their contribution to the T-PE model. Those who prefer this form of evaluation can now describe their practice in reference to this model.

The current research offers a visual representation of evaluation theory that will also aid practitioners in understanding their practice. Though this model may not represent everyone's practice in all evaluations, it can serve as a starting point for conversations with potential clients. Logic modeling is a common practice in evaluation, a tool in our evaluation toolbox. Therefore, practitioners may find this model a familiar and valuable representation of practice—one they can take apart and relate to their work. Dissecting the T-PE model and aligning it with their own practice can expand their understanding and interpretation of their practice and place it within the evaluation theory literature.

Use of this technology provides access to practitioners and in turn brings them into the theory building process. Though practitioners often feel very distant from the academic venture of theory building, their experiences can now be more easily included in the research. This has the potential to bridge the chasm of practice to theory and theory to practice by giving practitioners a direct connection to theory building.

In the process of engaging in theory building, practitioners will be invited to consider their own practice in a more reflective way than they might otherwise take on. By modeling their practice, they reflect on, and are forced to defend, the choices they make. Because we know that reflecting on one's actions has the potential to improve practice, this activity could improve evaluation. In fact, three of the 21 webinar participants mentioned that this experience made them consider their practice more deeply than they had in the past.

Participatory evaluation takes place most often in the complex world of real-time programs. By its very nature of attempting to engage stakeholders, PE is interpretive (Brisolara, 1998; Chambers, 2007). This model building exercise supports that notion. The wide variance in the

complexity of the T-PE model—543 different links for 142 people—shows that a relatively homogenous population of evaluators sees their practice in a wide range of ways.

Educational Implications

This study provides an accessible way of thinking about evaluation theory and practice, not only for practitioners, but also for students of the discipline. The model provides a visual representation of a theory of evaluation as defined by practitioners who use it. Visual representations are helpful in the classroom because of their accessibility. Just as the planetary representation of the atom helps us access the incredibly complex world of nuclear physics, the model created here provides students of evaluation access to a particular form of evaluation that might otherwise be overly complex. As the nuances of particular evaluation theories are defined and categorized, the process results in a tool that allows for diverse theories to be compared and understood.

Limitations

The statements used to distinguish T-PE practitioners from other evaluators, while deductively developed in close consultation with three experienced participatory evaluation theorists, may not have been restrictive enough and, as a result, may have under-identified the sample. The questions were piloted for clarity and understanding, but were not tested to determine how well they separated T-PE evaluators from other types of evaluators. The identification process was tested using other data gathered in this research, but the strength of these tests was limited by the fact that they used the data within the sample for testing. In fact, most of the participants indicated they agreed with all eight statements and the subgroups created were limited in their distinctiveness.

Logic modeling is usually accomplished with groups of practitioners working together to develop common understandings. That process is aided and accomplished through the social cognitions of a group. The modeling process used in this study occurred in a virtual vacuum, and therefore may suffer from a lack of interaction during the process. As well, it may have been difficult for respondents to conceptualize how the variables used in this stage would connect to one another. The webinars were designed to uncover and explain some of these difficulties, but they still provided only cursory considerations of the limitations.

The list of principles, activities, and outcomes developed and used in this research was necessarily limited in number because of the cognitive and algebraic complexity that would have occurred had the list included 40 or more potential variables. Moreover, the modeling results derived from a longer list would have had limited validity because of the potential complexity of an average model. But, had the list been longer and more inclusive, other key variables might have gained more importance and the resultant model might have looked quite different.

During development, a few variables were discarded because of these size limitations or because there was no agreement on a definition. The variables that were discarded generally reflected the transformative paradigm and, had they been included, they may have further informed the transformative nature of the T-PE model. On the other hand, given the unlimited complexity of any evaluation process, it is unrealistic that every key principle, activity, or outcome could be logically identified *a priori*. This is at the heart of the current criticism leveled at theory modeling (Cousins et al., in press).

From a list of common evaluation theories, participants were asked to select the one that most matched their practice. It is possible that practitioners who were not necessarily engaged

with the evaluation canon might not have known the difference between these theories and might have responded by selecting the one that most represented the theory guiding the research (C. A. Christie, personal communication, 2011). The comparison of subgroups in this research uses these theoretical orientation selections and should therefore be interpreted with some caution.

Future Research

In future studies, the list of 28 (26 original and 2 additions) variables produced by this research should be carefully examined. Specifically, the list might be narrowed or expanded so that T-PE is more accurately explained by the list. No list will ever cover all the possible variables, but using an Occam's razor perspective, the list could still be honed to include just those necessary to parse it from other general evaluations and from other participatory evaluations. Specifically, variables that help identify the transformative nature of participatory evaluation need more attention than this research was able to accomplish.

This model and list of variables can now be used in other research on evaluation. For example, the variables could be used in an analysis of case examples of T-PE to see if a) the variables are reflected in practice and b) the list is comprehensive enough to identify T-PE practice. The model could be compared with other participatory evaluations to deepen our descriptions of these theories. And finally, research could continue on the model to expand the qualitative interpretations. The variables and links would benefit from deeper descriptions.

If we continue the perspective that the model created here represents an evaluation theory, as a logic model represents a program theory, Chen (2005) might recommend this theory be tested for efficacy. For example, do participatory evaluations that develop shared understandings lead

to more credible findings? Other links in this model should be tested as well, until a substantial evidence base creates a more descriptive theory of transformative participatory evaluation.

Conclusions

Stufflebeam & Shinkfield's (2007) criticism of theory driven evaluation—that it expends valuable resources on explicating a program theory when well validated program theories rarely exist—may also inform the tremendous variance in the modeling data produced in this research. Though transformative participatory evaluation may have been well set out theoretically by Cousins & Whitmore (1998), it is extremely varied in practice and may not be easily defined by a set of variables. Is it then a waste of resources to try to describe it?

The current research stands with Chen (1992) in his valuing of greater understanding of theories in practice. Through deeper research and attempts at further articulation of the assumptions of different theories, our field, and the social sciences more generally, can reach a place of greater knowledge. Rather than discuss evaluation theory as a virtual “black box,” describing it in practice and attempting to discover the overlapping and divergent principles improves our abilities to interpret future evaluation practice.

This research deepens the discussion of participatory research by introducing principles and outcomes into the dialectic. This expanded schema, including the relative value among the variables, can be used to compare evaluation theories to increase our ability to describe the discipline's theories. Cousins and Whitmore (1998) used depth, control, and selection to describe T-PE and P-PE. Adding principles and outcomes to that schema will improve the definition of evaluation theories. For instance, while Patton (2008) clearly discusses the importance of a shared understanding in participatory evaluation, these T-PE evaluators endorsed it higher than

any other variable and linked it closely to credible findings and community trust. This lays the groundwork for articulating a key principle of T-PE where shared understanding is defined as an outcome that informs findings' credibility and the building of community trust. Other participatory evaluation theories, for instance P-PE, might not place such high value on community trust in developing shared understandings. This research brings empirical data to these conversations.

From a methodological perspective, this research uses a unique method of involving a large number of practitioners in the research process. Its use was not without difficulties, but it does have the potential to engage broader participation in a variety of research contexts. Gargani (2003) developed this technology for his own evaluation needs and it was repurposed for this research. With some further development, this technology can have broad application in both evaluation research and research on evaluation. It is the hope of this researcher that others take up this technology and apply it to the range of evaluation theories to test its meddle.

By engaging evaluators with a T-PE perspective, this research further develops our understanding of this evaluation theory by identifying those who actually prefer to practice this form of evaluation. No longer are T-PE evaluators only a theoretical type of evaluator. They were identified and are represented by this research. Though little data in this research help define these groups, other than their experience and education level and the eight identifying questions, the importance of their philosophical perspectives cannot be undervalued. It informs the choices they make and the evaluations they undertake.

The evaluation literature explains the transformative paradigm in a variety of ways. Cousins and Whitmore (1998) describe something similar to Mertens' transformative evaluation but stop

short of discussing the critical engagement with power struggles that is central to the underlying theories of TE (e.g., critical theory, feminist theory, indigenous theory, critical race theory, etc.). Sabo, in her dissertation (1999) and book (Sabo Flores, 2004), discusses a participatory evaluation model that is focused less on the broader social justice issue of TE and more on the transformation of the individual. Likewise, other writers who discuss youth participatory evaluation also focus on participation's value in affecting individual level change. This level seems neglected in the PE literature. In fact, this researcher struggled at the outset of the study with how "transformation" was defined in T-PE; one webinar participant touched on that subject by asking where transformation was in the T-PE model. It is the conclusion of this researcher that the transformation in T-PE first evolves from the philosophical perspective of the evaluator and that perspective directs the kinds of evaluations she undertakes. The nature of those evaluations is then oriented toward social justice and supports transformative axiology, epistemology, and ontology.

It should not be considered insignificant that this research was part catalyst to the work of Cousins, Whitmore, and Shulha (in press) on defining a key set of principles of participatory evaluation. It would be inappropriate to claim causation, but during the early stages of this research both Cousins and Whitmore were uncomfortable supporting the development of a model of T-PE. Surely their thinking had already started in that direction, but also as surely, their involvement in developing this research spurred them on to more deliberate effort at putting their thoughts to paper. While the current research may be an example of what they would not necessarily endorse, it has encouraged their research. In this small way, the current research has already contributed to the discipline.

References

- Alkin, M. C. (1979). An Approach to Evaluation Theory Development. *Studies In Educational Evaluation*, 5(2), 139-141.
- Alkin, M. C. (1985). *A Guide for Evaluation Decision Makers*. Beverly Hills, CA: Sage.
- Alkin, M. C. (1991). Evaluation Theory Development: II. In D. C. Phillips & M. W. McLaughlin (Eds.), *Evaluation and Education: At Quarter Century* (Part II., pp. 91-112). Chicago, IL: The National Society for the Study of Education.
- Alkin, M. C. (2004a). Comparing evaluation points of view. In M. C. Alkin (Ed.), *Evaluation Roots: Tracing Theorists' Views and Influences* (pp. 3-11). Thousand Oaks, CA: Sage.
- Alkin, M. C. (2007). Identifying critical processes and outcomes across evaluation approaches: Empowerment, Practical Participatory, Transformative, and Utilization-Focused. *American Evaluation Association Annual Conference*. Baltimore, MD.
- Alkin, M. C. (Ed.). (2004b). *Evaluation Roots: Tracing Theorists' Views and Influences*. Thousand Oaks, CA: Sage.
- Alkin, M. C., & Christie, C. A. (2004). An evaluation theory tree. In M. C. Alkin (Ed.), *Evaluation Roots: Tracing Theorists' Views and Influences* (pp. 12-66). Thousand Oaks, CA: Sage.
- Alkin, M. C., & House, E. R. (1992). Evaluation of programs. In M. C. Alkin (Ed.), *Encyclopedia of Educational Research* (6th ed., pp. 462-467). New York: Macmillan.
- Alkin, M. C., & Patton, M. Q. (1987). Working both sides of the street. *New Directions for Program Evaluation*, Winter(36), 19-32. doi:10.1002/ev.1464
- Alkin, M. C., Daillak, R. H., & White, P. (1979). *Using Evaluation: Does Evaluation Make A Difference?* Beverly Hills, CA: Sage.

- Alkin, M. C., Kosecoff, J., Fitz-Gibbon, C., & Seligman, R. (1974). *Evaluation and decision making: The title VII experience. CSE Monograph Series in Evaluation*. Los Angeles, CA: Center for the Study of Evaluation.
- American Evaluation Association. (2012). www.eval.org
- Amo, C., & Cousins, J. B. (2007). Going through the process: An examination of the operationalization of process use in empirical research on evaluation. *New Directions for Evaluation, Winter*(116), 5-26. doi:10.1002/ev
- Birckmayer, J. D., & Weiss, C. H. (2000). Theory-based evaluation in practice: What do we learn? *Evaluation Review, 24*(1), 407-431.
- Blackstock, K. L., Kelly, G. J., & Horsey, B. L. (2007). Developing and applying a framework to evaluate participatory research for sustainability. *Ecological Economics, 60*(4), 726-742. doi:10.1016/j.ecolecon.2006.05.014
- Brandon, P. (1998). Stakeholder participation for the purpose of helping ensure evaluation validity: Bridging the gap between collaborative and non-collaborative evaluations. *The American Journal of Evaluation, 19*(3), 325-337. doi:10.1016/S1098-2140(99)80215-X
- Brisolara, S. (1998). The history of participatory evaluation and current debates in the field. *New Directions for Evaluation, Winter*(80), 25-41. doi:10.1002/ev.1115
- Brunner, I., & Guzman, A. (1989). Participatory Evaluation: A Tool to Assess Projects and Empower People. *New Directions for Program Evaluation, Summer*(42), 9-18.
- Bryk, A. S. (1983). Editor's note. *New Directions for Program Evaluation, Spring*(17), 1-2. doi:10.1016/j.newast.2005.07.006
- Burke, B. (1998). Evaluating for a change: Reflections on participatory methodology. *New Directions for Evaluation, Winter*(80), 43-56. doi:10.1002/ev.1116

- Campbell, D. T. (1991). Methods for the Experimenting Society. *American Journal of Evaluation*, 12(3), 223-260. doi:10.1177/109821409101200304
- Chambers, R. (2007). *Who Counts? The Quiet Revolution of Participation and Numbers*. Brighton, UK: University of Sussex, Institute of Development Studies.
- Chambers, R. (2010). *Paradigms, Poverty and Adaptive Pluralism*. Brighton, UK: University of Sussex, Institute of Development Studies.
- Checkoway, B. N., & Gutierrez, L. M. (2006). Youth participation and community change: An introduction. In B. N. Checkoway & L. M. Gutierrez (Eds.), *Youth participation and community change* (pp. 1-9). New York: The Howarth Press.
- Checkoway, B., Dobbie, D., & Richards-schuster, K. (2003). Involving Young People in Community Evaluation Research. *Community Youth Development*, 4(1). Retrieved from <http://www.cydjournal.org/2003Spring/checkoway.html>
- Chen, H.-T. (1990). *Theory-Driven Evaluations*. Newbury Park, CA: Sage.
- Chen, H.-T. (2004). The roots of theory-driven evaluation: current views and origins. In M. C. Alkin (Ed.), *Evaluation Roots: Tracing Theorists' Views and Influences* (pp. 132-152). Thousand Oaks, CA: Sage.
- Chen, H.-T. (2005). *Practical Program Evaluation: Assessing and Improving Planning, Implementation, and Effectiveness*. Thousand Oaks, CA: Sage.
- Chen, H.-T., & Rossi, P. H. (1992). Introduction: Integrating theory into evaluation practice. In H.-T. Chen & P. H. Rossi (Eds.), *Using theory to improve program and policy evaluations* (pp. 1-11). New York: Greenwood Press.
- Chen, S., Poland, B. D., & Skinner, H. A. (2007). Youth Voices: Evaluation of participatory action research. *Society*, 22(1), 125-150.

