

Dissertation Proposal: Use of Research Tradition and Design in Program Evaluation:

An Explanatory Mixed Methods Study of Practitioners' Methodological Choices

Maggie Schultz Patel

Research Methods and Statistics, Morgridge College of Education

University of Denver

March 24, 2021

Table of Contents

Chapter 1: Introduction	4
Terms of Reference	5
Chapter 2: Literature Review	9
Use of Evaluation Methodology	9
Methodology Decisions in Evaluation	10
Chapter 3: Proposed Research Methodology Mixed Methods Research Design and Rationale for Design Selection	13
Research Question	14
Variables	14
Data Sources	14
Data Collection (processes) and Data Analysis (products)	14
Data Collection	15
Presentation of Results	19
Researcher Positionality	19
Ethical Considerations.....	19
Proposed Study Timeline	20

Abstract

The goal of the currently proposed explanatory sequential mixed method study is to assess whether there are observable trends in evaluation methodology by settings and content area in published evaluations from the past ten years (quantitative), illuminate how evaluation practitioners select these methodologies (qualitative), and assess how emergent findings from each phase fit together or help contextualize each other. In this study, methodology is operationalized as research tradition and method is operationalized as research design. For phase one, a systematic review of twelve peer-reviewed evaluation journals will be conducted and reported as frequency counts by tradition and design. These results will first be reported descriptively and then if appropriate, modeled inferentially with factorial repeated analysis of variance to assess for meaningful differences between groups (setting and content areas). For phase two (qualitative), interviews, which will be informed by the findings that emerge in the quantitative phase will be conducted with a purposive sample of practitioners to gain insight into how practitioners make methodological choices. In phase three (integration), findings will be integrated to contextualize emergent learnings from each phase.

Use of Research Traditions and Design in Program Evaluation:

An Explanatory Mixed Methods Study of Practitioners' Methodological Choices

Chapter 1: Introduction

As a still relatively new discipline, there is still a need, as some have pointed out, to increase self-knowledge in the field of program evaluation (e.g., Azzam, 2011). For instance, while debates on the merits of various methodological approaches are rampant and well-documented (e.g. Mertens & Hesse-Biber, 2013; Sechrest, Babcock, & Smith, 1993; Smith, 1994), how frequently each methodological approach is used in practice or how those approaches are selected is less well-documented. Similarly, in a survey of evaluators concerning what research on evaluation (RoE) questions they would most like answered, research on methods was one of the most frequently selected topics (Szanyi, Azzam, & Galen, 2013).

Problem Statement

There are very few, if any, systematic examinations of methodology and methods use or rationale for this use present in the literature. There is certainly conjecture in the literature about the way evaluation methods are selected, such as, “evaluators...have their favorite evaluation models and methods, usually those in which they were trained” (House, 1994, p. 241). This hypothesized practice of defaulting to favored models and methods as suggested by House, is problematic because each tradition and approach is designed to generate a particular type of evidence and answer a particular type of question. A mismatch between method and question would lead to limited utility, accuracy, and potentially validity of evaluation results. If the qualitative phase of the currently proposed study reveals that practitioners do in fact select

methods based on their preferences and comfort levels, this study will propel the field of evaluation forward by highlighting this weakness and prescribing improvements to practice.

Terms of Reference

A distinction should be made between *methods* and *methodology*. Both terms will be investigated throughout the course of this study. There are various characterizations of these terms in the literature. For example, in a discussion of how mixed methods have been defined across theorists, Creswell and Plano Clark distinguish methodology as “the process of research,” which they suggest includes underlying philosophy, methods, and interpretation of results, while method has been treated as the distinction of whether number or words will be the focus of data collection (p. 2-3, 2018). In this study, methodology refers to the research tradition (such as qualitative, quantitative, or mixed method), while method refers to the type of data collection and analysis strategies used (as discussed by Gliner & Morgan, 2000). In this text, *practitioner* refers to professional evaluators when they are conducting evaluations rather than conducting research on evaluation, developing evaluation theory, or teaching evaluation.

Rationale for Current Study

The currently proposed study will be conducted to interrogate whether there are observable group differences in the use of methodologies and methods (through a systematic review of published evaluations) as well as to explore what goes into the selection of these by evaluators (through semi-structured interviews with a protocol informed by learnings and questions that emerge from the quantitative phase). Findings from this study will contribute to the field of program evaluation practice in several ways. To begin with, this study will systematically collect evidence of which evaluation methods are most commonly used, as well as why. This

insight could be an important contribution to the field. Further, increased understanding of methodological trends over time as well as how methods are selected by practitioners is expected to contribute to improved practice. The theorized mechanism of change here is that as practitioners become more aware of their own habits and biases, they will gradually learn to select methods better-suited to the evaluation questions at hand. Further, this study is expected to illuminate the types of evidence privileged by the field as well as the types of organizations that tend to get evaluated. These insights will be useful not only to practitioners but also evaluation educators, evaluation clients, and professional evaluation associations. Understanding these insights may help these stakeholders in the field of evaluation advocate for more systematic, equitable, and pragmatic selection of methods. Additionally, this research will also be useful for theorists, as knowledge about practitioner decisions may provide fodder for future theory development. Awareness of these trends and practitioner rationale could encourage evaluation practitioners and commissioners to select their approaches more systematically and appropriately, given evaluation goals and program realities. Finally, findings from this research may generate recommendations related to new guidelines for credentials in program evaluation (such as for those being developed by the American Evaluation Association), particularly if findings suggest that practitioners are cherry-picking preferred methods rather than choosing those best suited to each evaluation.

The research questions that will be addressed by this study include:

1. Does practitioner use of evaluation methods and methodologies over the past ten years vary by setting or content area?

2. How do practitioners select evaluation methods and methodologies? What thought process do practitioners use to select methods and methodologies in light of practical considerations? What factors influence this process?
3. How do practitioners' explanations for how they select evaluation methods and methodologies thematically relate to observed group differences or similarities in practitioner use of evaluation methods and methodologies? How do these explanations contextualize observed differences or similarities?

The first question, which is quantitative, has been selected to address an established gap in the literature; there have been limited previous investigations of evaluation method and methodology use. Similarities and differences among groups will be assessed to account for contextual differences (year, setting, and content area). The goal for the quantitative phase of the research is to systematically assess group differences in methodologies and methods as indicated in published and peer-reviewed literature.

The second set of questions, which are qualitative, have been selected to explain the findings of the previous quantitative research question. The goal for this qualitative phase is to explore and contextualized emergent findings from the quantitative phase; specifically, to explore the factors that impact practitioners' methodological decision-making process and explain any observed group differences. Findings from the quantitative phase will be enhanced by practitioner perspectives, as the interview protocol will include questions meant to probe findings from the first phase. Research participants will be asked about quantitative trends that emerge.

The final, and mixed method set of questions, have been selected to integrate the findings of the quantitative and qualitative phases of this study. The goal for this phase is to contextualize

the quantitative and qualitative findings in light of each and ultimately, generate practice improvement recommendations to the field of evaluation.

Chapter 2: Literature Review

Use of Evaluation Methodology

Most of the literature available on use of methodology and methods in evaluation is non-empirical, consisting of reflections, editorials, and discussions (e.g., Norris, 2005; Smith, 1994; Stufflebeam, 2001). A typical example of this body of work is Norris' discussion of how important methodological choice is; Norris asserts that methodological creativity is superior to prescriptive approaches, which are of limited utility in a context-dependent field such as evaluation (2005). Another illustrative example is Stufflebeam's treatise on evaluation methods used in the 20th Century; in this piece, Stufflebeam categorizes which methods he feels are worth holding on to and not (2001). While useful to consider the opinions of often famed evaluators, there is little empirical basis for the assertions made in these types of articles.

There are, however, a few empirical examinations of trends in evaluation methodology evident in the literature (e.g., Christie & Nesbitt Fleischer, 2010; Galport & Galport, 2015). For example, following the scientific-based research movement that seemed to be taking off at the time, Christie & Nesbitt Fleischer (2010) conducted what they refer to as a content analysis of three evaluation-focused journals to determine whether there appeared to be a proliferation of randomized controlled trials. Ultimately, they found that non-experimental designs were used most frequently, followed by qualitative and mixed methods designs. This may be because practical realities dictate that the average program under evaluation is not ready for experimental study (e.g. due to lack of outcome evidence, data capacity, or newness). Another empirical study closely related to this proposed study concerns trends in Research on Evaluation (RoE) methods (Galport & Galport, 2015). Using a dataset of research on evaluation (RoE) articles published in

the *American Journal of Evaluation* from 1998-2014, the authors sought to categorize “methods-focused articles” to “uncover themes and trends in research on evaluation methodologies and techniques” (17). Most relevantly, they found nine themes related to why various evaluation methods were or should be chosen in these RoE articles, including: multiple units of analysis, maximizing data quality, determining evaluability, measuring fidelity, clarifying theories of change, an emphasis on low-cost or rapid results, a focus on qualitative or mixed methods approaches, and sampling concerns (24-25). The currently proposed study will build upon Galport & Galport’s research (2015) by cataloging methodologies used, and why they were used, in actual evaluation practice rather than in RoE.