- Coryn, C. L. S., Noakes, L. a., Westine, C. D., & Schroter, D. C. (2011). A Systematic Review of Theory-Driven Evaluation Practice From 1990 to 2009. *American Journal of Evaluation*, 32(2), 199-226. doi:10.1177/1098214010389321
- Coupal, F. P., & Simoneau, M. (1998). A case study of participatory evaluation in Haiti. *New Directions for Evaluation*, Winter(80), 69-79. doi:10.1002/ev.1118
- Cousins, J. B. (2006). [Review of the Book *Collaborative evaluations: A step-by-step model for the evaluator* by L. Rodriguez-Campos (2005)]. *Journal of Multidisciplinary Evaluation*, (5), 113-119.
- Cousins, J. B., & Chouinard, J. A. (2012). *Participatory evaluation up close: A review and integration of the research base*. Charlotte, NC: Information Age Press.
- Cousins, J. B., & Earl, L. M. (1992). The Case for Participatory Evaluation. *Educational Evaluation and Policy Analysis*, 14(4), 397. doi:10.2307/1164283
- Cousins, J. B., & Earl, L. M. (1995). Participatory Evaluation: Enhancing Evaluation Use and Organizational Learning Capacity. *Evaluation Exchange*, I(3&4).
- Cousins, J. B., & Leithwood, K. a. (1986). Current Empirical Research on Evaluation Utilization. *Review of Educational Research*, 56(3), 331-364. doi:10.2307/1170319
- Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation*, Winter(80), 5-23. doi:10.1002/ev.1114
- Cousins, J. B., Donohue, J. J., & Bloom, G. A. (1995). Collaborative Evaluation: Survey of Practice in North America. *Joint Meeting of the Canadian Evaluation Society and the American Evaluation Association*. doi:ERIC ID 413 364

- Cousins, J. B., Donohue, J. J., & Bloom, G. A. (1996). Collaborative Evaluation in North America: Evaluators' Self-reported Opinions, Practices and Consequences. *American Journal of Evaluation*, 17(3), 207-226. doi:10.1177/109821409601700302
- Cousins, J. B., Whitmore, E., & Shulha, L. M. (n.d.). Arguments for a Common Set of Principles for Collaborative Inquiry in Evaluation. *American Journal of Evaluation*.
- Crano, W. D., & Brewer, M. B. (2002). *Principles and Methods of Social Research* (2nd ed.). Mahwah, New Jersey: Lawrence Earlbaum Associates.
- Cronbach, L. J., & Associates. (1980). *Toward Reform of Program Evaluation*. Mahwah, New Jersey: Jossey-Bass.
- Daigneault, P.-M., & Jacob, S. (2009). Toward Accurate Measurement of Participation: Rethinking the Conceptualization and Operationalization of Participatory Evaluation. *American Journal of Evaluation*, 30(3), 330-348. doi:10.1177/1098214009340580
- DeVellis, R. F. (2003). *Scale Development: Theory and Applications* (2nd ed.). Thousand Oaks, CA: Sage.
- Dillman, D. A. (2007). *Mail and Internet Surveys: The Tailored Design Method* (2nd ed.). Hoboken, NJ: John Wiley & Sons.
- Dillman, L. (2010). Comparing evaluation activities across multiple theories of practice. *American Evaluation Association Annual Conference*, San Antonio, TX.
- Donaldson, S. (Claremont G. U. (2007). *Program Theory-Driven Evaluation Science: Strategies and Applications*. New York: Lawrence Earlbaum Associates.
- Donaldson, S. I., Christie, C. A., & Mark, M. M. (Eds.). (2009). *What Counts as Credible Evidence in Applied Research and Evaluation Practice?* Thousand Oaks, CA: Sage.

- Embretson, S. E., & Reise, S. P. (2000). *Item Response Theory for Psychologists*. Mahwah, New Jersey: Lawrence Earlbaum Associates.
- Estrella, M., & Gaventa, J. (1997). Who counts reality? Participatory monitoring and evaluation: A literature review. *International Institute for Rural Construction* (Vol. 5, pp. 1-70).
- Fetterman, D. M. (1994). Empowerment evaluation. *Evaluation Practice*, 15(1), 1-15. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/20520033>
- Fetterman, D. M. (1995). In Response. *American Journal of Evaluation*, 16(2), 179-199. doi:10.1177/109821409501600207
- Fetterman, D. M. (2004). Branching out or standing on a limb: Looking to our roots for insights. In M. C. Alkin (Ed.), *Evaluation Roots: Tracing Theorists' Views and Influences* (pp. 304-318). Thousand Oaks, CA: Sage.
- Fetterman, D. M. (2005). A window into the heart and soul of empowerment evaluation: Looking through the lens of empowerment evaluation principles. In D. M. Fetterman & A. Wandersman (Eds.), *Empowerment Evaluation Principles in Practice* (pp. 1-26). New York: The Guilford Press.
- Fetterman, D. M. (2008). personal communication.
- Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2010). *Program Evaluation: Alternative Approaches & Practical Guidelines*. Prentice Hall.
- Fleischer, D. N., & Christie, C. A. (2009). Evaluation Use: Results From a Survey of U.S. American Evaluation Association Members. *American Journal of Evaluation*, 30(2), 158-175. doi:10.1177/1098214008331009
- Frechtling, J. A. (2007). *Logic Modeling Methods in Program Evaluation*. San Francisco: Jossey-Bass.

- Gargani, J. (2003). A Novel Quantitative Method for Developing and Analyzing Program Theory. *unpublished manuscript*, 1-45.
- Gargani, J. (n.d.). What can practitioners learn from theorists' logic models? *Evaluation and Program Planning*, 1-8. Elsevier Ltd. doi:10.1016/j.evalprogplan.2012.03.020
- Gaventa, J. (1993). The powerful, the powerless, and the experts: Knowledge struggles in an information age. In P. Park, M. Brydon-Miller, B. Hall, & T. Jackson (Eds.), *Voices of Change: Participatory Research in the United States and Canada* (pp. 2-46). Toronto: OISE Press.
- Gerring, J. (1999). What makes a concept good? A critical framework for understanding concept formation in the social sciences. *Polity*, 31(3), 357-393.
- Gersten, R., & Hitchcock, J. (2009). What is credible evidence in education? The role of the What Works Clearinghouse in informing the process. *What Counts as Credible Evidence in Applied Research and Evaluation Practice?* (pp. 78-95). Thousand Oaks, CA: Sage.
- Goertz, G. (2006). *Social Science Concepts: A User's Guide*. Princeton, New Jersey: Princeton University Press.
- Gold, N. (1983). Stakeholder and Program Evaluation: Characterizations and Reflections. *New Directions for Program Evaluation*, March(17), 63-72.
- Goulet, D. (1989). Participation in development: New avenues. *World Development*, 17(2), 165-178. doi:10.1016/0305-750X(89)90242-8
- Greene, J. C. (1987). Stakeholder participation in evaluation design: Is it worth the effort? *Evaluation and Program Planning*, 10(4), 379-394. doi:10.1016/0149-7189(87)90010-3

- Greene, J. C. (1988a). Communication of results and utilization in participatory program evaluation. *Evaluation and Program Planning*, 11(4), 341-351. doi:10.1016/0149-7189(88)90047-X
- Greene, J. C. (1988b). Stakeholder Participation and Utilization in Program Evaluation. *Evaluation Review*, 12(2), 91-116. doi:10.1177/0193841X8801200201
- Greene, J. C. (2007). *Mixed Methods in Social Inquiry*. San Francisco: Jossey-Bass.
- Greene, J. C., & Caracelli, V. J. (1997). Defining and describing the paradigm issue in mixed-method evaluation. *New Directions for Evaluation*, Summer(74), 3-17.
- Greenwood, D., Whyte, W. & Harkavy, I. (1993). Participatory Action Research. *Human Relations*(46), 2, 175-192.
- Guba, E. G. (1990). The alternative paradigm dialog. In E. G. Guba (Ed.), *The Paradigm Dialog* (pp. 17-27). Newbury Park, CA: Sage.
- Hansen, M., Alkin, M. C., & Wallace, T. L. (n.d.). Depicting the logic of three evaluation theories. *Evaluation and program planning*. doi:10.1016/j.evalprogplan.2012.03.012
- Harnar, M. A., & Preskill, H. (2007). Evaluators' Descriptions of Process Use: An Exploratory Study. *New Directions for Evaluation*, Winter(116), 27-44. doi:10.1002/ev
- Henry, G. T., & Mark, M. M. (2003). Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions. *American Journal of Evaluation*, 24(3), 293-314.
doi:10.1177/109821400302400302
- Hickey, S., & Mohan, G. (Eds.). (2004). *Participation: From Tyranny to Transformation? Exploring New Approaches to Participation in Development*. New York: Zed Books.
- Ho, T. (2010). The utility of visual representations of evaluation theories. *American Evaluation Association Annual Conference*, San Antonio, TX.

- House, E. R. (1980). *Evaluation with Validity*. Beverly Hills, CA: Sage.
- House, E. R. (1991). Evaluation and social justice: Where are we? In M. W. McLaughlin & D. C. Phillips (Eds.), *Evaluation and Education: At Quarter Century. Ninetieth Yearbook of the National Society for the Study of Education* (pp. 233-247). Chicago, IL: University of Chicago Press.
- Howell, D. C. (2002). *Statistical Methods for Psychology* (5th ed.). Pacific Grove, CA: Duxbury.
- Hummelbrunner, R. (2010). Beyond logframe: Critique, variations and alternatives. In N. Fujita (Ed.) *Issues and Prospects of Evaluation for International Development Series IV: Beyond logframe; Using systems concepts in evaluation* (pp. 1-33).
- Institute of Development Studies. (n.d.). Retrieved April 2, 2012, from www.ids.ac.uk
- Isaac, S., & Michael, W. B. (1995). *Handbook in Research and Evaluation: For Education and the Behavioral Sciences* (3rd ed.). San Diego, CA: Educational and Industrial Testing Services.
- Jacob, S., Ouvrard, L., & Bélanger, J.-F. (2011). Participatory evaluation and process use within a social aid organization for at-risk families and youth. *Evaluation and program planning*, 34(2), 113-23. doi:10.1016/j.evalprogplan.2010.08.002
- Johnson, R. B. (1998). Toward a theoretical model of evaluation utilization. *Evaluation and Program Planning*, 21, 93-110.
- King, J. A. (1998). Making sense of participatory evaluation practice. *New Directions for Evaluation*, Winter(80), 57-67. doi:10.1002/ev.1117
- Kirkhart, K. E. (2000). Reconceptualizing evaluation use: An integrated theory of influence. *New Directions for Evaluation*, Winter(88), 5-23. doi:10.1002/ev.1188

- Lau, G., & LeMahieu, P. (1997). Changing roles: Evaluator and teacher collaborating in school change. *Evaluation and Program Planning*, 20(1), 7-15. doi:10.1016/S0149-7189(96)00032-8
- Lin, N. (1999). Building a Network Theory of Social Capital. *Connections*, 22(1), 28-51.
- Logic Model Development Guide: Using Logic Models to Bring Together Planning, Evaluation, and Action*. (2004).Development. Battle Creek, MI: W.K. Kellogg Foundation.
- Logic Model Training Module. (n.d.).*Program Development and Evaluation Logic Model Training*. University of Wisconsin-Extension. Retrieved April 2, 2012, from <http://www.uwex.edu/ces/pdande/evaluation/evallogicmodel.html>
- Luskin, R. J. (2010). Comparing the intended consequences of various theories of evaluation. *American Evaluation Association Annual Conference*, San Antonio, TX.
- Mark, M. M. (2008). Building a better evidence base for evaluation theory: beyond general calls to a framework of types of research on evaluation. In N. L. Smith & P. R. Brandon (Eds.), *Fundamental Issues in Evaluation* (pp. 111-134). New York: The Guilford Press.
- Mark, M. M., & Shotland, R. L. (1985). Stakeholder-based evaluation and value judgments. *Evaluation Review*, 9(5), 605-626.
- Mertens, D. M. (1999). Inclusive Evaluation: Implications of Transformative Theory for Evaluation. *American Journal of Evaluation*, 20(1), 1-14.
doi:10.1177/109821409902000102
- Mertens, D. M. (2005). Inclusive Evaluation. In S. Mathison (Ed.), *Encyclopedia of Evaluation*. Thousand Oaks, CA: Sage.
- Mertens, D. M. (2007). Transformative Paradigm: Mixed Methods and Social Justice. *Journal of Mixed Methods Research*, 1(3), 212-225. doi:10.1177/1558689807302811

- Mertens, D. M. (2009). *Transformative Research and Evaluation*. New York: The Guilford Press.
- Mertens, D. M., & Hopson, R. K. (2006). Advancing evaluation of STEM efforts through attention to diversity and culture. *New Directions for Program Evaluation*, 109(Spring), 35-51.
- Mertens, D. M., & Wilson, A. T. (2012). *Program Evaluation Theory and Practice: A Comprehensive Guide*. New York: Guilford Press.
- Meyer, L. H., Park, H., Grenot-scheyer, M., Schwartz, I., & Harry, B. (1998). Participatory research: New approaches to the research to practice dilemma. *The Journal of the Association for Persons with Severe Handicaps*, 23(3), 165-177.
- Miles, M. B., & Huberman, M. A. (1994). *Qualitative Data Analysis: An Expanded Sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.
- Miller, R. L. (2010). Developing Standards for Empirical Examinations of Evaluation Theory. *American Journal of Evaluation*, 31(3), 390-399. doi:10.1177/1098214010371819
- Miller, R. L., & Campbell, R. (2006). Taking Stock of Empowerment Evaluation. *American Journal of Evaluation*, 27, 296-319. doi:10.1177/1098214006291015
- Milstein, R. L., & Wetterhall, S. F. (1999). *MMWR 199 48(No. RR-11) Framework for Program Evaluation in Public Health*. Centers for Disease Control and Prevention. Retrieved from <http://www.cdc.gov/eval/framework/index.htm>
- Monkman, K., Miles, R., & Easton, P. (2007). The transformatory potential of a village empowerment program: The Tostan replication in Mali. *Women's Studies International Forum*, 30(6), 451-464. doi:10.1016/j.wsif.2007.09.005

- Morabito, S. M. (2002). Evaluator Roles and Strategies for Expanding Evaluation Process Influence. *American Journal of Evaluation*, 23(3), 321-330.
doi:10.1177/109821400202300307
- Murray, C. A. (1983). Stakeholders as deck chairs. *New Directions for Program Evaluation*, March(17), 59-61. doi:10.1002/ev.1325
- Naylor, P.-J., Wharf-Higgins, J., Blair, L., Green, L. W., & O'Connor, B. (2002). Evaluating the participatory process in a community-based heart health project. *Social science & medicine*, 55(7), 1173-87. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12365529>
- Onwuegbuzie, A. J., & Johnson, R. B. (2006). The Validity Issue in Mixed Research. *Research in the School*, 13(1), 48-63.
- Patton, M. Q. (1978). *Utilization-Focused Evaluation* (1st ed.). Beverly Hills, CA: Sage.
- Patton, M. Q. (1986). *Utilization-Focused Evaluation* (2nd ed.). Beverly Hills, CA: Sage.
- Patton, M. Q. (1997). *Utilization-Focused Evaluation* (3rd ed.). Thousand Oaks, CA: Sage.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Patton, M. Q. (2007). Process Use as a Usefulism. *New Directions for Evaluation*, Winter(116), 99-112. doi:10.1002/ev
- Patton, M. Q. (2008). *Utilization-Focused Evaluation* (4th ed.). Thousand Oaks, CA: Sage.
- Patton, M. Q., Grimes, P. S., Guthrie, K. M., Brennan, N. J., French, B. D., & Blyth, D. A. (1975). In search of impact: An analysis of the utilization of federal health evaluation research. doi:ERIC ID 135 938.
- Patton, M. Q., Grimes, P. S., Guthrie, K. M., Brennan, N. J., French, B. D., & Blyth, D. A. (1977). In search of impact: An analysis of the utilization of federal health evaluation

- research. In Carol H. Weiss (Ed.), *Using social research in public policy making* (pp. 141-164). Lexington, MA: D. C. Heath.
- Preskill, H., & Caracelli, V. (1997). Current and Developing Conceptions of Use: Evaluation Use TIG Survey Results. *American Journal of Evaluation*, 18(1), 209-225.
doi:10.1177/109821409701800122
- Preskill, H., Zuckerman, B., & Matthews, B. (2003). An Exploratory Study of Process Use: Findings and implications for future research. *American Journal of Evaluation*, 24(4), 423-442.
- Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A systematic approach*. Thousand Oaks, CA: Sage.
- Sabo Flores, K. (2008). *Youth Participatory Evaluation: Strategies for Engaging Young People*. San Francisco: Jossey-Bass.
- Sabo, K. (1999). *Young peoples involvement in evaluating the programs that serve them*. City University of New York. *ProQuest Dissertations and Theses*, 276 p.
<http://search.proquest.com/docview/304499267?accountid=10141>
- Sanders, J. & The Joint Committee on Standards for Educational Evaluation (1994). *The program evaluation standards: How to assess evaluations of educational programs*. Thousand Oaks, CA: Sage.
- Scriven, M. (1997). Empowerment Evaluation Examined. *Evaluation Practice*, 18(2), 165-176.
- Shadish, W. R. (1998). Evaluation Theory is Who We Are. *American Journal of Evaluation*, 19(1), 1-19. doi:10.1177/109821409801900102
- Shadish, W. R., Cook, T. D., & Leviton, L. C. (1991). *Foundations of Program Evaluation*. Newbury Park, CA: Sage.

- Sheirer, M. A. (1987). Program theory and implementation theory: Implications for evaluators. *New Directions for Program Evaluation, Spring*(33), 59-76.
- Smith, N. L. (1979). Evaluation reflections: Requirements for a discipline of evaluation. *Studies In Educational Evaluation, 5*, 5-12. doi:10.1016/0191-491X(81)90001-8
- Smith, N. L. (1993). Improving Evaluation Theory Through The Empirical Study of Evaluation Practice. *American Journal of Evaluation, 14*(3), 237-242.
doi:10.1177/109821409301400302
- Stake, R. E. (1975). Program Evaluation Particularly Responsive Evaluation. *Occasional Paper Series Paper #5*. Center for Instructional Research and Curriculum Evaluation, University of Illinois at Urbana-Champaign.
- Streiner, D. L., & Norman, G. R. (2008). *Health Measurement Scales: A Practical Guide to Their Development and Use* (4th ed.). New York: Oxford University Press.
- Stufflebeam, D. L. (1994). Empowerment Evaluation, Objectivist Evaluation, and Evaluation Standards: Where the Future of Evaluation Should Not Go and Where It Needs to Go. *American Journal of Evaluation, 15*(3), 321-338. doi:10.1177/109821409401500313
- Stufflebeam, D. L. (2001). Evaluation Models. *New Directions for Evaluation, Spring*(89), 7-98.
- Stufflebeam, D. L., & Shinkfield, A. A. (2007). *Evaluation Theory, Models, & Applications*. San Francisco: Jossey-Bass.
- Suarez-Balcazar, Y., & Harper, G. W. (2003). Community-Based approaches to empowerment and participatory evaluation. *Journal of Prevention & Intervention in the Community, 25*(3), 1-4.
- Tashakkori, A., & Teddlie, C. (1998). *Mixed Methodology: Combining Qualitative and Quantitative Approaches*. Thousand Oaks, CA: Sage.

- Teddlie, C., & Tashakkori, A. (2003). Major issues and controversies in the use of mixed methods in the social and behavioral sciences. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research* (pp. 3-50). Thousand Oaks, CA: Sage.
- Turnbull, B. (1999). The mediating effect of participation efficacy on evaluation use. *Evaluation and Program Planning*, 22(2), 131-140. doi:10.1016/S0149-7189(99)00012-9
- Vo, A. (2010). Visualizing context through theory decomposition. *American Evaluation Association Annual Conference*. San Antonio, TX.
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Cambridge, MA: Harvard University Press.
- Wallace, Tanner LeBaron, & Hansen, M. (2010). Visual representations of evaluation theories. *American Evaluation Association Annual Conference*. San Antonio, TX.
- Weaver, L., & Cousins, J. B. (2004). Unpacking the participatory process. *Journal of Multidisciplinary Evaluation*, 1(1), 19-40.
- Weiss, Carol H. (1972). *Evaluation Research: Methods of Assessing Program Effectiveness*. Upper Saddle River, NJ: Prentice-Hall.
- Weiss, Carol H. (1979). The Many Meanings of Research Utilization Published. *Public Administration*, 39(5), 426-431.
- Weiss, Carol H. (1983a). The stakeholder approach to evaluation: Origins and promise. *New Directions for Program Evaluation*, March(17), 3-14.
- Weiss, Carol H. (1983b). Toward the future of stakeholder approaches in evaluation. *New Directions for Program Evaluation*, March(17), 83-96.

- Weiss, Carol H. (1997). *Evaluation* (2nd ed.). Upper Saddle River, NJ: Prentice-Hall.
- What Works Clearinghouse. (n.d.). *Institute of Education Sciences*. Retrieved April 2, 2012, from <http://ies.ed.gov/ncee/wwc/>
- Whitmore, E. (1988). *Participatory evaluation approaches: Side effects and empowerment*. Cornell University. *ProQuest Dissertations and Theses*, 275 p.
<http://search.proquest.com/docview/303694502?accountid=10141>
- Whitmore, E. (1994). To tell the truth: Working with oppressed groups in participatory approaches to inquiry. In P. Reason (Ed.), *Participation in Human Inquiry* (pp. 82-98). Thousand Oaks, CA: Sage.
- Whitmore, E. (1998a). Final commentary: *New Directions for Evaluation*, Winter(80), 95-99.
- Whitmore, E. (1998b). We need to rebuild this house: The role of empowerment in evaluation of a Mexican farmers' cooperative. *Knowledge Shared: Participatory Evaluation in Development Cooperation* (pp. 217-230). Ottawa, Ontario, Canada: Kumarian.
- Whitmore, E., & McKee, C. (2001). Six street youth who could. In P. Reason & H. Bradbury (Eds.), *Handbook of Action Research* (pp. 396-402). London: Sage.
- Whitmore, E., Guijt, I., Mertens, D. M., Imm, P. S., Chinman, M., & Wandersman, A. (2006). Embedding improvements, lived experience, and social justice in evaluation practice. In I. F. Shaw, J. C. Greene, & M. M. Mark (Eds.), *Handbook of Evaluation: Policies, Programs, and Practices* (pp. 340-359). London: Sage.
- Whitmore, E., Wilson, M. G., & Calhoun, A. (Eds.). (2011). *Activism That Works*. Black Point, Nova Scotia: Fernwood.
- Wholey, J. S. (1979). *Evaluation: Promise and Performance*. Washington, D.C.: The Urban Institute.

- Wholey, J. S. (1987). Evaluability assessment: Developing program theory. *New Directions for Program Evaluation*, Spring(33), 77-92.
- Williams, B. (2010). Systems thinking and capacity development in the international arena. In N. Fujita (Ed.) *Issues and Prospects of Evaluation for International Development Series IV: Beyond logframe; Using systems concepts in evaluation* (pp. 35-53).
- Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2011). *The Program Evaluation Standards: A Guide for Evaluators and Evaluation Users/Editors* (3rd ed.). Thousand Oaks, CA: Sage.

APPENDIX A:

Online Editing Instructions To Theorists

Front Page Instructions on Google Wiki for Editing both Variables and Statements

Hello Brad, Bessa, and Donna.

Thank you for participating in developing my dissertation. Herein you should find all the resources you need to help identify a list of potential T-PE variables and write questions that will help identify T-PE practitioners from other participatory practitioners. Let me know if you don't find things you would like me to upload.

In the left hand navigation column there are two key links:

- [T-PE Variables List](#)
- [T-PE Questions](#)

These links take you to two different working documents. This is where you will do the bulk of your work. They are documents embedded into the web page that you can open in your browser as Google documents. Also, at the bottom of each page, there are sections for Attachments and Comments. The comment section is where you will send notes to us on your thoughts about these. I've written more directions on those pages, so click there and see what you think.

Two other key links are worth mentioning. The HOME link will always bring you back here. The FILES link takes you to a page where I've loaded some of the key articles relating to this process. If you think of something that I have not loaded and you think we should have it available, you can load it yourself, or tell me and I will load it.

Below is a place to add announcements about anything you might want all of us to see when we log in. Important articles or workshops can be listed here. Or just your morning ruminations about this process.

Thanks again! I look forward to your thoughts.

- Michael

Instructions for Editing Variables Document on Google Wiki

Here is a list of variables I've developed from Alkin's 2009 logic models work on P-PE and TE, and his group's work in 2010 on P-PE. As well, I've edited them to reflect my understanding of

T-PE. My first edit includes almost 50 variables, which I think might be too many.