Methodology Decisions in Evaluation

Similar to the previous research question, much of the published literature on the question of how methodology decisions are made consists of reflections from practice and editorials (e.g., Braverman & Arnold, 2008; Chelimsky, 1998; Chelimsky, 2007; Greene, Lipsey, Schwandt, Smith & Tharp, 2007; Kallemeyn, 2009; Schwandt, 2014, Smith, 1997, Spence & Lachlan, 2010) or prescriptive charges for how these decisions should be made (e.g., Braverman, 2012; Chelimsky, 2012; Mark, 2018; Maynard, Goldstein, & Nightingale, 2016; Julnes & Rog, 2007; Sechrest, Babcock, Smith, 1993).

An illustrative example of this subset of the literature comes from Chelimsky (2007), who states, “From an evaluator’s perspective, an a priori judgement about methods without a serious study of the context and specifics of a question is both unsuitable and imprudent in relation to likely evaluation success” (31). Another line of literature available on this topic is theoretical or prescriptive. For example, Kunding (2010) provides a framework for how to study evaluators’ decisions made in practice that emphasizes considering whether evaluators select

methodologies based on evaluation theory or if they use their own “practical knowledge,” consisting of assumptions, expertise, values, and judgement” (p. 347). Kundin’s suggested framework also includes considerations of evaluation context and real-time reflection based on changing environments. In a discussion piece published in 1994, Chen predicts that in the future, evaluation decisions will be made based on the specific evaluation question under study rather than a dogmatic attachment to quantitative or qualitative methods (Chen, 1994). Similarly, in the same year, House stated, “Originally only quantitative methods were deemed objective enough to be useful for evaluation, which followed beliefs then current in the social sciences...However, we have entered an ecumenical period in which qualitative techniques are seen as legitimate and mixed designs are recommended” (241).

These opinions and suggested frameworks, usually from venerated evaluation theorists or practitioners, are indicative of the type of literature that exists on this question. While this body of work is a useful starting point in documenting method use and how those choices are made or should be made, a more systematic assessment of method use along with further exploration of how those methods are selected would increase self-knowledge in the field.

Conversely, there are a handful of empirical and/or systematic examinations of how methods are selected by evaluators evident in the literature (Azzam, 2010; Azzam, 2011; Christie, 2003; Tourmen, 2009). For example, Azzam (2011) conducted a study that posed several evaluation questions to responding evaluators and asked them to propose designs. Azzam found that design choices are related to methods preferences and reported degree of focus on utility. Alternatively, there did not appear to be associations between design choice and evaluator gender, education, or level of stakeholder involvement in each evaluation scenario. While this study is an important step in pulling back the curtain on how evaluators select methods, this line

of research can be expanded on by examining evaluators' report of how they actually selected methods in previous experiences, rather than positing hypothetically how they might do so.

Similarly, Christie surveyed practitioners about whether they would use theory to inform method selection and found that only 10 percent reported their practice being informed by theory (2003). In a critique of this same research, Datta (2003) asserts that while what respondents purported to do was interesting, an even more useful task would be to review these respondents' evaluation reports to see what they actually do rather than what they say they do. This line of reasoning provides support for the current proposed study. While not exactly empirical, Datta (2007) attempted a somewhat systematic review of federal agency evaluation practice for the purpose of developing policies on method choice. Findings were based on a review of Governmental Accountability Office (GAO) reports, federal regulations, requests for proposals, grants, and reports, Evaltalk discussions (the American Evaluation Association discussion listserv), and personal experience. Ultimately, Datta found that different agencies tend to be inclined toward certain methods, while others are more versatile. These differences seem to be due to programs lending themselves more naturally to certain designs, agency culture preferring one kind of design over another, evaluator training and experience favoring certain methods, and the "politics of methodology" (44). The proposed study seeks to expand upon previous methods research in a more systematic and comprehensive manner.