You cannot edit this list on this page. You must open it into another web page. To do this, select the "Open..." link at the bottom of the list window and a new Google window will open in your browser. There will be a SAVE button in the top right hand corner of the page. Google generally saves your work often, but always check there before you close the window to make sure your work has been saved recently. When you are done editing the list simply close that window. Your changes will show on the below document when you refresh THIS page. If you want, you can keep both pages open and go between them, refreshing this page to show your work.

The two columns we will use in the model-building software are the "Variable" and the "Definition" columns. The notes column is for our work only.

If you find a variable that is unnecessary, rather than deleting it from the list, cross it out using "Strikethrough" under the FORMAT menu and write your reasons in the "Notes" column. When you write comments in the notes section or on any document on the wiki, please preface them with your initials.

If there are specific documents you'd like to attach for us reflect upon, you may use the "Attachments" link just above the Comments section.

Instructions for Statement Editing on Google Wiki

Below is a Google Docs document with the questions I've written (or taken from Cousins, Donohue, and Bloom, 1992) to help distinguish T-PE practitioners from other practitioners. I've included a few more questions than absolutely necessary so that I can do some stratifying of the final models across responses to some of these questions.

You cannot edit this document in this window. To edit this document, select the link at the bottom of its window and a new Google window will open in your browser with the document. This is a Google Doc, but it will work very much like a Word document, though the columns and numbering are more challenging to navigate. There will be a SAVE button in the top right hand corner. Google generally saves your work often, but always check there before you close the window to make sure your work has been saved recently. Your changes will show on the document below when you refresh THIS page.

If you find a question that is unnecessary, rather than deleting it from the document, cross it out using "Strikethrough" under the FORMAT menu and write your reasons in the "Comments" section on this page. I will follow up on your recommendations with comments of my own.

If there are specific documents you'd like to attach for us reflect upon, you may use the "Attachments" link just above the Comments section.

APPENDIX B:

Preliminary Variables And Definitions

Principle	Definition
Community Trust	Evaluator values community trust and works to build trusting relationship with community.
Power Structures	Evaluator analyzes program power relationships.
Evaluator Perception	Evaluator values how he/she is perceived by the community and works to manage that image.
Close Community Engagement	Evaluator engages in close involvement with community.
Negotiable Evaluation Focus	Evaluation focus is discussed by diverse stakeholder groups.
Methodological Pluralism	Evaluator embraces the idea of multiple methodologies as necessitated by the evaluation.
Diverse Perspectives	Evaluator ensures accurate representation of diverse perspectives.
Context-Sensitive Sampling	Sampling procedures are sensitive to diversity.
Reflective Evaluation	Evaluation reflects the concerns, values, and interests of collaborators.
Negotiable Decision Making	Technical decision-making roles for the evaluation are not predetermined.
Negotiable Participation	Stakeholder participation in evaluation knowledge production is not predetermined.
Activity	Definition
Smooth Implementation	Practical program implementation problems solved.
Program Theory Examined	Program theories are critically examined.
Participation Negotiated	Barriers to and supports necessary for participation are identified and negotiated.
Engage Marginalized Stakeholders	Evaluator engages marginalized stakeholders in meaningful participation.
Engage Program Beneficiaries	Evaluator engages program beneficiaries in meaningful participation.
Shared Decision-Making	Evaluator and stakeholders share evaluation decision-making on a negotiated basis.
Shared Responsibility	Evaluator and stakeholders share responsibility for evaluation.

Train Stakeholders	Evaluator trains stakeholders in necessary technical skills.
Educate Stakeholders on Evaluation	Evaluator educates stakeholders on the value of evaluation.
Evaluator Maintains Rigor	Evaluator's role includes maintaining sufficient technical rigor and adherence to professional standards of practice.
Local Knowledge Valued	Evaluation decisions are made using local knowledge.
Local Knowledge of Context Valued	Evaluation decisions are made using local knowledge of context.
Stakeholders Involved	Evaluator encourages participants to be involved in as many aspects of the evaluation as practical.
Outcome/Impact	Definition
Shared Understanding	Stakeholders develop shared understanding of program functions and processes.
Evaluation Is Integrated	Evaluation is integrated into organizational culture.
Use In Decision-Making	Stakeholders use evaluation findings in decision-making.
Increased Effectiveness	Increased program effectiveness.
Collaborators Learn	Collaborators learn technical skills.
Local Knowledge Developed	Local knowledge is developed and respected.
Credible Findings	Evaluation findings are seen as credible.
Timeliness	Evaluation findings are provided in a timely fashion.
Informed Decision-Making	Program decision-making is undertaken with information produced by the evaluation.
Evaluation Is Valued	Enhance evaluation value.
Evaluation Is Relevant	Enhance evaluation relevance.
Evaluation Is Used	Enhance utilization of the evaluation.
Improved Decision-Making	Improved organizational decision-making.
Individual Self-Critique	Increase individual capacity for self-critique.
Individual Self-Determination	Increase individual self-determination.
Systematic Inquiry Capacity	Increase capacity to engage in and use systematic inquiry.
Organizational Learning Capacity	Increase organizational learning capacity.
Social Justice	Enhance social justice.
Human Rights	Further human rights.
Social Action	Increase social action.
Outcomes Change	Outcome expectations change as a result of

the process.

APPENDIX C:

Revised Variables And Definitions (After Phase One)

1/30/2011

Principle	Definition
Community Trust	Evaluator works to build trust by developing working relationships with a broad range of community members.
Negotiable Focus	Evaluation focus is negotiated with diverse non-evaluator groups.
Method Pluralism	Evaluator embraces the idea of multiple methods as appropriate to the evaluation context.
Diverse Perspectives	Evaluator ensures accurate representation of diverse perspectives by addressing concerns, values, and interests of collaborators.
Negotiable Decision-Making	Technical decision-making roles for the evaluation are negotiated with diverse non-evaluator groups.
Negotiable Participation	Non-evaluator participation in evaluation knowledge production is not predetermined. Barriers to and supports necessary for participation are identified and negotiated.
Context-Sensitive Sampling	Sampling procedures are sensitive to community diversity.
Engage Marginalized Non-Evaluators	Evaluator engages marginalized non-evaluators in meaningful participation.
Engage Intended Program Beneficiaries	Evaluator engages intended program beneficiaries in meaningful participation.
Activity	Definition
Share Control	Evaluator negotiates divestment of control of the evaluation to non-evaluators.
Train	Evaluator trains non-evaluators in necessary technical skills.
Educate	Evaluator educates non-evaluators on the value of evaluation.
Value Local Knowledge	Evaluation decisions are made using local program knowledge.
Develop Questions	Evaluator involves non-evaluator participants in defining evaluation focus and question development.
Collect & Analyze Data	Evaluator involves non-evaluator participants in data collection and analysis.
Develop Judgments & Recommendations	Evaluator involves non-evaluator participants in formulating judgments and recommendations from the

Report & Disseminate	data. Evaluator involves non-evaluator participants in reporting and disseminating the findings.
Outcome	Definition
Shared Understanding	Non-evaluators develop shared understanding of program functions and processes.
Learning	Collaborators learn technical skills.
Integration	Evaluation is integrated into community culture.
Credible Findings	Evaluation findings are seen as credible.
Use	Enhance utilization of the evaluation.
Decision-Making	Improved organizational decision-making.
Systematic Inquiry	Increase capacity for individuals to engage in and use systematic inquiry.
Self-Critique	Increase individual capacity for self-critique.
Self-Determination	Increase individual self-determination, emancipation and empowerment.
Social Justice	Enhance social justice.
Social Action	Increase social action.
Outcomes Change	Outcome expectations change as a result of the process.
Program Improvement	Evaluation findings lead to improved program.

APPENDIX D:

Preliminary Statements

11/20/2010

Depth of Stakeholder Involvement

1. I do not make efforts to involve stakeholders in my evaluations.	SD	D	A	SA	N/A
2. I make some efforts to involve stakeholders in my evaluations, even if it's just to develop the evaluation framework or questions.	SD	D	A	SA	N/A
3. I involve stakeholders for a few key processes, like designing the framework and interpreting the findings.	SD	D	A	SA	N/A
4. I involve stakeholders in as many ways as possible in my evaluations.	SD	D	A	SA	N/A
5. An evaluation would not be a success if it did not have stakeholder participation. (I do not take on an evaluation unless it has a strong participatory component.)	SD	D	A	SA	N/A

Selection of Stakeholders

6. The more stakeholder groups involved in evaluation, the better.	SD	D	A	SA	N/A
7. Program beneficiaries should participate in carrying out evaluations.	SD	D	A	SA	N/A
8. Special interest groups should participate in carrying out evaluations.	SD	D	A	SA	N/A
9. People with a vital interest in programs (e.g., program developers, sponsors, directors) should participate in carrying out evaluations.	SD	D	A	SA	N/A
10. People responsible for implementing or delivering programs should participate in carrying out evaluations.	SD	D	A	SA	N/A

Control of the Evaluation

11. Evaluators should share control of evaluation projects equally with practitioners.	SD	D	A	SA	N/A
12. Evaluators should relinquish control of evaluation projects to stakeholders.	SD	D	A	SA	N/A

Philosophical Reasons for Stakeholder Involvement

13. Evaluators should educate practitioners about the power	SD	D	A	SA	N/A
---	----	---	---	----	-----

and value of evaluation as a planned change strategy.					
14. Evaluation should help practitioners improve practice.	SD	D	A	SA	N/A
15. Evaluation should stimulate practitioners to question fundamental beliefs and assumptions about practice.	SD	D	A	SA	N/A
16. Evaluation should result in fundamental changes in practice.	SD	D	A	SA	N/A
17. Evaluators should help train practitioners to do evaluations.	SD	D	A	SA	N/A
18. Practitioners' participation in evaluation should focus on enhancing the utilization of evaluation data.	SD	D	A	SA	N/A
19. Practitioners' participation in evaluation should focus on bringing about social justice.	SD	D	A	SA	N/A
20. Practitioners' participation in evaluation should focus on bringing about individual-level change.	SD	D	A	SA	N/A

APPENDIX E:

Revised Statements (After Phase One)

1/7/2011

Depth of Stakeholder Involvement

1. I always try to involve non-evaluator participants in my evaluations.	SD	D	A	SA
2. I prefer not to take on an evaluation unless it has a strong participatory component.	SD	D	A	SA

Selection of Stakeholders

3. Program beneficiaries should participate in carrying out evaluation.	SD	D	A	SA
4. People representing all important perspectives should be involved in any evaluation.	SD	D	A	SA

Control of the Evaluation

5. Evaluators should share technical decision-making with non-evaluator participants.	SD	D	A	SA
---	----	---	---	----

Philosophical Reasons for Stakeholder Involvement

6. Evaluators should help train all legitimate groups to do evaluation.	SD	D	A	SA
7. Evaluation should focus on bringing about individual empowerment, emancipation and self-determination.	SD	D	A	SA
8. Evaluation should focus on bringing about social justice.	SD	D	A	SA

APPENDIX F:

E-Mail Invitation For Pilot Study

From: Michael Harnar <eval.email.2010@gmail.com>
 Subject: Seeking volunteers to pilot my dissertation instrument
 Date: March 1, 2011 10:44:31 AM PST
 To: American Evaluation Association Discussion List <EVALTALK@bama.ua.edu>

Hello Eval Talkers,
 I am a doctoral candidate at Claremont Graduate University and I am looking for some practicing evaluators to help pilot my dissertation study.

My hope is to develop a preliminary model of transformative participatory evaluation. To do this I will ask the membership of the American Evaluation Association to take a survey. If a person's responses to the survey indicate he/she prefers to practice in a participatory fashion, they will be asked to model their preferred practice using an innovative online modeling software. Before I engage them in this process, I need a broad spectrum of practicing evaluators to take the survey and comment on the modeling process. You do not need to identify as "participatory" to help, all responses will be helpful.

Depending on how many notes you choose to provide, piloting my process should take less than an hour. The actual process participants will complete should take only about 30 minutes. For your time in this piloting, I can offer only my personal debt of gratitude, and the knowledge that you impacted ongoing research on evaluation and helped a struggling grad student move closer to completing. Your participation in, and contributions to, the pilot will be kept confidential. Your insights will be used only to improve this research. The Claremont Graduate University Institutional Review Board is currently reviewing this research and has provided conditional approval.

If you would like to participate, have any questions, or would like additional information about this research, please contact me directly at harnar.michael@gmail.com.

Thanks for your attention.

Best,

-Michael

Michael A. Harnar, M. A.

Doctoral Candidate in Evaluation and Applied Research Methods

Claremont Graduate University

School of Behavioral & Organizational Sciences

Claremont, CA 91711

www.cgu.edu/sbos

harnar.michael@gmail.com

"Question with wonder, rather than doubt."

APPENDIX G:

Final Survey Instrument

Participatory Evaluation Survey

Hello and welcome to my dissertation study. I am an evaluation doctoral candidate in the School of Organizational and Behavioral Sciences at Claremont Graduate University (CGU) and this is the informed consent page of the first phase of my dissertation study. You are being asked to participate because you are a member of the American Evaluation Association and I am interested in the experiences of AEA members.

The purpose of this study is to develop a model of Transformative Participatory Evaluation (T-PE) practice. If you choose to participate, you will help develop this model. There are three phases to my study and you may be invited to participate in all three. Though this first phase should take you fewer than 20 minutes, if you participate in all three phases participating could take you more than an hour.

In this first phase you will be asked questions that will help identify if you practice in accordance with theoretically described participatory evaluation. If you are not a practitioner, your participation will be relatively short. At your completion of this survey you will be offered the chance to opt into a drawing for the first incentive – a \$200 Amazon.com gift certificate.