The first line of research that emerged from this systematic review of the literature consisted of opinion-based essays or anecdotal reflections on method use and decision-making; this body of work may be instructive, but has limited generalizability or validity for the field. The second line of research that emerged from this review consisted of two past empirical attempts to systematically assess the use of methods in published evaluations and a few surveys

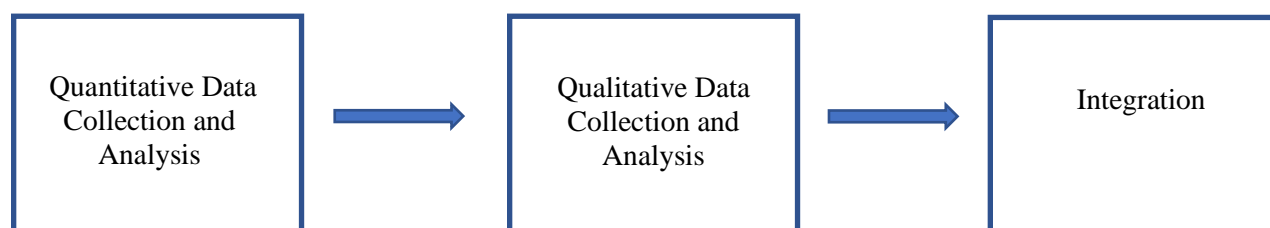
of practitioners about how they would hypothetically select evaluation methods. These studies, while more relevant to this currently proposed study, were limited by the following factors; one of these studies was conducted over a decade ago and could stand to be updated; the other was focused on methods used in research on evaluation (RoE) rather than evaluation per se; and the surveys concerned hypothetical situations rather than actual practice. In summary, while there is some literature on the topic of method use and decision-making in evaluation, there has been very limited research or empirical investigation on the topic. This means that there are very limited data available on this subject. The proposed study will generate empirical data and insights to build upon and expand these important foundations.

Chapter 3: Proposed Research Methodology Mixed Methods Research Design and Rationale for Design Selection

The currently proposed research will be a mixed methods study using an explanatory sequential design (Creswell & Plano Clark, 2018). The first phase, which will be quantitative, will be used to conduct a systematic review of published evaluations to address an observed gap in the literature, as there is limited past research conducting quantitative analysis of observed trends in evaluation methods and methodology. The second phase, which will be qualitative, will be informed by the findings of phase one. For example, persistent trends that emerge from the first phase will be explored in qualitative interviews with practitioners. The findings from phase two will be used to explain how and why practitioners choose various methods and methodologies. Finally, the integration phase will allow for the researcher to weave practitioner rationales for methodology choices together with the observed quantitative trends in practitioner

use (see Figure 1). This process is expected to generate unique insights into the authenticity of these observed trends.

Figure 1. Diagram of Sequential Exploratory Mixed Methods Research



The research questions that will be addressed by this study include (also summarized in Table 1):

1. Does practitioner use of evaluation methods and methodologies followed over the past ten years vary by setting or content area?
2. How do practitioners select evaluation methods and methodologies? What thought process do practitioners use to select methods and methodologies in light of practical considerations? What factors influence this process?
3. How do practitioners' explanations for how they select evaluation methods and methodologies thematically relate to observed group differences in practitioner use of evaluation methods and methodologies? How do these explanations contextualize observed differences or similarities?

Table 1. Research Matrix

Research Question	Variables	Data Sources	Data Collection (processes) and Data Analysis (products)
1. Does practitioner use of evaluation methods and methodologies followed over the past ten	Research tradition, research design, year of publication,	<i>American Journal of Evaluation, New Directions for Evaluation, Journal of Canadian Program Evaluation, Journal of</i>	systematic review (process) and database (product)

years vary by setting or content area?	content area, author setting	<i>MultiDisciplinary Evaluation, Evaluation and Program Planning, Evaluation: The International Journal of Theory, Research, and Practice, Studies in Educational Evaluation, Practical Assessment, Research and Evaluation, Evaluation & the Health Professions, African Evaluation Journal, Evaluation Journal of Australasia, Evaluation Review, Research Evaluation, Evaluation and Program Planning, Studies in Educational Evaluation, Educational Evaluation and Policy Analysis</i>	
2. How do practitioners select evaluation methodologies? What thought process do practitioners use to select methods and methodologies in light of practical considerations? What factors influence this process?	practitioner perspective, thought process, and identified contextual factors that contribute to method selection	purposively sampled practitioners	semistructured interview protocol (process) and interview transcripts (product)
3. How do practitioners' explanations for how they select evaluation methodologies thematically relate to observed trajectories in practitioner use of evaluation methodologies? How do these explanations contextualize observed trends?	to be determined (depends on results of first two phases of research)	data collected in phase one and two of study	integration of quantitative and qualitative inferences (process) and joint display linking themes from interviews (product)