If your answers to those questions confirm that your practice is at all participatory, you will be invited to continue to the next phase where I will ask you to use an online software to build a model of your participatory practice. Following your model building you will again be offered the opportunity to opt into another drawing for a \$200 Amazon.com gift certificate and to participate in the third phase -webinars to discuss a common model of practice. While phases 1 and 2 will occur back-to-back in this session, the webinars will be scheduled for a later date.

The potential risks associated with this study are not greater than those ordinarily encountered in daily life or during the performance of routine work. There is a risk that your practice may not look like the model that is produced by this research. The model is not likely to be reflective of every practitioner's practice and you should not consider your practice "wrong" if it is not reflected in the final model. I expect this research to benefit the evaluation community by providing a clearer picture of TPE practice and to better inform evaluation research schemas, and the science community at large because the resultant picture of T-PE practice can be applied to various disciplines.

Participation is completely voluntary. You may withdraw at any time and refuse to answer any question for any reason without penalty. Your decision whether or not to participate will in no way affect your current or future relationship with Claremont Graduate University, its faculty, students, or staff, or the American Evaluation Association. Your responses will be kept completely confidential and will not appear in any publication related to this research. The data

collected will be kept on a password protected computer and deleted when this research is completed.

If you have any questions or would like additional information about this research, please contact me at 909-524-7800, 315 Adams Ave., Pomona, CA, 91767, or harnar.michael@gmail.com. You can also contact my research advisor, Dr. Christina Christie at tina.christie@ucla.edu. The CGU Institutional Review Board, which is administered through the Office of Research and Sponsored Programs (ORSP), has approved this project. You may also contact ORSP at (909) 607-9406 with any questions.

Because our work will take place entirely online, this form allows no opportunity to attain your signature. By choosing to continue past this page, you attest to understanding the above information and that you voluntarily consent to participate in the survey that follows. If you would like a copy of this, please print it from your web browser window now.

-Michael A. Harnar, M.A.

There are 26 questions in this survey

Practice

1. On average, how many evaluations do you conduct per year? (By conducting evaluations I mean any role in designing and/or implementing evaluations, including supervising evaluations)

Please choose only one of the following:

- 1-3
- 4-6
- 7 or more
- None

2. Which of the following types of evaluations do you usually conduct?

Please choose all that apply:

- | | |
|---|--------------------------------|
| • Curricula evaluations | • Program evaluations |
| • Consumer evaluations | • Policy evaluations |
| • Performance auditing/monitoring/reviewing | • Evaluation of research |
| • Personnel evaluations | • Student/trainee evaluations |
| • Product evaluations | • I do not conduct evaluations |
| | • Other: |

Participatory Activity

3. In answering these questions, please think about how you prefer to practice evaluation. I know that answers to these questions are almost always context dependent, and "it depends" might be your answer choice. But, I would like you to think of your ideal evaluation situation.

The term "stakeholder" is used here to mean anyone, other than the evaluator, with a vested interest in the entity (evaluand) being evaluated.

"Participants" are those stakeholders who take an active role in the evaluation.

"Participation" is any active role and may vary widely in breadth and depth.

Please choose the appropriate response for each item:

Strongly Disagree; Somewhat Disagree; Somewhat Agree; Strongly Agree

- Intended program beneficiaries should participate in carrying out evaluation.
- People representing all important perspectives should be involved in any evaluation.
- People responsible for implementing or carrying out evaluations.
- Only key decision-makers (e.g., program developers, sponsors, directors) should participate in carrying out evaluations.
- Evaluators should share technical decision-making (e.g., survey instrument selection, statistical analyses, data presentation) with stakeholders.
- Evaluators should give technical decision-making (e.g., survey instrument selection, statistical analyses, data presentation) to stakeholders.
- Evaluators should maintain technical decision-making (e.g., survey instrument selection, statistical analyses, data presentation) of evaluation projects.
- Evaluators should help train all legitimate groups to do evaluation.
- Evaluation should focus on bringing about individual empowerment, emancipation, or self-determination.
- Evaluation should focus on bringing about social justice.
- Stakeholder involvement in an evaluation should focus on enhancing the utilization of evaluation data.

Participatory Practice Questions

4. Again, I would like you to think of your ideal evaluation situation. *

Please choose the appropriate response for each item:

Strongly Disagree; Somewhat Disagree; Somewhat Agree; Strongly Agree

- I always try to involve stakeholders in my evaluations.
- I prefer not to take on an evaluation unless it has a strong participatory component.
- I prefer to involve stakeholders in every possible stage of the evaluation.
- I prefer to involve stakeholders in very limited ways.

Practice

5. In which areas do you do your evaluation-related work?

Please choose all that apply:

- Adult education
- Advocacy and Policy Change
- Alcohol or Drug Abuse
- Business and industry
- Child care/early childhood education
- Disaster/Emergency management
- Educational technologies
- Environmental programs
- Evaluation methods
- Foundations
- Gender rights
- Government
- Health/Public health
- Higher education
- Human development
- Human resources
- Human services
- Indigenous peoples
- Information systems
- International/Cross cultural
- K-12 education
- Law/Criminal justice
- Lesbian, gay, bisexual and transgender issues
- Media
- Medicine
- Non-profits
- Organizational behavior
- Public policy/Public administration
- Science, technology, engineering, math (STEM)
- Social justice
- Social work
- Special needs populations
- Workforce/Economic development
- Youth development
- Other:

6. Describe the populations served by the programs you usually work with? (not only in the last year, but generally, what population do you work with?)

Please write your answer here: (open text box)

7. Currently, in which type of organization do you predominantly work?

Please choose all that apply:

- College or university
- School system
- State/Provincial government agency
- Federal government agency
- Local government agency
- Private business or consulting
- Nonprofit organization
- Other:

8. What size evaluations do you typically work on?

Please choose only one of the following:

- Small scale evaluations (i.e., smaller sample size, single-site or small multi-site, single-year)
- Large scale evaluations (i.e., larger sample size, multi-site, multi-year)
- A combination of both small scale and large scale evaluations

9. Which statement best describes your role as an evaluator?

Please choose only one of the following:

- I conduct evaluations primarily where evaluation services are external to the organization.
- I conduct evaluations primarily where evaluation services are internal to the organization.
- I conduct primarily a mix of both external and internal evaluations.

10. Where do you do most of your evaluation work?

Please choose all that apply:

- Africa
- Asia
- Australia/New Zealand
- Europe
- Central America
- North America
- South America

Experience

11. Dating from your first significant experience with an evaluation project, for how many years have you been involved in evaluation?

Please choose only one of the following:

- Less than two
- 2-5
- 6-10
- 11-15
- 16-20
- More than 20

12. How would you describe your professional identity as it relates to evaluation?

Please choose only one of the following:

- Evaluation is my primary professional identity
- Evaluation is my secondary professional identity after another discipline (e.g., education, psychology)
- Evaluation is not part of my professional identity

13. How would you categorize yourself in terms of evaluation knowledge and experience?

Please choose only one of the following:

- A relative beginner
- At an intermediate level
- At an advanced level

Training

14. What is the highest education level you have completed?

Please choose only one of the following:

- High school/some college
- Associate's degree
- Bachelor's degree
- Master's degree
- Doctoral degree

15. What is the field of your highest degree?

Please choose only one of the following:

- | | |
|---------------------------------|-------------------------|
| • Anthropology | • Psychology |
| • Advanced qualitative methods | • Public health |
| • Advanced quantitative methods | • Public policy |
| • Art/Music | • School administration |
| • Business | • Social welfare |
| • Economics | • Sociology |
| • Education | • Decline to answer |
| • Evaluation/Research methods | • Other |
| • Nursing/Medicine | |

16. Do you hold certificates in a field other than that of your highest degree?

Please choose only one of the following:

- Yes
- No

17. Please list any certificates in fields other than that of your highest degree.

Please write your answer here: (open text box)

Theoretical Orientation

18. Is your preferred theoretical orientation similar to any of these?

I know that many evaluators say that they design evaluations that are context specific and none of these orientations covers every evaluation. But, I also know that you probably have a perspective you 'prefer'. This is what I am asking for with this question -your preferred orientation.

Please choose only one of the following:

- | | |
|--------------------------------|---|
| • CIPP Model | • Theory driven |
| • Connoisseurship evaluation | • Utilization focused |
| • Developmental evaluation | • My theoretical orientation is not listed here |
| • Empowerment evaluation | • I do not know enough about these to select one |
| • Evaluation research | • I do not have a preferred theoretical orientation |
| • Fourth generation evaluation | |
| • Participatory evaluation | |
| • Social justice driven | |
| • Stakeholder evaluation | |

19. Is there a particular book or reference you use to guide your evaluation work? Please provide the title and author (e.g., Evaluation 2nd Edition; Carol Weiss)

Please write your answer here: (open text box)

Variables

20. On this and the next two pages I list a number of variables that are theorized to be important to participatory evaluation. I would like to know how important they are in your practice. To narrow the effort, I have categorized these variables and I ask you to only rank the top few in each category.

From the top box, select in order of importance, from highest to lowest, the 4 principles or activities you see as most important in your participatory evaluation practice.

- Community Trust -Evaluator works to build trust by developing working relationships with a broad range of stakeholders.
- Negotiable Purpose -The purpose of the evaluation is negotiated with stakeholders.
- Multiple Method Perspective -Evaluator applies multiple methods as appropriate to the evaluation context.
- Diverse Perspectives -Evaluator ensures representation of diverse perspectives by including concerns, values, and interests of stakeholders.
- Negotiable Decision Making -Technical decision-making (e.g., survey instrument selection, statistical analyses, data presentation) is negotiated with stakeholders.
- Negotiable Participation -Scope of stakeholder participation in evaluation is not decided ahead of time. Barriers to and supports necessary for participation are identified and negotiated.
- Community-Sensitive Sampling -Sampling procedures account for community diversity.
- Engage Marginalized Stakeholders -Evaluator engages marginalized program stakeholders (e.g., those who might otherwise lack representation) in meaningful participation.
- Engage Intended Beneficiaries -Evaluator engages intended program beneficiaries in meaningful participation.

Clicking on a variable moves it to the table.

21. Please select from this list the 4 most important principles or activities in your participatory evaluation practice.

- Build Capacity -Evaluator trains stakeholders in the necessary skills to participate in the evaluation.
- Share Control -Evaluator negotiates the giving of control of the evaluation to program stakeholders.
- Educate -Evaluator educates stakeholders on the value of evaluation.
- Use Local Program Knowledge -Evaluation decisions are made using local program knowledge.
- Develop Questions -Evaluator collaborates with stakeholders in defining evaluation purpose and evaluation questions.
- Collect & Analyze Data -Evaluator collaborates with stakeholders in data collection and analysis.
- Develop Judgments & Recommendations -Evaluator collaborates with stakeholders in interpreting findings, and formulating judgments and recommendations from the data.
- Report & Disseminate -Evaluator collaborates with stakeholders in reporting and disseminating the findings.

22. Please select the 4 most important outcomes of your participatory evaluation practice.

- Shared Understanding -All participants develop shared understanding of program functions and processes.
- Learning -All participants learn new skills.

- Credible Findings -Participants see evaluation findings as credible.
- Increase Systematic Inquiry - Increase capacity for participants to engage in and use systematic inquiry.
- Increase Self-Critique - Increase participants' capacity for self-critique.
- Increase Self-Determination - Increase individual self-determination, emancipation and empowerment.
- Increase Social Justice - Enhance social justice.
- Increase Social Action - Increase social action.
- Outcomes Change -Program outcome expectations change as a result of the process.

Opt In

23. Thank You for participating! Would you like to opt into the drawing for a \$200 Amazon.com gift card? Your contact information will be separated from your responses to this survey and will not be used for any purpose but to contact you if you win the drawing.

Please choose only one of the following:

- Yes
- No

24. Please provide the best contact information if you win the drawing. (open text box)

Link

25. As mentioned in the opening description, to help understand how participatory practice works, I am looking for practitioners who involve stakeholders in their practice to help me understand what that practice looks like.

Because your survey responses show you practice in a participatory fashion, I want to invite you to map your practice for me. As with the survey, upon completing this next level of participation you will have the chance to opt in to another drawing for a \$200 Amazon.com gift certificate.

This next stage should take you about 20-30 minutes to complete, depending on how complex you choose to make your model. Either way, your participation is greatly appreciated. Would you like to continue?

Please choose only one of the following:

- Yes
- No

26. Please click this link to be forwarded to the model building stage. When you get to the next page, you will be provided an abbreviated informed consent for that phase and asked to log in with the email address I used to first contact you for this survey.

Thanks for continuing.

datagraph.gcoinc.com/part_eval_survey/login.php

Disregard the open text box above. No information needs to be included, just click the above link to be taken to the model-building software.

Thank you very much for your participation. If you have any questions, please feel free to contact me at harnar.michael@gmail.com.

APPENDIX H:

Statements (After Phase One & Pilot Testing)

1/9/2011

“In answering these questions, please think about how you prefer to practice evaluation. I know that answers to these questions are almost always context dependent, and “it depends” might be your answer choice. But, I would like you to think of your ideal evaluation situation.”