Data Collection

In phase one (quantitative), data will be collected through a systematic review of the past ten years of issues of the American Journal of Evaluation, Journal of Canadian Program Evaluation, Journal of MultiDisciplinary Evaluation, Evaluation and Program Planning, Evaluation: The International Journal of Theory, Research, and Practice, Studies in Educational Evaluation, Practical Assessment, Research and Evaluation, Evaluation & the Health

Professions, African Evaluation Journal, Evaluation Journal of Australasia, Evaluation Review, Research Evaluation, Evaluation and Program Planning, New Directions in Evaluation Studies in Educational Evaluation, Educational Evaluation and Policy Analysis. These fifteen journals were selected based on the precedent of past research and because they are evaluation-focused, are published in English, and have at least ten years of issues available online (Coryn, Noakes, Westine, & Schroter, 2007). *New Directions for Evaluation* was excluded due to the concentration of reflective articles about topics in evaluation rather than articles featuring evaluations. While relying on the peer-reviewed literature will exclude a substantial portion of evaluation work (such as grey literature or unpublished but utilized work), this pragmatic approach will allow for a systematic approach.

In phase two (qualitative), data will be collected from 30 to 50 practitioners using a grounded theory design. Grounded theory is intended to generate a theory of a particular process grounded in the perspective of participants; this mirrors the purpose of this phase, which is to develop a theory of how practitioners select methodology that has naturalistically emerged from participant interviews (Creswell, 2000). The proposed sample size is based on the conventions of a grounded theory approach (i.e., capturing a theory that applies across participants, rather than an in-depth focus on the perspectives or experiences of a few participants. Participants will be recruited through the American Evaluation Association (AEA) listserv, professional contacts of the researcher, and snowball sampling from each. To increase the representativeness of the sample, a purposive sample of evaluation practitioners will be interviewed. To ensure a purposive sample, pre-interview demographic data will be collected (See Appendix A). To ensure representativeness, interviewees will be selected to represent a broad swath of evaluation practitioners in the United States in terms of details related to academic degree/credential,

practice setting, years of experience, and field of practice. Semi-structured interview questions will be posed in an open-ended fashion to allow for participants to comment without being influenced by the researcher's preconceived thoughts. Interview questions will include questions related to emergent findings from phase one, as well as questions about the considerations that go into selecting research tradition and design and how these considerations may be influenced by factors such as evaluator training, funders, evaluation purpose, or content area. As indicated by grounded theory, data will be collected and coded inductively until a coherent theory starts to emerge (Creswell, 2000). This theory will then be tested with participants and settled upon once saturation is reached. The end product will be a cohesive theory of how evaluators select evaluation methodologies across settings and content areas.

Phase three (mixed method) will not require any new data collection.

Data Analysis

In phase one (quantitative), data analysis will consist of first a descriptive analysis and then, if appropriate, inferential analysis. To begin with, the number of articles from each journal that met study criteria and were therefore included in the systematic review will be reported in a frequency table. Then, articles will be coded by year, research tradition, research design, practitioner setting, and content area. These codes will be quantified and operationalized into variable counts (e.g, number of articles using each type of research tradition); an example of how this would look is in Table 2 below. To increase the reliability of this coding process, an independent rater will be engaged to independently code a sample of articles. Interrater reliability will be assessed and reported; in the case of disagreement, a third party will be engaged to resolve the matter. Then, these variable counts will be analyzed with descriptive statistics and reported in a frequency table. Finally, if appropriate given the nature of data yielded in this

process (in terms of sample size and group balance), a factorial repeated measures analysis of variance will be conducted to explore group differences (setting and content area) in evaluation methodologies and methods over time (up to five time points consisting of two-year periods). This analytic approach is best suited to answering research questions exploring between and within group differences over time.

As recommended by previous researchers, the analysis of the data gathered in the systematic review will be implemented in a manner intended to maximize trustworthiness, including during the preparation phase, the organization phase, and reporting phase (e.g., Elo, Kaariainen, Kanste, Polkki, Utrainen, & Kyngas, 2014). This will include determining the utility of each category contained within each code, considering whether categories are truly distinct, determining the degree of interpretation involved in each categorization, and ensuring that categorizations accurately reflect the information provided by article authors.

Table 2.

Year	ResearchTraditi on_Quan	ResearchTraditi on_Qual	ResearchTraditi on_MM	ResearchDe n_Ex
2010	90	50	60	100
2011				
2012				
2013				
2014				
2015				
2016				
2017				
2018				

Qualitative data collected in the second phase will be inductively coded and analyzed for themes (Creswell, 2012). These data will be analyzed using Atlas.ti. For the third and final phase, these qualitative data will then be combined into pivot tables with the previously collected quantitative data; quantitative data will be dimensionalized by subgroup.

Presentation of Results

Results will be presented in a formal report. These findings will also be disseminated through conference presentations and a brief article to be submitted for publication within an *American Evaluation Association* journal.

Researcher Positionality

This researcher is an evaluator with over ten years of experience as a practitioner. Further, this researcher tends to subscribe to the philosophy that a multi-method or mixed-method approach is the most comprehensive. This researcher believes that there is often a mismatch between evaluation questions and methods used and evaluators should not shy away from using less familiar methods if they would best serve the evaluation questions under study. This perspective will likely influence the researcher's initial reaction to explanations of method choices, but will not affect final interpretations. Reflexivity journaling will be used to minimize this bias.

Ethical Considerations

Given that phase one will involve the analysis of secondary data, there are limited ethical concerns for this phase. In phase two, which will involve primary data, the rights of research participants will be protected through the use of Institutional Research Board (IRB) approval,

consent forms, secure data storage, and confidentiality. These processes will include COVID-19 protections and protocols.

Proposed Study Timeline

Following approval, this study is expected to begin in March 2021. Data collection will span March 2021 through August 2021. Data will be analyzed from June through September 2021. Final results are expected to emerge in November 2021.

References

- Azzam (2010). Evaluator responsiveness to stakeholders. *American Journal of Evaluation* 31(1), 45-65.
- Azzam, T. (2011). Evaluator characteristics and methodological choice. *American Journal of Evaluation*, 32(3), 376 - 391.
- Braverman, M. T., & Arnold, M. E. (2008). An evaluator's balancing act: Making decisions about methodological rigor. In M. T. Braverman, M. Engle, M. E. Arnold, & R. A. Rennekamp (Eds.), *Program evaluation in a complex organizational system: Lessons from Cooperative Extension. New Directions for Evaluation*, 120, 71–86.
- Braverman, M. (2012). Negotiating measurement: methodological and interpersonal considerations in the choice and interpretation of instruments. *American Journal of Evaluation*, 34(1), 99-114.
- Cordray, D. (1993). Synthesizing evidence and practice. President's address. *Evaluation Practice*, 14(1), 1992, 1-8.
- Chelimsky, E. The Role of Experience in Formulating Theories of Evaluation Practice. (1998). *American Journal of Evaluation*, 1998, 19(1), 35–55.
- Chelimsky, E. (2007). Factors influencing the choice of methods in federal evaluation practice. *New Directions in Evaluation*, 113, 13-33.
- Chelimsky (2012). Balancing theory and practice in the real world. *American Journal of Evaluation* 34(1) 91-98.
- Creswell, J. & Plano Clark, V. (2018). *Designing and Conducting Mixed Methods research* (third edition). Los Angeles: Sage Publications.

- Chen H. (1994). Current trends and future directions in program evaluation. *Evaluation Practice*, 15(3), 229 - 238.
- Coryn, C., Noakes, L., Westine, C., & Schroter, D. 2011. A systematic review of theory-driven evaluation practice from 1990 to 2009. *American Journal of Evaluation*, 32(2), 199-226.
- Christie, C. (2003). What guides evaluation? A study of how evaluation practice maps onto evaluation theory. *New Directions for Evaluation*, 97, 7-35.
- Christie, C. & Nesbitt Fleischer, D. (2010). Insight into evaluation practice: A content analysis of designs and methods used in evaluation studies published in North American evaluation-focused journals. *American Journal of Evaluation* 31(3) 326-346.
- Cresswell, J. (2000). *Qualitative Inquiry and Research Design: Choosing Among the Five Approaches* (first edition). Los Angeles: Sage Publications.
- Datta, L. (2003). Important questions, intriguing method, incomplete answers. *New Directions for Evaluation*, 97, 37-46.
- Datta, L. (2007). Looking at evidence: what variations in practice might indicate. *New Directions for Evaluation*, 113, 35-54.
- Elo, S., Kaariainen, M., Kanste, O., Polkki, T., Utrainen, K. & Kyngas, H. (2014). Qualitative content analysis: A focus on trustworthiness. *SAGE Open*, 4(1).
doi.org/10.1177/2158244014522633.
- Galport, M., & Galport, N. (2015). Methodological trends in research on evaluation. In Paul R. Brandon (Ed.), *Research on evaluation*. *New Directions for Evaluation*, 148, 17–29.
- Gliner, J. A., & Morgan, G. A. (2000). *Research methods in applied settings: An integrated*