Depth of Stakeholder Involvement

1. I always try to involve non-evaluator participants in my evaluations.	SD	D	A	SA
2. I prefer not to take on an evaluation unless it has a strong participatory component.	SD	D	A	SA

Selection of Stakeholders

3. Program beneficiaries should participate in carrying out evaluations.	SD	D	A	SA
4. People representing all important perspectives should be involved in any evaluation.	SD	D	A	SA
5. Evaluators should help train all legitimate groups to do evaluation.	SD	D	A	SA

Control of the Evaluation

6. Evaluators should share technical decision-making with non-evaluator participants.	SD	D	A	SA
---	----	---	---	----

Philosophical Reasons for Stakeholder Involvement

7. Evaluation should focus on bringing about individual empowerment, emancipation and self-determination.	SD	D	A	SA
8. Evaluation should focus on bringing about social justice.	SD	D	A	SA

APPENDIX I:

Survey Invitation

Dear {FIRSTNAME},

Let me first tell you that this email is an invitation to participate in my dissertation study. The Executive Director of the American Evaluation Association (AEA) and the Claremont Graduate University Institutional Review Board have approved this invitation. The intent of my study is to develop a greater understanding of participatory evaluation practice and develop a practice-informed model of transformative participatory evaluation. You have been invited to participate because you are a member of AEA. The AEA Executive Director provided me access to your email address for this study only. You may opt out of receiving any further communication related to this study by selecting this “opt out” link {OPTOUTURL}.

You have the potential to be involved in up to three stages of the research and at each stage there is a drawing for a \$200 gift certificate to Amazon.com. As you can imagine, your chances of being rewarded increase as you complete more and more phases.

The first phase is to complete a short survey that will take you 15–20 minutes to complete. When you complete this survey you may opt in to a drawing for a \$200 Amazon.com gift certificate. Depending on how you respond to the survey questions, you may be invited to participate in the second and third phases. Each phase is described more fully in its subsequent invitation, but essentially, in phase two participants will create a logic model-like representation of their evaluation practice and in phase three practitioners will participate in webinars to discuss a common model.

The product of this study will be a theory- and practice-informed model of transformative participatory evaluation practice that will have a qualitative component that further explains some of the elements.

If you are interested in participating, please click the link below and read a more thorough informed consent before continuing to the survey.

If you are invited to the second phase of the study, you will be asked to log in using the email address I sent this invitation to: {EMAIL} (all lower case letters)

Thank you very much for your willingness to participate.

- {ADMINNAME} ({ADMINEMAIL})
909-524-7800

You are receiving this email as a member of the American Evaluation. This research request was reviewed by the AEA Executive Director. If you have concerns about the survey and would like

to express them to the AEA leadership, please email them to aea@eval.org. Any concerns raised will be shared, confidentially, with the Executive Committee of the association. AEA allows its membership list to be used infrequently for research that focuses on the field of evaluation. If you would like to optout of AEA's research list, please send an email request to heidi@eval.org. Please note that we encourage you to consider remaining on the list as such research strengthens and furthers the field's knowledge base.

To participate, please click on the link below.

Click here to do the survey:
{SURVEYURL}

APPENDIX J:

Survey Reminder

To: [firstname] [lastname]

From: Michael A. Harnar, Claremont Graduate University

Last week you received an invitation to participate in a descriptive study of participatory evaluation aimed at developing a practice-informed model of transformative participatory evaluation. You have been invited to participate in this research on evaluation because you are a member of AEA and your perspective is important. I hope you will consider participating before June 13th when this phase's data collection ends.

The Claremont Graduate University Institutional Review Board has approved this study and AEA has provided me access to your email address for this study only. You may opt out of receiving any further communication related to this study by selecting the "opt out" link below.

You have the potential to be involved in up to three stages of this research. Phase 1 is a survey, phase 2 is an online modeling exercise, and phase 3 is a webinar. The first two phases each take about 20 minutes and the third phase may take as much as an hour. How you answer the questions in phase 1 will decide if you are a candidate for phase 2. If invited to participate in phase 2, it would be best if you continue on straight through from phase 1 directly into phase 2. Therefore, consider starting this survey only when you have 40 minutes available. At the end of each stage there is an option to participate in a drawing for a \$200 gift certificate to Amazon.com as an incentive for your participation.

Any data you provide will be kept completely confidential and no participants will be named in any publication. You may refuse to answer any question and you may terminate your participation at any time.

If you are interested in participating, please click the link below and read a more thorough informed consent before continuing to the survey.

Thank you very much for your interest in research on evaluation.

Sincerely,

Michael A. Harnar, michael.harnar@cgu.edu

909-524-7800

You are receiving this email as a member of the American Evaluation. This research request has met the requirements for use of the membership list. If you have concerns about the survey and would like to express them to the AEA leadership, please email them to aea@eval.org. Any concerns raised will be shared, confidentially, with the Executive Committee of the association. AEA allows its membership list to be used infrequently for research that focuses on the field of

evaluation. If you would like to optout of AEA's research list, please send an email request to heidi@eval.org. Please note that we encourage you to consider remaining on the list as such research strengthens and furthers the field's knowledge base.

Click here to do the survey:

{SURVEYURL}

If you do not want to participate in this survey and don't want to receive any more invitations to participate in Michael Harnar's dissertation study please click the following link:

{OPTOUTURL}

APPENDIX K:

Modeling Reminder

To: {FIRSTNAME} {LASTNAME}

From: Michael A. Harnar, Claremont Graduate University

Thank you very much for completing phase 1 of my research project. You are receiving this follow up because I noticed that you agreed to participate in phase 2 but have not yet completed a model. Many people have requested that I send them the link to phase 2 because they completed phase 1 but ran out of time and wanted to return later for phase 2. Therefore, in the event that you ran out of time to complete phase 2, I wanted to provide you with the link.

You have until midnight PDT June 13th to complete a model. At that time I will close this phase and begin analysis.

As before, your email address serves as both login and password for this section. It should take you about 20 minutes to complete, depending on how long you take with the modeling. Click this link to begin:

http://datagraph.gcoinc.com/part_eval_survey/login.php

Whether or not you complete this next phase will have no impact on your relationship with me, AEA, or Claremont Graduate University. Your information will be kept completely confidential and once I have downloaded the data your email login will be stripped away and replaced by a serial number. No information you provide will be traced back to you and your identity will not be revealed in any publication.

Thanks again for completing phase 1 and I hope you can make time to continue through phase 2. As a reminder, you will have the option to enter a drawing for a \$200 Amazon.com gift certificate upon completion of a model.

Best,

- Michael A. Harnar (michael.harnar@cgu.edu)

APPENDIX L:

Modeling Interface



Login and Informed Consent

Thanks for continuing your participation in my research!

In this 2nd phase of my dissertation research you will use web-based model building software to depict your practice. This may take you anywhere from 20-30 minutes, depending on how detailed you choose to make your model.

Though you will receive no direct compensation for your participation, you may opt into a drawing for a \$200 Amazon.com gift certificate at the end of this phase.

Your participation is completely voluntary. You have the right to withdraw at any time or refuse to answer any question(s) for any reason, without penalty.

If you have any questions, please contact me at harnar.michael@gmail.com. You can also contact my research advisor, Dr. Christina Christie at tina.christie@ucla.edu.

By continuing with the model building you attest to this statement: "I understand the above information and have had all of my questions about participation in this research project answered. I voluntarily consent to participate in this research." If you would like a copy of this, please print it from your web browser window now.

Please log in below using the email address that I used to first contact you.

email address (all lowercase letters):

password (email address -- all lowercase letters)

© Copyright 2009 Gargani + Company, Inc. All Rights Reserved.
Designed, developed and hosted by Dayspring Technologies, Inc.



Welcome to the model building portion of the research

In this phase of the research you will be asked to:

1. Draw a logic model-like depiction of your evaluation practice.
2. Respond to 1 short-answer question about your logic model.
3. Answer 2 questions about your participation in this research.

[Let's begin!](#)

© Copyright 2009 Gargani + Company, Inc. All Rights Reserved.
Designed, developed and hosted by Dayspring Technologies, Inc.

INSTRUCTION PAGE

INSTRUCTION PAGE

datagraph.gcoinc.com/p/

Google

Most Visited Latest Headlines Bookmarks

GO

INSTRUCTION PAGE

Great, thanks for getting this far! This page provides instructions on how I would like you to model your practice. On the next page, you will be asked to use a unique, web-based software to create a model that reflects your participatory evaluation practice. You will be provided a list of principles, activities, and outcomes that are theorized to be part of participatory evaluation, and a drawing space where you will place those principles. I am interested in how many you use and how they relate to one another. In building that model, I would like you to answer this question:

How do you ensure stakeholder involvement and what outcomes do you intend to create?

You may believe that the answer to this question is context dependent. In practice, I know it is almost always context dependent. But, I would like you to consider your IDEAL evaluation experience and build a model to reflect that experience. This question will be available on the next page while you build your model.

I have included one variable at the center of the workspace for you to build from: Stakeholder Involvement - Stakeholders are involved in the evaluation. You can move this variable around the space as necessary and add as many factors as you like, but you cannot remove the blue rectangular variable.

You may not find all the outcomes, principles, and actions you would like to use. Do your best with what is available. Also, do not feel like you must use all the variables. After you are done with your model, you will have the opportunity to name and describe a key variable that was missing.

Here are some instructions on the using the software you will see on the next page, these are also covered in a video embedded below these instructions:

- Identify as many or few factors listed on the left that you wish to include in your logic model.
- Drag and drop one factor at a time from the list onto the drawing area.
- When you place the cursor over the center of a factor in the drawing area, it will turn into a pencil.
- Use the pencil cursor to draw an arrow by clicking and dragging the cursor from one concept to another.
- You use arrows to indicate that the concept at the beginning of the arrow influences or helps bring about the concept that the arrow points to.
- When you place the cursor near the border of a factor in the drawing area, it will turn into a hand.
- Use the hand cursor to position the factor on the drawing area.
- You can delete a factor or an arrow by selecting it and then pressing the delete key.
- You cannot delete the rectangular factor "Stakeholder Involvement."
- There are no right or wrong answers. Draw a model that best fits your beliefs.



[Begin drawing your model.](#)

© Copyright 2009 Gargani + Company, Inc. All Rights Reserved.
Designed, developed and hosted by Dayspring Technologies, Inc.

ABP 1P

DRAWING AREA

datagraph.gcoinc.com/part_eval_survey/graph.php

Drag idea on to drawing

- ☐ Increase Self-Determination
- ☐ Develop Questions
- ☐ Share Control
- ☐ Negotiable Purpose
- ☐ Outcomes Change
- ☐ Shared Understanding
- ☐ Collect & Analyze Data
- ☐ Community-Sensitive Sampling
- ☐ Community Trust
- ☐ Increase Social Justice
- ☐ Build Capacity
- ☐ Diverse Perspectives
- ☐ Negotiable Decision Making
- ☐ Negotiable Participation
- ☐ Increase Self-Critique
- ☐ Engage Intended Beneficiaries
- ☐ Increase Social Action
- ☐ Increase Systematic Inquiry
- ☐ Learning
- ☐ Credible Findings
- ☐ Develop Judgments & Recommendations
- ☐ Educate
- ☐ Engage Marginalized Stakeholders
- ☐ Use Local Program Knowledge
- ☐ Report & Disseminate
- ☐ Multiple Method Perspective

How do you ensure stakeholder involvement and what outcomes do you intend to create?

Done

Stakeholder Involvement



COMMENT PAGE

Use this space to name and define ONE key outcome, principle, or action that was not provided on the previous page. (Note: this box is required. If I provided all the variables you needed on the previous page and you have no comments just type "none" in the box. Thanks!)

Submit



SOME FINAL QUESTIONS

Would you like to have your contact info added to the drawing for a \$200 Amazon.com Gift Certificate?

- ☐ No
☐ Yes

As I mentioned in the opening description, to help understand how participatory practice works, I am looking for practitioners who involve stakeholders to a certain degree in their practice. My research involves one more step and I hope you will continue your participation. The information you just provided in the survey and by modeling your ideal practice will be combined with other's responses. By combining them, I will create one representative model of transformative participatory evaluation practice. Then, I will hold 2-3 webinars where I would like transformative practitioners to interpret the model as it reflects their practice. Everyone who participates in a webinar will also be offered the option to enter another drawing for a \$200 Amazon.com gift certificate. If your survey responses reflect a transformative approach to participatory evaluation, I would like to contact you for the webinar. Can I contact you to participate? Again, your experiences in practice are very important to the field.

Would you like to participate in a webinar and have one more chance at an Amazon.com gift card?

- ☐ No
☐ Yes

© Copyright 2009 Gargani + Company, Inc. All Rights Reserved.
Designed, developed and hosted by Dayspring Technologies, Inc.



Thank you for participating in my dissertation research!

If you have any questions, please contact me at harnar.michael@gmail.com or 909-524-7800.

© Copyright 2009 Gargani + Company, Inc. All Rights Reserved.
Designed, developed and hosted by Dayspring Technologies, Inc.

APPENDIX M:

Webinar Invitation

Hello,

You are receiving this email because you participated in my survey and model-building study and agreed to be contacted about participating in the final phase's webinar. Thank you very much for participating up to this point and for agreeing to participate in a webinar.

During the webinar we will discuss a model of transformative participatory evaluation that was created by combining your model with others' whose survey responses were similar to yours. The webinar will use a semi-structured design that will leave room for open discussion of the model around a few key topic areas. Namely, it is important for me to know how this model reflects current practice in the field and how context influences choices around this model. Your insight will be key to this understanding.

As with other pieces of this research, you have the option of participating as much or as little as you wish. You may choose not to answer any question without any repercussion. Each webinar is expected to take about an hour. At the end of the webinar you will have the option to opt into a drawing for a \$200 Amazon.com gift certificate.

I will record the webinars for analysis purposes only. You may choose to participate using a pseudonym to protect your identity. The recording will be kept in my possession and will only be used to verify memory of the conversation. No individual identity will be linked to comments on the webinar and no names will be used in any publication or presentation of this research. The recordings will be deleted when my dissertation is approved for publication.

Please click this link _____ that will take you to a survey page where you can select which date and time you will participate. If you cannot make any of the offered dates, you may participate after the webinars. I will post the final model online and offer opportunities for comments by participants. You can provide your comments about the final model there for consideration.

Thanks again for your participation and I look forward to your comments in the webinar.

Best,

- Michael

APPENDIX N:

Webinar Protocol

I will start with a 5-10 minute opening to introduce myself and outline the webinar process.

I will then ask for volunteers to speak for a few minutes at a time. I will moderate the conversation by using the webinar software control panel to manage the “audio space” in the beginning. After the first person speaks, I will open all the microphones to allow for a dialogue.

Background

In the survey, I asked you to answer 15 questions related to your “preferred evaluation practice.” Given the choices of Strongly Disagree, Slightly Disagree, Slightly Agree or Strongly Agree, you somewhat or strongly agreed with these eight statements:

- Evaluators should share technical decision-making with stakeholders.
- I always try to involve stakeholders in my evaluations.
- I prefer not to take on an evaluation unless it has a strong participatory component.
- Intended program beneficiaries should participate in carrying out evaluation.
- People representing all important perspectives should be involved in any evaluation.
- Evaluators should help train all legitimate groups to do evaluation.
- Evaluation should focus on bringing about individual empowerment emancipation or self-determination.
- Evaluation should focus on bringing about social justice.

These statements, collaboratively produced for this study with Cousins, Whitmore, & Mertens, were designed to identify evaluators who prefer to practice transformative participatory evaluation (T-PE) (Cousins & Whitmore, 1998). Because you and others (n=142) agreed with these statements, I am making the assumption that your preferred form of evaluation practice fits the Cousins and Whitmore definition of T-PE.

In the modeling phase, I provided 26 variables (also collaboratively produced w/Cousins, Whitmore, & Mertens) and asked you to use these variables to “model” your practice by answering the question “How do you ensure stakeholder involvement and what outcomes do you intend to create?” The complexity of models created by those who agreed with the above eight statements was extreme. Some used just 1 or 2 variables and drew 2 arrows (links) between them while one person created 59 links between almost all 27 variables. The challenge for me was to produce a model that most represents your work and yet is not too complex to describe. I decided to choose the median number of links (median=21) as the model’s level of complexity.

The attached model uses the 21 most-used links (arrows from one variable to another). This level of complexity (21 links) is the median number of links of the models created by those who, like

you, agreed with the above eight statements. My research thesis statement is that this model is highly representative of T-PE preferred practice.

You were invited to the webinar because not only did you elect to be included, but also because your model used many of these 21 links. Therefore, this model should look somewhat familiar and you are in the best position to describe this model in practice.

The focus of the webinar is to obtain qualitative insight into this quantitatively created model. In particular, I am interested in where this model *actually* reflects your practice and any stories or experiences that help bring this model to life. To that end, I ask that you take a look at the model and the variable definitions included and consider how your practice might be reflected there.

There will be 2–3 people on the webinar. If you do not have access to a phone or VoIP capability and cannot call into the conference call, I hope that you will type up your comments and share them with us during the webinar.

The guiding questions I would like you to consider as you prepare for the webinar are:

1. How is your *actual* practice reflected in this model?
 - a. In particular, can you provide real-world examples of one or two (or more) of the links in this model?
 - b. What contextual variables enable or hinder a particular variable or link between variables?

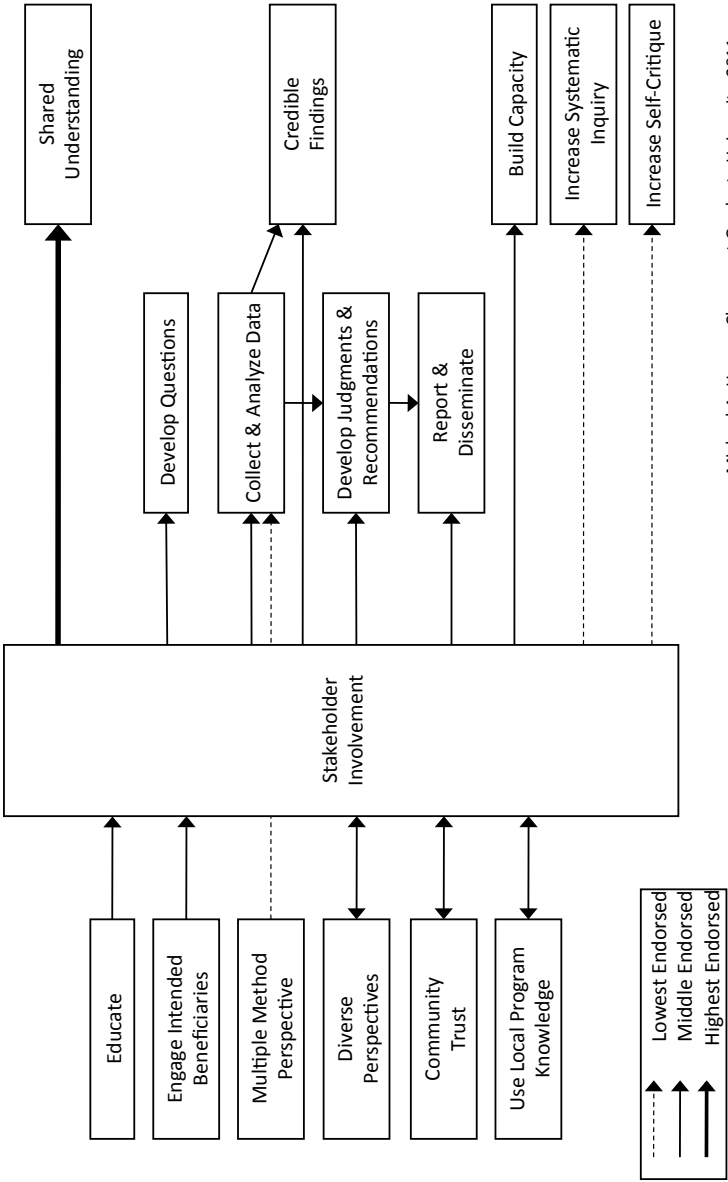
A few links and points of interest, given models created by other practitioners, are:

- Stakeholder Involvement's links to
 - Increased Self-Critique and
 - Increased Systematic Inquiry
- The reciprocal relationships between Stakeholder Involvement and
 - Community Trust and
 - Diverse Perspectives
- Of the 21 top-endorsed links, only 4 of them are *not* connected directly to Stakeholder Involvement.
- Two variables noted in the survey as important outcomes of evaluation did not rise to be included in the model: "Increase Social Justice" and "Increase Self-Determination."

Thank you again for participating in my dissertation study. Your insight from the practice arena is critical to forwarding the area of research on evaluation. I am honored by your willingness to share your experience and I look forward to talking with you.

T-PE Model

[DRAFT 10/30/2011]
(top 21 links; n=142 participants)
(25-37 endorsers per link, or 18%-26% of 142 participants)



Principles and Definitions

3/17/2011

“Stakeholder” is a broad term intended to mean community and program members who have a vested interest in the outcome of an evaluation. In this survey, evaluators are not considered stakeholders.

“Participants” are those stakeholders who take an active role in the evaluation.

“Participation” is any active role and may vary widely in intensity.

Principle	Definition
Community Trust	Evaluator works to build trust by developing working relationships with a broad range of stakeholders.
Negotiable Purpose	The purpose of the evaluation is negotiated with stakeholders.
Multiple Method Perspective	Evaluator applies multiple methods as appropriate to the evaluation context.
Diverse Perspectives	Evaluator ensures representation of diverse perspectives by including concerns, values, and interests of stakeholders.
Negotiable Decision-Making	Technical decision-making (e.g., survey instrument selection, statistical analyses, data presentation) is negotiated with stakeholders.
Negotiable Participation	Scope of stakeholder participation in evaluation is not decided ahead of time. Barriers to and supports necessary for participation are identified and negotiated.
Community-Sensitive Sampling	Sampling procedures account for community diversity.
Engage Marginalized Stakeholders	Evaluator engages marginalized program stakeholders (e.g., those who might otherwise lack representation) in meaningful participation.
Engage Intended Beneficiaries	Evaluator engages intended program beneficiaries in meaningful participation.
Share Control	Evaluator negotiates the giving of control of the evaluation to program stakeholders.
Build Capacity	Evaluator trains stakeholders in the necessary skills to participate in the evaluation.
Educate	Evaluator educates stakeholders on the value of evaluation.
Use Local Program Knowledge	Evaluation decisions are made using local program knowledge.
Develop Questions	Evaluator collaborates with stakeholders in defining evaluation purpose and evaluation questions.

Collect & Analyze Data	Evaluator collaborates with stakeholders in data collection and analysis.
Develop Judgments & Recommendations	Evaluator collaborates with stakeholders in interpreting findings, and formulating judgments and recommendations from the data.
Report & Disseminate	Evaluator collaborates with stakeholders in reporting and disseminating the findings.
Shared Understanding	All participants develop shared understanding of program functions and processes.
Learning	All participants learn new skills.
Credible Findings	Participants see evaluation findings as credible.
Increase Systematic Inquiry	Increase capacity for participants to engage in and use systematic inquiry.
Increase Self-Critique	Increase participants' capacity for self-critique.
Increase Self-Determination	Increase individual self-determination, emancipation and empowerment.
Increase Social Justice	Enhance social justice.
Increase Social Action	Increase social action.
Outcomes Change	Program outcome expectations change as a result of the process.

APPENDIX O:

Final Variables Set

Principle	Definition
Use Local Program Knowledge	Evaluation decisions are made using local program knowledge.
Educate	Evaluator educates stakeholders on the value of evaluation.
Engage Marginalized Stakeholders	Evaluator engages marginalized program stakeholders (e.g., those who might otherwise lack representation) in meaningful participation.
Engage Intended Beneficiaries	Evaluator engages intended program beneficiaries in meaningful participation.
Share Control	Evaluator negotiates the giving of control of the evaluation to program stakeholders.
Multiple Method Perspective	Evaluator applies multiple methods as appropriate to the evaluation context.
Community Sensitive Sampling	Sampling procedures account for community diversity.
Diverse Perspectives	Evaluator ensures representation of diverse perspectives by including concerns, values, and interests of stakeholders.
Stakeholder Support	Commitment to participatory evaluation by program leaders and other key stakeholders.
Negotiable Purpose	The purpose of the evaluation is negotiated with stakeholders.
Activity	Definition
Negotiable Participation	Scope of stakeholder participation in evaluation is not decided ahead of time. Barriers to and supports necessary for participation are identified and negotiated.
Negotiable Decision-Making	Technical decision-making (e.g., survey instrument selection, statistical analyses, data presentation) is negotiated with stakeholders.
Build Capacity	Evaluator trains stakeholders in the necessary skills to participate in the evaluation.
Develop Questions	Evaluator collaborates with stakeholders in defining evaluation purpose and evaluation questions.
Collect & Analyze Data	Evaluator collaborates with stakeholders in data collection and analysis.
Develop Judgments & Recommendations	Evaluator collaborates with stakeholders in interpreting findings, and formulating judgments and recommendations from the data.
Report & Disseminate	Evaluator collaborates with stakeholders in reporting and disseminating the findings.
Outcome	Definition
Shared Understanding	All participants develop shared understanding of program

	functions and processes.
Community Trust	Evaluator works to build lasting trust by developing working relationships with a broad range of stakeholders through designing and implementing an evaluation, and providing credible findings.
Learning	All participants learn new skills.
Credible Findings	Participants see evaluation findings as credible.
Increase Self-Critique	Increase participants' capacity for self-critique.
Increase Systematic Inquiry	Increase capacity for participants to engage in and use systematic inquiry.
Increase Social Action	Increase social action.
Increase Self-Determination	Increase individual self-determination, emancipation and empowerment.
Increase Social Justice	Enhance social justice.
Learning Loop	Learning generated by the evaluation is incorporated into the organization and improves future stages of the evaluation.
Outcomes Change	Program outcome expectations change as a result of the process.

APPENDIX P

Variables and Statements Editing Evolution

	Evolution of Key Variable Names & Definitions <i>Variable names that are shown in parentheses had slight variations through the three stages.</i> <i>Each accurate name is included with the respective definition.</i>			
		DEFINITIONS		
		PRELIMINARY	AFTER PHASE ONE	FINAL
	PRINCIPLES			
Context	Use Local Program Knowledge	(See Activities.)	(See Activities.)	Evaluation decisions are made using local program knowledge.
Relationships	Community Trust	Evaluator values community trust and works to build trusting relationship with community.	Evaluator works to build trust by developing working relationships with a broad range of community members.	(See Outcomes.)
	Power Structures	Evaluator analyzes program power relationships.	n/a	n/a
	Evaluator Perception	Evaluator values how he/she is perceived by the community and works to manage that image.	n/a	n/a
	Close Community Engagement	Evaluator engages in close involvement with community.	n/a	n/a
Input/ Perspective	Diverse Perspectives	Evaluator ensures accurate representation of diverse perspectives.	Evaluator ensures accurate representation of diverse perspectives by addressing concerns, values, and interests of collaborators.	Evaluator ensures representation of diverse perspectives by including concerns, values, and interests of stakeholders.
	Reflective Evaluation	Evaluation reflects the concerns, values, and interests of collaborators.	n/a	n/a
Engagement	Engage Marginalized Non-Evaluators	(See activities.)	Evaluator engages marginalized non-evaluators in meaningful participation.	
	(Engage Intended Program Beneficiaries)	(See activities.)	Engage Intended Program Beneficiaries: Evaluator engages intended program beneficiaries in	Engage Intended Beneficiaries: Evaluator engages intended program beneficiaries in

	Evolution of Key Variable Names & Definitions <i>Variable names that are shown in parentheses had slight variations through the three stages. Each accurate name is included with the respective definition.</i>			
		DEFINITIONS		
		PRELIMINARY	AFTER PHASE ONE	FINAL
			meaningful participation.	meaningful participation.
	Engage Marginalized Stakeholders	(See activities.)	n/a	Evaluator engages marginalized program stakeholders (e.g., those who might otherwise lack representation) in meaningful participation.
	Share Control	n/a	(See activities.)	Evaluator negotiates the giving of control of the evaluation to program stakeholders.
Information Sharing	Educate	(See activities.)	(See activities.)	Evaluator educates stakeholders on the value of evaluation.
Buy-In	Stakeholder Support	n/a	n/a	Commitment to participatory evaluation by program leaders and other key stakeholders.
Goals	(Negotiable Focus/Purpose)	Negotiable Evaluation Focus: Evaluation focus is discussed by diverse stakeholder groups.	Negotiable Focus: Evaluation focus is negotiated with diverse non-evaluator groups.	Negotiable Purpose: The purpose of the evaluation is negotiated with stakeholders.
Evaluation Design	(Methodological Pluralism)	Methodological Pluralism: Evaluator embraces the idea of multiple methodologies as necessitated by the evaluation.	Method Pluralism: Evaluator embraces the idea of multiple methods as appropriate to the evaluation context.	Multiple Method Perspective: Evaluator applies multiple methods as appropriate to the evaluation context.
	(Context-Sensitive Sampling)	Context-Sensitive Sampling: Sampling procedures are sensitive to diversity.	Context-Sensitive Sampling: Sampling procedures are sensitive to community diversity.	Community-Sensitive Sampling: Sampling procedures account for community diversity.
General	Negotiable Decision-Making	Technical decision-making roles for the evaluation are not predetermined. (See also Activities.)	Technical decision-making roles for the evaluation are negotiated with diverse non-evaluator groups.	(See Activities.)
	Negotiable Participation	Stakeholder participation in evaluation knowledge production is not predetermined. (See also Activities.)	Non-evaluator participation in evaluation knowledge production is not predetermined. Barriers to and supports necessary for participation are identified and	(See Activities.)

	Evolution of Key Variable Names & Definitions <i>Variable names that are shown in parentheses had slight variations through the three stages. Each accurate name is included with the respective definition.</i>			
		DEFINITIONS		
		PRELIMINARY	AFTER PHASE ONE	FINAL
			negotiated.	
	ACTIVITIES			
Process	Smooth Implementation	Practical program implementation problems solved.	n/a	n/a
Program	Program Theory Examined	Program theories are critically examined.	n/a	n/a
Engagement & Participation	(Negotiable Participation)	Participation Negotiated: Barriers to and supports necessary for participation are identified and negotiated.	(See Principles.)	Negotiable Participation: Scope of stakeholder participation in evaluation is not decided ahead of time. Barriers to and supports necessary for participation are identified and negotiated.
	Engage Marginalized Stakeholders	Evaluator engages marginalized stakeholders in meaningful participation.	n/a	(See principles.)
	Engage Program Beneficiaries	Evaluator engages program beneficiaries in meaningful participation.	(See principles.)	(See principles.)
	Stakeholders Involved	Evaluator encourages participants to be involved in as many aspects of the evaluation as practical.	n/a	n/a
	Share Control	n/a	Evaluator negotiates divestment of control of the evaluation to non-evaluators.	(See Principles.)
Decision-Making	(Negotiable Decision-Making)	Shared Decision-Making: Evaluator and stakeholders share evaluation decision-making on a negotiated basis. (See also Principles.)	(See principles.)	Negotiable Decision-Making: Technical decision-making (e.g., survey instrument selection, statistical analyses, data presentation) is negotiated with stakeholders.
Relationships	Shared Responsibility	Evaluator and stakeholders share responsibility for evaluation.	n/a	n/a

	Evolution of Key Variable Names & Definitions <i>Variable names that are shown in parentheses had slight variations through the three stages. Each accurate name is included with the respective definition.</i>			
		DEFINITIONS		
		PRELIMINARY	AFTER PHASE ONE	FINAL
Information Sharing	(Build Capacity)	Train Stakeholders: Evaluator trains stakeholders in necessary technical skills.	Train: Evaluator trains non-evaluators in necessary technical skills.	Evaluator trains stakeholders in the necessary skills to participate in the evaluation.
	(Educate Stakeholders)	Educate Stakeholders on Evaluation: Evaluator educates stakeholders on the value of evaluation.	Educate: Evaluator educates non-evaluators on the value of evaluation.	(See principles.)
Evaluation Design	Evaluator Maintains Rigor	Evaluator's role includes maintaining sufficient technical rigor and adherence to professional standards of practice.	n/a	n/a
Context	(Local Knowledge)	Local Knowledge Valued: Evaluation decisions are made using local knowledge.	Value Local Knowledge: Evaluation decisions are made using local program knowledge.	(See principles.)
	Local Knowledge of Context Valued	Evaluation decisions are made using local knowledge of context.	(See above.)	(See Principles.)
	Develop Questions	n/a	Evaluator involves non-evaluator participants in defining evaluation focus and question development.	Evaluator collaborates with stakeholders in defining evaluation purpose and evaluation questions.
	Collect & Analyze Data	n/a	Evaluator involves non-evaluator participants in data collection and analysis.	Evaluator collaborates with stakeholders in data collection and analysis.
Findings	Develop Judgments & Recommendations	n/a	Evaluator involves non-evaluator participants in formulating judgments and recommendations from the data	Evaluator collaborates with stakeholders in interpreting findings, and formulating judgments and recommendations from the data.
	Report & Disseminate	n/a	Evaluator involves non-evaluator participants in reporting and disseminating the findings	Evaluator collaborates with stakeholders in reporting and disseminating the findings.
	OUTCOMES/IMPACTS			
Individual Development	(Learning)	Collaborators Learn: Collaborators learn technical skills.	Learning: Collaborators learn technical skills.	Learning: All participants learn new skills.

	Evolution of Key Variable Names & Definitions <i>Variable names that are shown in parentheses had slight variations through the three stages.</i> <i>Each accurate name is included with the respective definition.</i>			
		DEFINITIONS		
		PRELIMINARY	AFTER PHASE ONE	FINAL
	Shared Understanding	Stakeholders develop shared understanding of program functions and processes.	Non-evaluators develop shared understanding of program functions and processes.	All participants develop shared understanding of program functions and processes.
	(Self-Critique)	Individual Self-Critique: Increase individual capacity for self-critique.	Self-Critique: Increase individual capacity for self-critique.	Increase Self-Critique: Increase participants' capacity for self-critique.
	Individual Self-Determination	Individual Self-Determination: Increase individual self-determination.	Self-Determination: Increase individual self-determination, emancipation and empowerment.	Increase Self-Determination: Increase individual self-determination, emancipation and empowerment.
	(Systematic Inquiry)	Systematic Inquiry Capacity: Increase capacity to engage in and use systematic inquiry.	Systematic Inquiry: Increase capacity for individuals to engage in and use systematic inquiry.	Increase Systematic Inquiry: Increase capacity for participants to engage in and use systematic inquiry.
Organizational Development	Increased Effectiveness	Increased program effectiveness.	n/a	n/a
	Informed Decision-Making	Program decision-making is undertaken with information produced by the evaluation.	n/a	n/a
	(Improved Decision-Making)	Improved Decision-Making: Improved organizational decision-making.	Decision-Making: Improved organizational decision-making.	n/a
	Organizational Learning Capacity	Increase organizational learning capacity.	n/a	n/a
	Outcomes Change	Outcome expectations change as a result of the process.	Outcome expectations change as a result of the process	Program outcome expectations change as a result of the process.
	(Integration)	Evaluation is Integrated: Evaluation is integrated into organizational culture.	Integration: Evaluation is integrated into community culture.	n/a
	Program Improvement	n/a	Evaluation findings lead to improved program.	n/a
	Learning Loop	n/a	n/a	Learning generated by the evaluation is incorporated into the organization and improves future

	Evolution of Key Variable Names & Definitions <i>Variable names that are shown in parentheses had slight variations through the three stages. Each accurate name is included with the respective definition.</i>			
		DEFINITIONS		
		PRELIMINARY	AFTER PHASE ONE	FINAL
				stages of the evaluation.
Community Development	Local Knowledge Developed	Local knowledge is developed and respected.	n/a	n/a
	Community Trust	(See Principles.)	(See Principles.)	Evaluator works to build lasting trust by developing working relationships with a broad range of stakeholders through designing and implementing an evaluation, and providing credible findings.
Social Development	(Social Justice)	Social Justice: Enhance social justice.	Social Justice: Enhance social justice.	Increase Social Justice: Enhance social justice.
	Human Rights	Further human rights.	n/a	n/a
	(Social Action)	Social Action: Increase social action.	Social Action: Increase social action.	Increase Social Action: Increase social action.
Use of Evaluation	(Use In Decision-Making)	Use in Decision-Making: Stakeholders use evaluation findings in decision-making.	Use: Enhance utilization of the evaluation.	n/a
	Evaluation Is Valued	Enhance evaluation value.	n/a	n/a
	Evaluation Is Relevant	Enhance evaluation relevance.	n/a	n/a
	Evaluation Is Used	Enhance utilization of the evaluation.	n/a	n/a
Findings	Credible Findings	Evaluation findings are seen as credible.	Evaluation findings are seen as credible.	Participants see evaluation findings as credible.
	Timeliness	Evaluation findings are provided in a timely fashion.	n/a	n/a

Evolution of Identifying Statements		
DEPTH OF STAKEHOLDER INVOLVEMENT		
PRELIMINARY	AFTER PHASE ONE	FINAL
I do not make efforts to involve stakeholders in my		

Evolution of Identifying Statements		
evaluations.		
I make some efforts to involve stakeholders in my evaluations, even if it's just to develop the evaluation framework or questions.		
I involve stakeholders for a few key processes, like designing the framework and interpreting the findings.		
I involve stakeholders in as many ways as possible in my evaluations.	I always try to involve non-evaluator participants in my evaluations.	I always try to involve non-evaluator participants in my evaluations.
An evaluation would not be a success if it did not have stakeholder participation. (I do not take on an evaluation unless it has a strong participatory component.)	I prefer not to take on an evaluation unless it has a strong participatory component.	I prefer not to take on an evaluation unless it has a strong participatory component.
SELECTION OF STAKEHOLDERS		
PRELIMINARY	AFTER PHASE ONE	FINAL
The more stakeholder groups involved in evaluation, the better.		
Program beneficiaries should participate in carrying out evaluations.	Program beneficiaries should participate in carrying out evaluation.	Program beneficiaries should participate in carrying out evaluation.
Special interest groups should participate in carrying out evaluations.	People representing all important perspectives should be involved in any evaluation.	People representing all important perspectives should be involved in any evaluation.
People with a vital interest in programs (e.g., program developers, sponsors, directors) should participate in carrying out evaluations.		
People responsible for implementing or delivering programs should participate in carrying out evaluations.		Evaluators should help train all legitimate groups to do evaluation.
CONTROL OF THE EVALUATION		
PRELIMINARY	AFTER PHASE ONE	FINAL
Evaluators should share control of evaluation projects equally with practitioners.	Evaluators should share technical decision-making with non-evaluator participants.	Evaluators should share technical decision-making with non-evaluator participants.
Evaluators should relinquish control of evaluation projects to stakeholders.		
PHILOSOPHICAL REASONS FOR STAKEHOLDER INVOLVEMENT		
PRELIMINARY	AFTER PHASE ONE	FINAL
Evaluators should educate practitioners about the power and value of evaluation as a planned change strategy.		
Evaluation should help practitioners improve practice.		
Evaluation should stimulate practitioners to question		

Evolution of Identifying Statements		
fundamental beliefs and assumptions about practice.		
Evaluation should result in fundamental changes in practice.		
Evaluators should help train practitioners to do evaluations.	Evaluators should help train all legitimate groups to do evaluation.	
Practitioners' participation in evaluation should focus on enhancing the utilization of evaluation data.		
Practitioners' participation in evaluation should focus on bringing about social justice.	Evaluation should focus on bringing about social justice.	Evaluation should focus on bringing about social justice.
Practitioners' participation in evaluation should focus on bringing about individual-level change.	Evaluation should focus on bringing about individual empowerment, emancipation and self-determination.	Evaluation should focus on bringing about individual empowerment, emancipation and self-determination.