- approach to design and analysis*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Greene, J., Lipsey, M., Schwandt, T., Smith, N., & Tharp, R. (2007). Method choice: Five discussant commentaries. *New Directions for Evaluation*, 113, 111-127.
- House, E. (1994). The future perfect of evaluation. *Evaluation: review of the past, preview of the future*. *Evaluation practice*, 15(3), 239-247.
- Julnes, G. & Rog, D. (2007). Current federal policies and controversies over methodology in evaluation. *New Directions for Evaluation*, 113, 1-12.
- Kallemeyn, L. (2009). Methodological changes and respecting stakeholder dignity. *American Journal of Evaluation*, 30(4) 575-580.
- Edmonds, W.E, & Kennedy, T.D. (2017). *An applied guide to research designs: Quantitative, qualitative, and mixed methods* (2nd ed). Sage Publications.
- Kundin, D. (2010). A conceptual framework for how evaluators make everyday practice decisions. *American Journal of Evaluation*, 31(3), 347 - 362.
- Spence, P. & Lachlan, K. (2010). Disasters, crises, and unique populations: suggestions for survey research. *New Direction for Evaluation*, 126, 95-106.
- Maas, C. & Hox, J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology*, 1(3), 86-92.
- Mark, M. (2018). Strengthening links between evaluation theory and practice, and more: Comments inspired by George Grob's 2017 Eleanor Chelimsky Forum Presentation. *American Journal of Evaluation*, 39(1) 133-139.
- Maynard, R., Goldstein, N., & Nightingale, D. S. (2016). Program and policy evaluations in

- practice: Highlights from the federal perspective. In L. R. Peck (Ed.), *Social experiments in practice: The what, why, when, where, and how of experimental design & analysis*. *New Directions for Evaluation*, 152, 109–135.
- Norris (2005). The politics of evaluation and the methodological imagination. *American Journal of Evaluation*, 26(4), 584-586.
- Sechrest, L., Babcock, J., Smith, B. (1993). An invitation to methodological pluralism. *Evaluation Practice*, 14(3), 227-235.
- Smith, M.F. (1994). Evaluation: review of the past, preview of the future. *Evaluation practice*, 15(3), 215-227.
- Schwandt, T. (2014). On the mutually informing relationship between practice and theory in evaluation. *American Journal of Evaluation*, 35(2) 231-236.
- Smith, N. (1997). Functions of the evaluation proposal in preordinate and emergent studies. *Evaluation Practice*, 18, (f), 1997, 17-24.
- Stufflebeam, D. (2001). Evaluation Models. *New Directions for Evaluation*, 89, 7-98.
- Stufflebeam, D. L. (2016). Factors that influenced my conduct of evaluations and evaluation training programs. In D. D. Williams (Ed.), *Seven North American evaluation pioneers*. *New Directions for Evaluation*, 150, 41–49.
- Tourmen, C. (2009). Evaluators' decision making: The relationship between theory, practice, and experience. *American Journal of Evaluation*, 30(1), 7-30.

Appendix A: Pre-Interview Screening Survey

1. How many years have you been practicing as an evaluator?
2. What is the primary setting in which you practice evaluation?
 - a. University
 - b. Private Research/Evaluation Firm
 - c. Nonprofit/ Community-Based Service Provider
 - d. Government
 - e. Other (Please fill in:_____)
3. What is the primary topic you evaluate (e.g., human services, public health, education, economics, etc.)?
4. What state are you practicing in?
5. What is your highest academic degree?
6. What subject is your degree in ?
7. Were there any field practice requirements in this degree program? If so, please describe:
8. Have you earned any post graduate credentials? If so, please describe:

