

Evaluator Exclusion in the Program Planning and Design Phase

by

Matthew Gallagher

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2022 by the
Graduate Supervisory Committee:

Jesse Lecy, Chair
Richard C. Knopf
Megha Budruk
Roseanne Schuster

ARIZONA STATE UNIVERSITY

May 2022

ABSTRACT

Program leadership's decision to include an evaluator during the program planning and design phase is the critical first step necessary for evaluators to provide the programmatic benefits associated with the evaluation profession. Several recent developments have promoted evaluator inclusion in program planning and design activities, including federal legislation that mandates evaluator inclusion and advocacy efforts from evaluation academics. However, the evaluation literature presents a collective frustration within the evaluation field due to ongoing exclusion from program planning and design activities. Utilizing the defensive attribution hypothesis, this quantitative study gathered responses from 260 American Evaluation Association members and 61 Project Management Institute members to determine an evaluator exclusion rate, develop a taxonomy of exclusion factors, and explore the extent to which program leaders and program evaluators demonstrate defensive attributions when rating these factors' influence on evaluator exclusion in program planning and design activities. Results indicated an approximately 70% evaluator exclusion rate in respondents' most recent program experiences. Furthermore, the defensive attribution hypothesis was not supported in the study, as program evaluators more strongly attributed their lack of inclusion to deficiencies outside of the evaluation practice, but program leaders also more strongly attributed evaluator exclusion to deficiencies outside of the evaluation practice. Program evaluators most strongly attributed their exclusion to program leaders' insufficient training and knowledge on the role of evaluation during the program planning and design phase. Program leaders most strongly attributed evaluator exclusion to their own staffing decisions, indicating a preference to not include evaluators in program

planning and design activities due to achieving previous program success without them, assigning evaluation activities to non-evaluation staff, and a funding process that allows the practice to occur. As the first study to explore evaluator exclusion in the program planning and design phase, it sets a foundation for future research studies to corroborate and build upon its findings, identify policies that encourage evaluator inclusion, and continue efforts to establish mutually beneficial relationships in the program planning and design phase.

Keywords: program evaluation, program leadership, program management, attribution theory, defensive attribution hypothesis, evaluator exclusion, American Evaluation Association, Project Management Institute, program planning, program design

ACKNOWLEDGMENTS

The academic necessity of completing this dissertation has been at once a daily pedestrian chore and a continually exhilarating endeavor.

It is upon myself which I place blame for the doldrums of the former.

For the peaks of the latter, it is necessary to focus some spotlights:

- To Naomi: I have a lifetime together with you to express my thanks. Let me start by saying thank you for being my daily support system, my sounding board when I've decided to verbally process the trials and tribulations of a doctoral student, and for your keen awareness of when I needed time and space to think and type.
- To Miriam: You have become my main motivation for everything I want to achieve. Your wonderment at the prospect of me becoming 'Dr. Daddy' have kept me going.
- To Drs. Knopf, Budruk, and Schuster: Each of you coaxed me to continue on this path, and provided me with the support, suggestions, and constructive feedback I needed to keep taking the necessary steps forward.
- And to Dr. Jesse Lecy: It was an absolute pleasure to simply sit and listen to your methods of critique and process of thought. I am grateful for your quick comprehension of my ideas and your always superior suggestions. Thank you for being a wonderful mentor.

TABLE OF CONTENTS

	Page
LIST OF TABLES	vii
LIST OF FIGURES	viii
CHAPTER	
1 INTRODUCTION	1
2 LITERATURE REVIEW	3
Definitions and Delimitations	3
Origins of Evaluation	6
Recent History of Evaluation	8
Expansion of the Program Evaluator Role in Evaluation.....	11
The Funding Opportunity Announcement and Request for Proposal.....	12
The Program Evaluator in the Program Planning and Design Phase.....	15
Advocacy for Evaluator Inclusion in the Program Planning and Design Phase..	15
Developmental Evaluation	19
Exploration of Reasons for Evaluator Exclusion and their Effects.....	21
Symbolic Evaluations	26
Attribution Theory	27
The Historical Roots and Models of Attribution Theory	28
Using Attribution Theory to Explore Evaluator Exclusion.....	33
Attribution Theory Use in Research	35
Attribution Theory in the Context of Evaluation Research.....	36
Purpose and Research Questions	42

CHAPTER	Page
3	METHODS 45
	Respondent Set from American Evaluation Association45
	Respondent Set from Project Management Institute46
	Questionnaires50
	Data Analysis.....55
4	RESULTS 58
	Questionnaire Response Rates58
	Respondent Demographics.....58
	Evaluator Exclusion Rate60
	Evaluator Exclusion – Program Type60
	Taxonomy of Evaluator Exclusion Factors61
	Analysis of Exclusion Factor Responses – Internal and External Factor Sets64
	Analysis of Exclusion Factor Responses – Individual Factors66
	Analysis of Exclusion Factor Responses – Program Evaluator Type.....68
	Analysis of Exclusion Factor Responses – Program Leader Type71
5	DISCUSSION 74
	Evaluator Exclusion Rate74
	Taxonomy76
	Evaluator Exclusion – Internal and External Factor Sets78
	Evaluator Exclusion – Individual Factors.....81
	Evaluator Exclusion – Individual Factors Summary95
6	CONCLUSION 98

CHAPTER	Page
Results and Findings.....	98
Value and Contribution	100
Limitations	104
Future Opportunities.....	105
REFERENCES	108
APPENDIX	
A QUESTIONNAIRE – PROGRAM EVALUATOR VERSION	118
B QUESTIONNAIRE – PROGRAM LEADER VERSION	128
C CHI-SQUARE TEST RESULTS	138
D TESTS OF NORMALITY	140
E PAIRED T-TEST RESULTS	142
F INDEPENDENT SAMPLE T-TEST RESULTS – EXCLUSION FACTORS ..	144
G INDEPENDENT SAMPLE T-TEST – EVALUATOR TYPE	149
H INDEPENDENT SAMPLE T-TEST – PROGRAM LEADER TYPE	154
I IRB APPROVAL LETTER	159

LIST OF TABLES

Table	Page
1. Respondent Demographics	59
2. Evaluator Exclusion Rate	60
3. Evaluator Exclusion Breakdown by Program Type	61
4. Taxonomy of Factors Influencing Evaluator Exclusion	63
5. Paired T-Test Results of Factor Sets by Respondent Group.....	66
6. Independent Sample T-Test Results for Exclusion Factors by Respondent Group	67
7. Independent Sample T-Test Results for Exclusion Factors by Evaluator Program Type.....	69
8. Independent Sample T-Test Results for Exclusion Factors by Program Leader Program Type.....	71

LIST OF FIGURES

Figure	Page
1. Attribution Model	30

CHAPTER 1

INTRODUCTION

Program leadership's decision to include an evaluator during the program planning and design phase is the critical first step necessary for evaluators to provide the benefits associated with the discipline of evaluation. However, the current evaluation literature presents a collective frustration within the evaluation field. The frustration has its roots in the general sentiment within the program management field that evaluation is a non-essential or back-end program activity. The current paradigm within the field of evaluation includes federal legislation that promotes the inclusion of evaluators in the program planning and design phase and evaluation academics and practitioners advocating for the same. However, evaluators are consistently excluded from the program planning and design phase due to a system that almost insists upon it: funding opportunity announcements that request inclusion of evaluation activities in proposed programs, but do not require evidence of any collaboration with a program evaluator, and program leaders with little to no training in or understanding of the benefits of evaluation during the planning and design phase. The current system produces programs that are plagued by rushed and symbolic evaluations that are of limited use, thereby frustrating program leaders and demoralizing evaluation practitioners.

To date, studies have examined approaches on how to incorporate an evaluator into a program's planning and design phase (Fitzpatrick, 1988) and multiple researchers have advocated on behalf of the inclusion of the evaluator throughout all the phases of the project cycle (Patton, 1978; Stufflebeam, 2001; Preskill & Torres 2001; Mark, 2012; Scheirer, 2012). However, no studies have examined the extent to which evaluator

exclusion occurs, or explored the reasons why a program leader decides whether or not to include an evaluator in a program's planning and design phase. As this decision is based on the antecedents of information, beliefs, and motivation, Heider's attribution theory provides an ideal foundation for exploring a program's pre-relationship environment between a program leader and evaluator. More specifically, the locus of control component of attribution theory provides a starting point for assessing whether a particular action should be attributed to internal characteristics or the external environment. Further, within the locus of control component, the defensive attribution hypothesis proposes that participants in a failure event tend to explain the failure in a way that minimizes their personal responsibility, by externalizing causality and offering alternative explanations in order to protect their self-esteem (Zuckerman, 1979). This study will measure the extent to which program leaders and program evaluators demonstrate defensive attributions when reflecting on the exclusion of the evaluator during the program planning and design phase.

CHAPTER 2

LITERATURE REVIEW

Definitions and Delimitations

Evaluation

The field of evaluation encompasses multiple schools of evaluation thought and practices, such as impact evaluation, participatory evaluation, systems evaluation, and performance evaluation (Grob, 2017). Chelimsky (1997) outlined three main purposes of evaluation, which are: 1) evaluation for accountability to funders and other stakeholders to ensure that funding is being used appropriately; 2) evaluation for program improvement by using data for managing a program; and 3) evaluation for causal knowledge about a program to generate strong evidence that the intervention independently causes the intended outcomes. Mark et al. (2000) offered four purposes of evaluation, of which two mirror Chelimsky's: 1) program oversight and compliance, which matches Chelimsky's accountability purpose; and 2) program and organizational improvement, which matches Chelimsky improvement purpose, but expands the idea to organizations as well. Mark et al. breaks Chelimsky's third purpose for evaluation, the generation of causal knowledge, into two parts: 3) assessing the merit and worth of a program; and 4) knowledge development.

For this paper, Scheirer's (2012) broad definition of evaluation will be adopted, which states that evaluation is the systematic collection and use of quantitative and qualitative data to aid program teams in developing knowledge about and managing a targeted set of program activities. It is a practice-based definition, focusing on the benefits that evaluation offers to program staff. Expanding on this point, Weiss (1972,

p.318) stated, “The basic rationale for evaluation is that it provides information for action. Its primary justification is that it contributes to the rationalization of decision-making. Unless it gains serious hearing when decisions are made, it fails in its major purpose.” Implicit in Weiss’s explanation is that the contribution an evaluator provides is determined by the strength of their relationship with program leadership.

Program Evaluator

The program evaluator can be defined simply as the person implementing the responsibilities of the evaluation profession. The American Evaluation Association defines the competencies a program evaluator must possess and practice. These include 49 competencies in the areas of professional practice, methods, program context, planning and management, and interpersonal interactions (AEA, 2018). The program evaluator’s role and responsibilities have expanded greatly in recent years, most notably since the passage of the Government Performance and Results Act in 1992, after a period of professional decline throughout the 1980s (Volkov, 2011; Wargo, 1994). This phenomenon will be explored more deeply in the ensuing sections.

Program

In Ward’s (2011) dictionary of program management terms, a project is defined as a temporary undertaking to create a unique product or service. A project has a defined start and endpoint as well as specific objectives that, when attained, signify completion. Ward differentiates between a project and a program, as a program is defined as a group of related projects managed in a coordinated way to obtain benefits not available from managing the projects individually, and may also include elements of on-going, operational work. It should be noted that not all definitions offered in the literature

make this distinction between a project and a program, and some authors have proposed definitions that invert the two terms. For example, Newcomer, Hatry, & Wholey (2004) define a program as the development and delivery of a specified set of activities intended to address one or more identified outcomes – a definition similar to Ward’s definition of project. This paper will adopt Ward’s broader definition and utilize the term ‘program.’

Program Leader

Despite setting firm boundaries around the differences between a project and a program, Ward (2011) sees overlap between the role and responsibilities of a project manager and program manager, as both need to monitor the achievement of planned objectives. However, Ward (2011) notes the program manager position requires more refined skills in business areas such as negotiation, organizational change management, financial management, consensus building, and political savvy, in addition to being responsible for overseeing all projects within a program. Some federal agencies involved in community development, such as USAID and Peace Corps, describe similar responsibilities using the title of Program Officer (USAID, 2015) and Program Manager (Peace Corps, 2020). This paper adopts the terms ‘program leaders’ and ‘program leadership’ to collectively represent the roles of program director, program officer, and program manager in programs. Overall, these terms position program leadership as the person (or people) who oversees the day-to-day implementation of a program, provides strategic direction for a program’s implementation, and with whom all final program based-decisions rest.

Project Cycle - Planning and Design Phase

Fifty years ago, Baum (1970, 1978) introduced the project cycle while working at the World Bank as a logical way to conceptualize and manage community development programs. It has since become standard practice for program staff to organize their programs using Baum's cycle (Biggs & Smith, 2003). In broad terms, the project cycle consists of progressive phases that lead from the planning and design of a program, through implementation of activities to achieve planned outcomes, to evaluation and reporting of the program's outcomes. Evaluation methods are useful during every phase of project development, implementation, and delivery (Scheirer, 2012). The central concept behind the use of a project cycle is one of use: the feedback, results, and findings from each phase will be utilized during the subsequent phases, including a completed program's next iteration (Patton, 1978; Biggs & Smith, 2003; Scheirer, 2012). The simplicity of the project cycle depicts an easy flow of information, data, and feedback across all program phases. In reality, this information flow requires committed coordination from all program staff. This paper, primarily concerned with the planning and design phase of a program, defines the phase as the timeframe when a program is being designed and developed for subsequent implementation.

Origins of Evaluation

Researchers trace the modern roots of evaluation back to the poverty-reduction programs enacted during President Lyndon Johnson's Great Society agenda in the 1960s (Alkin & King, 2016; Carman et al., 2008). Prior to implementation of the Great Society programs, tools to evaluate the effectiveness of federal programs and federal grants were much less prevalent and rigorous (Weiss, 1972). To study these new interventions, new

research and evaluation methods emerged from within the federal sector (Scherer, 2015), and led to a tenfold increase in government funding for program evaluations during the 1970s (Carman et al., 2008). During the 1970s and 1980s, these new evaluation tools and approaches, backed by financial support from federal agencies, expanded the evaluation components included in programs funded by private funders as well (Campbell, 1994). Private funders, including corporations and philanthropic foundations, followed federal efforts to measure grantee performance by incorporating evaluative activities into their requests for proposals (Mitchell, 2012; Thomson, 2010).

Ever since their advent, policymakers have debated and discussed the need to rigorously evaluate programs that annually consume billions of dollars (Coryn, Hattie, Scriven, & Hartmann, 2007; Alkin & King, 2016). Indeed, the evaluation discipline has endured responses ranging from ignorance to apathy to skepticism to derision. Rippey (1973, p. 9) provided one of the earliest reviews of federal program evaluations, commenting, “At the moment, there seems to be no evidence that evaluation, although the law of the land, contributes anything to educational practice other than headaches for the researcher, threats for the innovators, and depressing articles for journals devoted to evaluation.”

By the 1980s, the field of evaluation’s twenty-year foothold in the federal government was crumbling (Chelimsky, Cordray, & Datta, 1989). After tracking federal investment in program evaluation throughout the 1980s, Chelimsky, Cordray, & Datta (1989) declared the state of program evaluation to be in poor health, with large declines in funds allocated to evaluation as well as the overall number of evaluation staff. Freeman (1982) attributed this decline to an overall reduction of federal expenditures on

social programs, thereby reducing the need for evaluators at the federal level. For federal agencies, the relationship between evaluators and program leaders suffered, as program leaders lacked the information to improve or make informed decisions about their programs (Chelimsky, Cordray, & Datta, 1989; Freeman, 1982).

Recent History of Evaluation

In the 1990s, this situation changed. A new federal administration, which espoused a renewed belief in the capacity of government to address societal problems, found bi-partisan support with those on the other side of the aisle who held concerns about government accountability and performance (Wargo, 1994). Passed in 1993, the Government Performance and Results Act (GPRA) was designed to improve program management throughout the federal government. In the area of program evaluation specifically, the legislation mandated all federal agencies to submit annual strategic plans to Congress that contain “a description of the program evaluations used in establishing or revising general goals and objectives, with a schedule for future program evaluations” (Government Performance and Results Act, 1993, p.286). Additionally, the legislation mandated all federal agencies to submit annual performance plans to Congress: “Each report shall...evaluate the performance plan for the current fiscal year relative to the performance achieved toward the performance goals in the fiscal year covered by the report...and include the summary findings of those program evaluations completed during the fiscal year covered by the report” (Government Performance and Results Act, 1993, p.288-289). In spirit as well as practice, GPRA was the federal expression of calls from within the evaluation community for programs of all types to be more accountable for the outcomes committed to by those who had designed and funded the program, and

for the evaluation findings to, in turn, be utilized by program leaders to improve the current and next iteration of the program (Wargo, 1994; Wargo, 1995).

In 2010, Congress passed the Government Performance and Results Modernization Act, which updated and expanded the scope of GPRA. The new legislation created the positions of Chief Operating Officer and Performance Improvement Officer at every federal agency, and assigned them the responsibilities originally outlined in GPRA. The law also directed the Office of Personnel Management (OPM) to create a new position classification specifically for federal program evaluators. To do so, the law directed OPM to “identify the key skills and competencies needed by Federal Government personnel for developing goals, evaluating programs, and analyzing and using performance information for the purpose of improving Government efficiency and effectiveness” (GPRA Modernization Act, 2010, p.3882). The Modernization Act also required agencies to disclose information about the accuracy and validity of program performance data, to ensure program performance information is both useful and used in decision-making, and to demonstrate the commitment of agency leaders to program improvement efforts, as evidenced through their personal involvement in these efforts (GAO, 2011).

In 2019, efforts to improve federal program performance were further solidified through the signing of the Foundations for Evidence-Based Policymaking Act, known colloquially as the Evidence Act. The legislation mandated that every agency develop an evaluation plan “describing evaluation activities the agency plans to conduct...during the fiscal year following the year in which the performance plan is submitted” (Foundations for Evidence-Based Policymaking Act, 2019, p.5530). The legislation also mandated

every agency to designate an Evaluation Officer whose responsibilities were to “1) continually assess the coverage, quality, methods, consistency, effectiveness, independence, and balance of the portfolio of evaluations, policy research, and ongoing evaluation activities of the agency; 2) assess agency capacity to support the development and use of evaluation; and 3) establish and implement an agency evaluation policy” (ibid., p.5531). Additionally, the law directed policymakers and program planners at every federal agency to collect rigorous evidence and data prior to commencing any internal decision-making processes, and then use that information to craft policies, as well as design and fund programs, that will achieve expected results (The Evidence-Based Policy Commission Act, 2015).

Over the past 30 years, these three laws have breathed new life into the field of evaluation (Carman et al., 2008), and represent a profound change in the relationship between program leaders and program evaluators. Taken together, evaluation at the federal level now serves an instrumental role as both a management tool designed for accountability and as an integral provider of information to aid program leaders’ decision-making processes (Chouinard, 2013). Additionally, state, local, and private funders have seized upon the precedent that the federal government set through these laws by integrating more comprehensive accountability and evaluation requirements into their own grant programs (Carman, 2009; Newcomer, 1997). Indeed, over the past thirty years, state, local, and private funders have increased pressure on grantees to measure program results (Carman, 2009; Newcomer et al., 2004; Bozzo, 2000; Forbes, 1998), quantify a program’s financial and operational performance (Stone et al., 1999), and be

held accountable for achieving a program's intended outcomes (Stone & Cutcher-Gershenfeld, 2001).

Expansion of the Program Evaluator Role in Evaluation

In response to the legislative mandates of GPRA, GPRAMA, and the Evidence Act, program leaders began to broadly expand the role of the evaluator. The responsibilities of evaluators were no longer limited to simply assessing and reporting on the performance and progress of programs at their conclusion, but were broadened to include their input and expertise throughout all phases of the program: planning and design, implementation and monitoring, and evaluation and reporting (Wargo, 1994; Wargo, 1995; Howell & Yemane, 2006; Huberty, 1988). Volkov (2011) noted that a primary theme in the evaluation literature after the passage of GPRA was the broadening role of the evaluator, expanding from primarily a technical role to consultative and advisory roles. In 2010, the American Evaluation Association (AEA, 2010), the primary professional development organization for practitioners in the field of evaluation, fielded a survey of their membership to understand the breadth of responsibilities evaluators had been entrusted with. These included designing and implementing evaluations, building evaluation capacity, facilitating evaluations, managing internal evaluations, collecting and analyzing data, analyzing and judging programs, and maintaining monitoring and evaluation (M&E) systems (Volkov, 2011). Volkov (2011) took these survey responses and integrated them into a comprehensive analysis of literature that explicitly discussed the responsibilities of an evaluator. He then categorized his findings, which illustrate the multiplicity of roles the evaluator fills, as defined by researchers and program leadership alike:

- Change Agent
- Educator about Evaluation
- Practitioner
- Management Decision-making Supporter
- Consultant
- Researcher/Technician
- Advocate
- Organizational Learning Supporter

From these categories, it is clear that the role of the evaluator is no longer to simply conduct evaluations upon completion of a program, but to fill a variety of roles that spread across all program phases. However, while federal legislation has certainly led to the expansion of the evaluator role in both federally and privately-funded programs, the laws do not actively address the lack of collaboration between a program leader and evaluator *prior* to program funding; i.e., during the planning and design phase of a program proposal. To be fair, they were not meant to. In the next section, we will explore the federal funding opportunity announcement and request for proposal, which provides guidance to potential grantees in the timeframe immediately prior to funding awards.

The Funding Opportunity Announcement and Request for Proposal

The path to program funding in the U.S. begins with a Funding Opportunity Announcement (FOA) from a federal agency or a Request for Proposal (RFP) from a private funding organization. The FOA and RFP provide all the instructions needed for interested citizens to apply for grant funds. Cozzens (2000) notes that the federal funding system in the United States, unlike many other nations, has no single council providing operational oversight for its funding, but rather distributes funds via mission-oriented agencies through a peer and expert panel review process. Currently, 26 federal agencies

offer grant funding through thousands of programs (Grants.gov, 2020). As of this writing, 2,271 FOAs from these 26 agencies are currently open for the submission of proposals.¹ Of these, 28 are categorized as community development grants.

FOAs and RFPs request the inclusion of evaluation activities in proposal submissions at varying levels, and expect these evaluation activities to occur in order to fulfill the organization's reporting mandate². Responsibility for including and strengthening the evaluative components requested by FOAs and RFPs lies with the funding organization's leadership (Loo, 1985; Huberty, 1988). As this document is the catalyst for the future relationship between program leaders and the program evaluator, the evaluative language and requirements included are of critical importance. However, no meta-level research could be found on the strengths, weaknesses, or overall themes of the evaluative components included within FOAs or RFPs. This is possibly due to the sheer volume of FOAs and RFPs, as well as their dynamic and extremely devolved nature. A 2007 study by Coryn, Hattie, Scriven, and Hartmann comes closest, where the authors reviewed and appraised the national models for assessing federally-funded programs in sixteen countries, including the United States. The authors noticed that large-scale evaluations of government-financed programs were underway on a significant scale in nearly every major economy, and that the evaluation of government-financed programs had become a more prevalent method for increasing the quality of, and payoff from, the programs being implemented (Coryn, Hattie, Scriven, and Hartmann, 2007).

¹ For a look at the most current list, all open federal FOAs are available through the following link: <https://www.grants.gov/web/grants/search-grants.html>

² For federal agencies, program reporting is mandated by Congress. For private funders, program reporting is typically mandated by the organization's board.

Despite the lack of research, the academic literature does suggest what evaluative components should be included in FOAs and RFPs. Recommendations focus on activities meant to establish a collaborative foundation between the potential program leader and program evaluator at the time of proposal development. These include a program needs assessment, the explication or co-construction of a program's theory of change, and a literature-based assessment of the plausibility of key links in the program's logic model (Mark, 2012). In support of a program logic model, the literature advocates for FOAs and RFPs to require an evaluation plan, agreed upon by the potential program leader and evaluator, which specifies how program outcomes will be measured. According to Huberty (1988, p.25), "There is little doubt that a sound evaluation plan can enhance not only the likelihood of a proposed project being funded, but also the credibility of project outcomes." Overall, the inclusion of these items in a submitted program proposal offer evidence to the funding organization that collaboration between the program leader and evaluator has occurred during the program design phase. Finally, the literature recommends that the revision of the FOA/RFP language on evaluative activities, as well as the review of submitted proposals, should be the responsibility of employees with expertise and experience in evaluation. These staff have valuable insight on determining a program's need, internal logic, and are able to identify the basic concepts of evaluative thinking within program proposals (Grob, 2017). Furthermore, these staff can determine if a program's evaluation budget reflects the entirety of the responsibilities assigned to the evaluator (ibid., 2017).

The Program Evaluator in the Program Planning and Design Phase

The evaluation literature has established the unique skills that evaluators provide which benefit the early planning and design phase of a program. These include refining the program's design by conducting needs assessments (Fitzpatrick, 1988; Rossi & Freeman, 1993; Kane & Trochim, 2007; Carman, 2007), assessing a program's evaluability (Scheirer, 2012), and providing useful information and feedback for decisions on program structure (Stufflebeam, 2001; Cousins & Earl, 1995; Patton, 1997). The evaluation literature has also established that program evaluators' skills in designing an effective program structure lead to future cost-savings and increased program effectiveness. In their study of twelve federal evaluations, Howell and Yemane (2006) found that among the less successful programs, the development of evaluative activities during the program's design phase was incomplete. The incomplete design phase led to a waste of program resources due to false starts, after which impractical program designs needed to be changed in critical ways during implementation, thereby costing additional time and money. Furthermore, the evaluation efforts that result from this constricted timeframe and lack of collaboration between program leader and evaluator negatively influence the quality, timeliness, and usefulness of the collected data (Weiss, 1988; Cronbach & Associates, 1980).

Advocacy for Evaluator Inclusion in the Program Planning and Design Phase

The field of evaluation research, particularly in the area of advocacy for evaluator involvement throughout the project cycle, has experienced broad growth in recent decades. Within the literature, numerous researchers have advocated for the integration of an evaluator into the program planning and design phase. The pioneering work of

Stufflebeam (see: Guba and Stufflebeam, 1968; Stufflebeam, 1969) provided the earliest foundation for a utilization-focused theory, where evaluators were encouraged to link findings from the final evaluative phase of a program to the decisions made during the early planning and design phase of the next iteration of the program. Patton (1978, 1986) then expanded upon this theory, stating that it is the obligation of the evaluator to seek out potential users of evaluation findings as early as possible in the program, in order to ensure utilization of an evaluation's findings. The influence of this utilization-oriented evaluation approach reached the federal government mostly due to the work of Wholey (1981, 1983), who proposed a four-stage evaluation procedure called the 'sequential purchase of information' where evaluators would provide support to each phase of a program with the intention of improving program management and operations and influencing federal policy.

By the 1990s, there was general agreement among researchers and evaluation practitioners that the utilization-oriented approach to evaluation was not a single concept, but a multi-dimensional phenomenon characterized by the interaction of several dimensions: the instrumental (supporting decision-makers and solving programmatic problems), the conceptual (contributing to a learning-focused organization), and symbolic (serving a political function) (Shulha & Cousins, 1997). This general agreement spawned what is called Use Theory, which is a branch of the evaluation theory tree that groups "decision-oriented theories" together (Alkin & Christie, 2004). Decision-oriented evaluation theorists feel it is critical to conduct evaluations that are designed specifically to assist program leaders in program decision-making and to provide the information necessary for organizational change (ibid., 2004). Indeed, a multitude of evaluation

researchers and practitioners joined the chorus of utilization-oriented theory, as it wiped away the idea of the context-blind external evaluator in favor of injecting the evaluator directly into the program leadership team. Encouraged by this convergence, Smith (1994) stated that evaluation seemed destined to play a major role in the formulation of new programs and policies.

Use Theory advocates for evaluator influence in all program phases, with inclusion during the planning and design phase deemed the most significant development and, perhaps, as the literature will show, the most contentious. Fitzpatrick (1988) called for the expertise of evaluators to be utilized during the program planning and design phase to conduct needs assessments, assist in the development of program plans, and assess the overall evaluability of the program. Rossi and Freeman (1993) defined ‘innovative programs’ as those programs where the evaluator was included in the program planning and design phase to help identify goals and assist in the development of the program plan to achieve those goals. Stufflebeam (2001) updated his original theory, calling for evaluators to engage program leaders and stakeholders in all program phases to ensure that relevant information to assist in programmatic decision-making is produced and provided. Preskill and Torres (2001) proposed that evaluators should expand beyond the traditional boundaries between evaluator and program staff to create a seamless blend of program work, research, evaluation, and organizational development. Scheirer (2012) presented an evaluation roadmap, illustrating the variety of evaluative methods that are appropriate for the different phases of the project cycle. By doing so, Scheirer (2012) argued that evaluation can become a managerial function that is integrated throughout the project cycle, providing information for program planning,

decision-making during program implementation, and continuous program improvement. Looking to the future, Smith (1994) envisioned that program failures “could be lessened or even perhaps eliminated if evaluators become more involved in program development” (Smith, 1994, p.220).

What is clear from the literature is that evaluation researchers recognize the unique skillset of the program evaluator, and researchers have advocated convincingly for them to be provided an opportunity to apply these skills in the program planning and design phase (Preskill 2001; Scheirer 2012; Mark 2012). However, a large gap in the evaluation literature exists: the extent to which evaluators are included in a program’s planning and design phase is not known. Fitzpatrick’s (1988) single case study, which detailed her involvement as an evaluator in the planning and design phase of a program, provides some guidance on how evaluators and program leaders form a relationship in a program’s conceptual phase. However, Scheirer (2012) notes that little guidance in the form of practical guidelines, practices, activities, and descriptive case studies is currently available to program leaders, evaluators, and policymakers interested in incorporating evaluation into a program’s planning and design phase. Mark (2012) noted that the evaluation and program management fields need people who know when, where, why, and how different methods could and should be used during this phase. Huberty (1988) notes that this literature should not be solely directed towards evaluators, who have generally embraced the advocacy offered on their behalf, but towards program leaders, who still need to be convinced why evaluation activities need to be included in a program, beyond the reason of requirement or mandate.

Developmental Evaluation

Alkin and Christie (2004) organized all evaluation approaches into three theory branches: Methods, Valuing, and Use. All three theories are practice-based; they are meant to guide evaluators in their planning and implementation of data collection methods. However, all three theories are only applicable once a working relationship between a program leader and evaluator has already been established. This study addresses the space in time prior to this relationship, exploring the program leader's decision-making process to include or exclude an evaluator from a program's planning and design phase. All evaluation approaches and theories are rendered effectively useless if evaluators are not afforded the opportunity to apply them. Therefore, this pre-relationship environment needs exploration.

The practice of developmental evaluation comes closest to addressing this phenomenon by transforming advocacy for evaluator inclusion in the program planning and design phase into action. Through developmental evaluation, Patton (2011, p.vii) operationalized the expectation that "evaluators' involvement in program development would increase as the profession became recognized as having contributions to make at the front-end design stage of new programs based on general knowledge about patterns of effectiveness." Patton (1997, p. 106) crafted and operationalized developmental evaluation as a practice within the use theory branch of evaluation, where "the evaluator becomes part of the program design team or an organization's management team, not apart from the team . . . but fully participating in decisions and facilitating discussion."

Broadly, developmental evaluation establishes evaluative practices for programs operating in dynamic, novel environments with complex and changing interactions. It is

an approach focused on evaluator involvement in the early program planning and design activities, as it positions evaluation as a team function, integrated into program design and implementation activities and ongoing interpretive processes (Patton, 2006). The evaluator is “part of a design team whose members collaborate to conceptualize, design, and test new approaches in a long-term, ongoing process of development, adaptation, and intentional change” (Smith, 1994, p.220). The evaluator’s role in developmental evaluation is two-fold: 1) to elucidate program team discussions with evaluative data and logic; and 2) establish a feedback loop between the evaluator and program leader where the program leader is clearly stating the information needed from a program, and the evaluator is collecting high quality data, conducting rigorous data analysis, and feeding the program leader with findings and data-informed recommendations to improve program operations (Patton, 2011). The result is a collaborative and constructive approach that will positively influence the quality, timeliness, and utilization of program data.

Developmental evaluation is currently considered a niche evaluation approach. Patton (2011) notes that it is appropriate for times before a program model is in place and for dynamic programs where staff and funders expect to keep developing and adapting the program, never intending to conduct a summative evaluation. The advent of developmental evaluation is important because it provided evaluator inclusion advocates with guidance on building the relationship between program leader and evaluator during the program planning and design phase - earlier than any prior evaluation approaches had outlined. However, it does not explore the timeframe immediately preceding the establishment of this relationship, nor the information, beliefs, or attitudes held by

program leaders who must take the initiative to invite an evaluator into the program planning and design process.

Exploration of Reasons for Evaluator Exclusion and their Effects

To summarize, several developments over the past few decades have promoted the increased inclusion of program evaluators: 1) federal legislation has mandated the inclusion of evaluators in federal programs and led to the expansion of the evaluator role in both federally and privately-funded grant programs; 2) leading academics in the field of evaluation have led advocacy efforts towards the inclusion of evaluators in the planning and design phase of programs; and 3) the advent and rise of developmental evaluation from one of the evaluation field's most influential and respected theorists has moved the idea of evaluator inclusion in a program's planning and design phase from advocacy to action. In this section, we will explore the reasons that program leaders continue to exclude evaluators from the planning and design process despite these developments.

This study will develop a taxonomy of factors that influence evaluator exclusion during the program planning and design phase. From the literature, a loosely themed grouping of these factors can be attempted:

- 1. Program leaders lack the knowledge to include evaluators during the program planning and design phase.** Carman (2007) and Preskill (2014) noted that program leaders typically have little training in the benefits that evaluation can provide to programs. Rather, Carman (2007) continues, evaluation is typically covered in a single chapter in public administration and nonprofit management textbooks, or it is presented as the last stage in the policy-making process. Sanders (2003) pointed out

that evaluation is often not a required course in undergraduate or graduate education. Indeed, a brief review of Project Management Institute's³ basic PMP-training curriculum revealed a single fleeting reference to evaluation in Domain V, a section that focuses on processes for closing a program. Therefore, many program leaders enter into the program planning and design phase knowing very little about the practice and potential benefits of program evaluation. On this topic, Alkin and King (2017, p.435) noted, "Evaluation means different things to different people." In the case of program leaders, the idea of evaluator inclusion in the planning and design phase of programs may be an unknown concept.

It is unclear if this lack of knowledge is a recent phenomenon or not. During the mid-1960s, evaluation was a formalized part of the program planning and design process, where a needs assessment and evaluation plan were expected to be present in every funded program (Patton, 1978). But during the 1980s, as presented previously, the emphasis on evaluative thought and practice began to decline when federal funding was cut for many social services (Chelimsky, Cordray, and Datta, 1989; Carman et al., 2008).

- 2. Program leaders view evaluation as an end-of-program activity.** McLaughlin and Jordan (2004) and Rossi and Freeman (1993) noted that program leaders that have received formal training in evaluation typically only possess a working knowledge of the formal, outdated linear model of program evaluation, where evaluation is an activity that occurs at the end of the program. Similarly, Preskill and Torres (2001)

³ Project Management Institute (PMI) is the largest educator of project managers, with almost three million individuals having taken one of their courses and over 650,000 individuals possessing their Project Management Professional (PMP) certification (PMI, 2018).

and Scheirer (2012) forwarded the idea that there is a general sentiment within the program management field that all evaluation activities should be confined to the end of a program, i.e. during the evaluation phase of the project cycle. This sentiment would understandably lead program leaders to exclude evaluators from the planning and design phase of their programs.

This approach forces evaluators to forfeit their access to the program's planning, design, and implementation phases. This limits evaluators' opportunities to collect feedback on participants' program needs and baseline data, to gain a complete understanding of a program's intent, and to build collaborative relationships with the program team. Once provided access to the program through invitation by the program leader, evaluators will need to quickly jump into collecting data from program beneficiaries and participants. This is the unfortunate tradeoff that must occur between adequate evaluation planning and the need to quickly begin evaluative activities (Howell & Yemane, 2006). This constricted timeframe necessitates evaluators to be flexible and nimble to quickly design and collect data that will be immediately useful to the program manager (Scheirer, 2012). This urgency may lead evaluators to abandon critical evaluative activities, such as the collaborative development of a logic model and evaluation plan with the program team. Abandonment of this step, which typically provides evaluators with a method for understanding a program's intended outcomes, will produce inappropriate or misaligned data collection methods and tools. Furthermore, with little understanding on the part of the evaluators on what a program is intending to achieve, data collection and analysis efforts will ultimately yield little insight in the way of

actionable information to improve program operations or inform program decisions. The data produced from hasty and impractical data collection designs are ultimately only useful for fulfilling the requirement of submitting program progress reports to the program funder. All of these knock-on effects decrease the benefits that evaluators can offer to program leaders, thereby reducing the perceived quality and necessity of the evaluation altogether.

3. Program leaders actively exclude evaluators. To ensure the quality and use of evaluation findings resulting from data collection, the competency of the evaluator cannot be understated (Loo, 1985). For the evaluator, demonstrating proficiency in the basic skills of evaluation is of the utmost importance, as there is widespread concern among program leaders that many who take on the job of conducting an evaluation lack formal training or experience, resulting in evaluations that are poorly conceived, poorly executed, and poorly managed (Schwandt, 2015). Indeed, there is a long intellectual history of both normative and critical literature on the utility and benefits of program evaluation. For just as long, program leaders have been skeptical of the benefits of evaluation (Nature, 2006). Within the critical literature, themes tend to focus on the considerable expense necessary to establish and implement evaluative activities required by funders (Weiss, 1988) only for the information collected to be rarely utilized by program leaders (Cooke & Kothari, 2000; Biggs & Smith, 2003). Furthermore, few program leaders genuinely believe that the results of evaluations are actually used by funders to drive future funding decisions (Nature, 2006; Weiss, 1988).

Indeed, researchers have noted that current methods of evaluation are not always sufficient for the needs of program leaders (Coryn & Scriven, 2008). However, the possibility exists that these evaluation efforts are deemed inadequate because evaluators were not provided an opportunity to collaborate with program leaders during the planning and design phase to develop evaluative activities appropriate to the stated needs of program leaders. The data collection efforts that result from this lack of collaboration and constricted timeframe negatively influence the quality, timeliness, and usefulness of the collected data (Weiss, 1988; Cronbach & Associates, 1980).

Preskill (2014) also notes that there is a discipline-wide lack of understanding of the knowledge and skills necessary to be a professional evaluator. She called on the evaluation community to fill this research gap, as findings may explain why organizations, and more specifically, program leaders, may eschew evaluators from their programs. Fierro and Christie (2017) heeded this call, leading a study of how both evaluators and program leaders assessed evaluators' capacity, skills, and practices. Their results found that the evaluator was more likely to provide a lower rating of their own evaluation capacity than the program leader. However, they pointed out that these findings still allow for evaluators to systematically overestimate their own evaluation capacity. While not explicitly stated in Fierro and Christie's findings, this situation may lead a program leader to attribute their dissatisfaction with evaluation activities to evaluator ineffectiveness - and then applying this individual experience to their future decisions. Conversely, the program leader may lack the skills or motivation to appropriately apply the results, findings, and/or

recommendations provided to them by an evaluator. Producing high quality and rigorous data will add little value to programs unless program leaders possess the skills to interpret evaluation findings, think critically about how they affect program processes, and use data to make appropriate decisions about revising or adapting the program to better achieve planned outcomes (Patton, 2011; Sparrow, 1999). There is also a growing literature on the information overload experienced by program leaders as data has become more easily generated and available (Sparrow, 1999; Saxena & Lamest, 2018; O'Reilly, 1980). In some cases, this leads to program or organizational inertia where information and data, and the evaluators who produce it, are actively avoided (Saxena & Lamest, 2018).

Symbolic Evaluations

Taken together, the three areas of reasoning for excluding evaluators from the planning and design phase of programs explored above too often result in what Knorr (1977, 1980) termed 'symbolic evaluations.' Knorr presented symbolic evaluation as a situation where a program leader launched an evaluation of their program to signal that they were actively addressing a societal problem, but then simply postponed or ignored altogether the recommended actions resulting from the evaluation. Similarly, symbolic evaluations also occur when an evaluation is requested by program leaders because they are mandated, not because they are genuinely interested in the benefits that an evaluation could provide. These situations systematically restrict evaluator access to programs as the potential programmatic benefits of evaluation are never seriously desired, pursued, or known (Scheirer, 2012). As stated earlier, restricted access from the program planning and design phase effectively eliminates evaluators' ability to collaborate with program

leaders to logically structure a program's implementation plan, set practical and measurable benchmark indicators of program success, conduct a needs assessment, and collect potential participants' voice to inform a program's structure.

These hobbled attempts at program evaluation provide little added value to a program, leaving program leaders to perceive evaluation as a symbolic back-end activity that is necessary only to fulfill reporting requirements. This perception will be taken into the next program proposal phase, where a program leader will likely be faced once again with FOA (in the case of federally-funded programs) or RFP (in the case of all non-federally funded programs) language that requires little to no evidence of collaboration with an evaluator during the planning and design phase. As Bartle (2007) observed, evaluation deliverables have had little, if any, influence on these types of policy decisions. The cycle of symbolic evaluation is reinforced and continues unbroken, leaving fertile ground for frustration among funding organizations, program leaders, and program evaluators.

Attribution Theory

We have explored three broad themes on why program leaders would exclude evaluators from their programs. Some of these theme areas support each other: Program leaders may lack the knowledge to include evaluators in the program planning and design phase, while also viewing evaluation as a back-end program activity. Other possibilities are contradictory: Program leaders who actively exclude evaluators due to poor previous experiences at one point possessed the knowledge to include an evaluator in their program. To better understand the link between the situational antecedents to a relationship between a program leader and an evaluator, and the program leader's

decision whether or not to include an evaluator in a program's planning and design phase, attribution theory provides an ideal exploratory lens.

Attribution theory is the study of perceived causation. Wood (2008, p.51) provided a simple definition for the term attribution, stating that it is "the act of explaining why something happens or why a person acts a particular way." While many theories of attribution exist, the common foundation for all is that people interpret a behavior in terms of its causes, and these interpretations play an important role in determining our reactions to the behavior (Newcombe & Rutter, 1982). Furthermore, the process of attributing causes to a behavior is universal, pervasive, and predictable (Manusov & Spitzberg, 2008). Orth et al. (2012) provided a social science perspective on this important point, stating, "the central theme underlying attribution theories is that causal analysis is inherent in an individual's need to understand events." In other words, as humans, we have an innate need to understand why things happen. And the subsequent attributions that we inevitably make are the result of multiple influences, causes, and motivations. A critical component of attribution theory, which is also critical to exploring this study's research questions, is that it not only provides a foundation to investigate the causes of others' behaviors, but also allows others to investigate the causes for their own behaviors (Kelley and Michela, 1980).

The Historical Roots and Models of Attribution Theory

Most attribution theories arose in, and are most commonly researched as part of, the field of psychology. Heider provided the conceptual roots for attribution theory in his 1958 book, *The Psychology of Interpersonal Relations*. Heider (1958) described humans as naïve psychologists, believing they can explain any phenomena. Heider (1958, p.296)

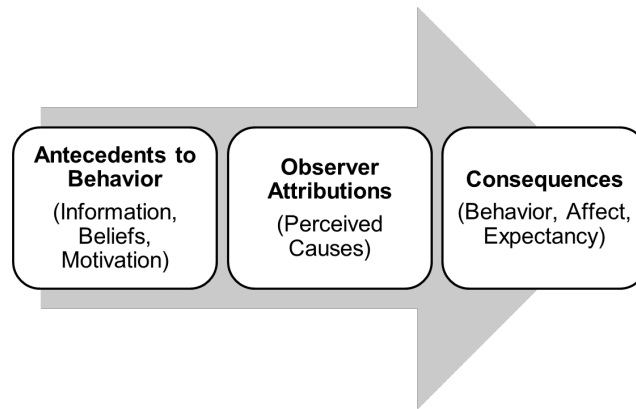
describes his contribution as a “pre-theory,” providing a transition from implicit theories based on intuitive thinking to the systematic thinking of a developed science. Heider’s main contribution to attribution theory was the concept of locus of control, which states that all behavior is considered to be determined by either internal (dispositional) factors or external (situational) forces.

As a concept, locus of control was originated by Rotter in 1954, where he noted that an individual should first examine the extent of the influence that they have over the observed outcome. An internal locus of control is identified when an outcome is significantly related to the individual’s own skills, abilities, and efforts. An external locus of control is identified when an outcome is determined to be independent of an individual’s own skills, abilities, and efforts (Rotter, 1966). Rotter uses the locus of control concept as a primary component of his social learning theory, which proposes that new behaviors can be acquired by observing and imitating others. After the publication of his theory, Rotter (1975) wrote extensively on misconceptions he observed with the usage of the locus of control concept. First, he noted that locus of control was not a typology, but instead represented a continuum of outcome attributions. Therefore, an outcome can be placed along a spectrum of control. Second, as personality results from an individual’s interactions with their environment, locus of control should be seen as a generalized expectancy of predictable behavior, but unique situations may arise where individuals may alter their behavior. It should be noted that even though the concept of locus of control is associated with both Heider and Rotter, and their descriptions are essentially identical, neither cited the other’s work in their own work. In general, it is

understood that Rotter originated the concept, but that Heider originated attribution theory utilizing the concept as a primary component.

Research studies utilizing attribution theory investigate the antecedents to a behavior, the attribution made for a behavior, and the consequences of the attribution that was made. This broad outline of attribution theory research is illustrated in Figure 1.

Figure 1. Attribution Model (adapted from Kelley and Michela, 1980)



Social scientists interested in the cognitive processes preceding behavioral attribution focus primarily on the link between antecedents to behavior and observer attributions portion of the attribution model seen in Figure 1. Harold Kelley's Covariation Model, which built upon Heider's initial contribution, is the most referenced theory for studies in this area. Kelley's model, first published in 1967, provided the foundation for analysis of the attribution process (Weiner, 2000). Kelley's model focuses on how people use information to make attributions for others', or for their own, behavior. It is a logical model that aims to understand the underlying structure that we use when we attribute causes of behaviors (Kelley and Michela, 1980). According to Kelley's model, people take into account three kinds of information, or evidence, before they assign causation to an observed behavior. The first is consistency, which answers our reflections on whether

the individual usually exhibits the observed behavior in a similar situation (Kelley, 1973). The observer is able to attribute a behavior to a cause when it is consistent over time and modality (Kelley and Michela, 1980). The second concept in Kelley's model is distinctiveness, which answers our reflections on whether the individual varies their behavior across different situations (Kelley, 1973). The observer is able to attribute a behavior to a certain cause when the behavior is distinctive from responses to other stimuli (Kelley and Michela, 1980). The third concept in Kelley's model is consensus, which answers our reflections on whether most individuals display a certain behavior in a certain situation (Kelley, 1973). The observer is able to attribute a behavior to a cause when either they have observed others to behave in a similar way in a situation, or if they observe others to behave in a situation that mirrors their own behavior in the same situation (Kelley and Michela, 1980). Together, these three types of information form the pattern of information that leads observers to attribute behavior to a single, or group of, causes.

Social scientists interested in behavior dynamics that are the consequences of our attributions primarily focus on the attributions-consequences link illustrated in the latter half of Figure 1. Bernard Weiner's attribution model, first presented in 1972, is the most referenced theory for studies in this area. The basic premise of Weiner's model is that individuals' causal explanations for their successes and failures will affect their future expectations, affect, and behavior in similar situations. Weiner (1986) posited that there are three main dimensions of our attributions that will affect future behavior. The first is locus of control, which is the degree to which a person perceives an occurrence to be caused by internal (also referred to as dispositional) or external (also referred to as

situational) factors (Weiner, 1972). As presented earlier, this dimension of Weiner's model was adopted from Heider's original theory. Weiner posited that the most important internal factors affecting locus of control attributions are the actor's ability and effort, while the most important external factors are task difficulty and luck (Forsyth & McMillan, 1981). The second dimension of Weiner's model is stability, which refers to whether the causes of a behavior or outcome are expected to change over time or situations (Weiner, 1979). Our perceived stability of the factors affecting an outcome will weigh into our future expectations for a similar outcome to occur. The third dimension of Weiner's model is controllability, and refers to the extent to which the cause of an outcome is under the control of the observer (Weiner, 1979). Weiner's original model included only the locus of control and stability dimensions, but he added the controllability dimension to distinguish between the causes of an actor's mood and effort (Forsyth & McMillan, 1981). Weiner (1979) granted that an actor's mood and effort are both internal and unstable causes, but that mood is less controllable than a person's level of effort. Therefore, he revised his model to incorporate the controllability dimension.

Overall, Kelley's Covariation Model and Weiner's Three-Dimensional Model have different areas of scope and application, which presents a problem for researchers interested in applying attribution theory to a study. However, the foundation for both models of attribution theory is Heider's original concept of locus of control, which will be the central component of attribution theory explored in this study. Further, the defensive attribution hypothesis, a hypothesis that utilizes the locus of control component to explain why participants in failure events minimize their personal responsibility, will be utilized to explore evaluator exclusion during the planning and design phase of

programs. The defensive attribution hypothesis will be discussed in more depth below, explaining its utility in the context of research into the field in evaluation in general and in the context of this study specifically.

Using Attribution Theory to Explore Program Evaluator Exclusion

With no prior research studies utilizing attribution theory to explore the program leader and program evaluator relationship, it is necessary to consider the strengths and weaknesses of attribution theory and how they may affect this study's research questions. To do so, the criteria posed by Braithwaite and Schrodt (2014) to assess theories (explanatory power, scope and generality, conditionship specification, and verifiability/falsifiability) will be utilized.

For explanatory power, attribution theories have the advantage of making good intuitive sense, as they were developed to account for laypersons as naïve scientists (Heider, 1958). Additionally, most of the dimensions and principles of attribution theories are immediately recognizable in our daily interpersonal interactions (Manusov and Spitzberg, 2008). Additionally, attribution theories embrace a logical, empirical view of the world (Manusov and Spitzberg, 2008). In other words, any behavior may be viewed as an effect that has some defined cause. Therefore, through the use of attribution theory in this study, an assumption is being made that program leaders make rational decisions on evaluator inclusion/exclusion based on antecedent information. However, as explored above, the literature has shown that some program leaders lack training on the benefits of evaluation and evaluator inclusion, which may indicate that a decision to exclude an evaluator from a program's planning and design phase is not an active rational

decision, but a passive decision based on ignorance. The methods utilized in this study will take steps to explore and take this assumption into account.

In the area of scope and generality, attribution theories were originally developed as a universal theory of human sense-making (Manusov and Spitzberg, 2008). Therefore, they are highly applicable to a wide breadth of phenomena. Reflecting on the theory that he contributed to, Weiner (2019) noted that attribution theory has enjoyed longevity in academic studies. Weiner attributed this longevity to the theory's ability to generate reliable data as well as its relational fertility - its ability to generate a range of predictions due to multiple interrelated constructs. Additionally, a similarity between attribution theory and the field of evaluation are their breadth of application. Both aim to address a wide range of issues using logical, pragmatic approaches (Patton, 2011; Mark et al., 2000).

In the area of conditionship specification, a theory must clearly articulate the nature of the relationship among its concepts (Braithwaite and Schrodt, 2014). In Heider's pre-theory, attributions are assigned to an internal or external locus of control. In Kelley's model, the informational antecedents to assigning causation to a behavior are limited to a behavior's distinctiveness, consistency, and consensus (Kelley, 1967). Newcombe and Rutter (1982) argue that Kelley's inclusion of the vague and ambiguous concepts of distinctiveness, consistency, and consensus in his model invariably result in study findings supporting some, or all, of the concepts. Manusov and Spitzberg (2008) point out that, almost 50 years after its inception, it is still not entirely clear how much the results from investigations of Kelley's model of attribution theory support his condition specifications. This points to the overall verifiability/falsifiability criterion of

Braithwaite and Schrodt's criteria for assessing attribution theory as a whole and Kelley's model in particular. Weiner (2000) noted that the vagueness and ambiguity of the concepts in Kelley's model may be one of the main reasons for its longevity.

Finally, any study utilizing attribution theory must address the problem of fundamental attribution error. Psychologists have determined that most people are biased in their judgment of who or what is responsible for an observed outcome (Ross, 1977; Newcombe & Rutter, 1982; Manusov & Spitzberg, 2008). This bias manifests when we tend to attribute the successes of others and our own failures to external factors, while we tend to attribute the failures of others and our own successes to internal factors.

Attribution Theory Use in Research

As noted above, Weiner reflected that attribution theory has enjoyed longevity in academic studies. However, he also lamented that attribution theory has lacked more widespread influence (Weiner, 2019).

As attribution theory focuses on interpersonal relations and perceptions, it found an immediate home in social psychology, clinical psychology, and cognitive psychology (Weiner, 2019). Martinko and Mackey (2019) noted that significant progress has been made more recently regarding the application and generalizability of attribution theory to organizational sciences. In this context, attribution theory has been used most often in studies on human resource management (see: Hewett et al., 2018; Sanders & Yang, 2016; and Martinko, 2004). More specifically, attribution theory has been applied to studies in the areas of personnel decisions (Carless & Waterworth, 2012; Struthers et al., 1998), organizational change (Chen & Wang, 2014), employee discipline (Cole, 2008), employee commitment and satisfaction (Fontinha et al., 2012; Tandung, 2016),

performance reviews (Levy et al., 1998), and work safety (Gonçalves et al., 2008; Gyekye, 2010).

In a 2000 article that reads more like a plea than a work of academia, Weiner (2000, p.383) wondered why studies on consumer behavior had not been more influenced by attributional thinking: “After all, consumers purchase products; these acquisitions meet with positive or negative end states; the consumer then reaches an attributional conclusion regarding why the outcome was ‘good’ or ‘poor’...and this conclusion surely influences their subsequent consumer-related behavior.” Heeding this call, researchers studying the topic of consumer behavior began to publish studies utilizing attribution theory (see: Mayo & Mallin, 2010; Pardo & Alfonso, 2017; Camilleri, 2017).

Studies utilizing attribution theory have struggled to gain a foothold in any other major topic areas. Martinko and Mackey (2019) note that many areas of inquiry utilizing attribution theory remain, particularly in studying the relationships between attributions and emotional processes. The authors write that there is still much to understand on the topics of how attributional processes influence emotions and the behaviors associated with those emotions in both organizations and individuals. Indeed, Weiner (2019) declared that if the influence of attribution theory is to match its longevity, researchers will need to reach out into new areas, expanding studies into new phenomena that attribution theory has yet to address.

Attribution Theory in the Context of Evaluation Research

Attribution theory has not yet been used as a foundation to study the pre-relationship environment between the program leader and program evaluator specifically, nor has it been used for research in the broader field of program evaluation. As a

practice-based discipline, the three main theories (Methods, Valuing, and Use) that currently guide the evaluation field are narrowly based on principles of how evaluative activities should be applied, albeit once a relationship between program leader and evaluator has already been established. Chelimsky (2013) noted that the practice of evaluation has little grounding in theory, and that current evaluation practice largely ignores theory. She explains this limited interaction as the result of the different interests of theorists and practitioners, with theorists focusing on methodological processes internal to evaluation and practitioners focusing on the application of these processes to an external environment. This study, therefore, explores the middle ground between theory and evaluation practice, and in so doing, offers a rare but practical link between the two. Additionally, this study will explore the pre-relationship environment between a program leader and evaluator, a situation for which there is a large gap in the academic literature.

Pragmatism and Sense-making

This study will adopt a pragmatic research paradigm. Pragmatism emerged as a method of inquiry for more practical-minded researchers, as it orients itself toward solving practical problems in the real world (Kaushik & Walsh, 2019). It is therefore an applied, problem-centered approach that is not committed to any one philosophy. In terms of methods, pragmatic approaches focus on the research problem and use all available methods to understand the problem being studied (Cresswell, 2014). Instead of focusing on methods, pragmatists emphasize the research problem as the most important determinant of research philosophy, and use all the approaches available to understand the problem (Cresswell, 2014). This study will utilize a survey approach, focusing on

questionnaires with quantitative components. This approach will be explored further in the Methods section.

In his seminal book on attribution, Heider (1958, p.5) noted, “scientific psychology has a good deal to learn from common-sense psychology.” In this study, evaluation, a pragmatic practice-based field, will be studied using the basic concepts of attribution theory, a psychology grounded in common sense. Attribution theory’s focus on sense-making finds a kindred spirit in the field of evaluation. Mark et al. (2000) refer to evaluation as ‘assisted sense-making,’ where the main responsibility of evaluators is to make sense of social policy and programs. Wye and Sonnichsen (1992) noted that the future and strength of evaluation, like pragmatism and attribution theory, is its protean nature. Evaluation adheres to a core conceptual framework and professional standards, yet is a flexible technique that can be adapted to a diverse array of organizational settings, client requirements, and program participant needs. This study, therefore, will utilize a pragmatic approach to studying the practice-based field of evaluation through the lens of the common-sense psychology of attribution theory.

Clarification and Organization of Terms

Heider (1958) described his early work on attribution theory as a pre-theory. He stated that his first attempt at exploring the psychology of interpersonal relations should not be taken as a full-grown system, as “it contains pre-theoretical speculations, hunches, and suggestions” (Heider, 1958, p.296). Additionally, Heider noted that he sought a clarification of the concepts and framework underlying interpersonal relations before a more robust theory could be constructed. Future theorists, most notably Kelley and Weiner, then built upon the foundation that Heider built. While attribution theory is a

critical component of this study, equally important to the study is the creation, clarification, and organization of terms used to explore the pre-relationship environment between a program leader and program evaluator. Doing so will: 1) offer a starting place for constructing the language to be used in future discussions of this pre-relationship environment, and; 2) provide a basis for analysis of the concepts and factors underlying the problem of evaluator exclusion.

In addition to the clarification and analysis of the terms and concepts underlying the pre-relationship environment between program leader and evaluator, this study will also organize these factors into a taxonomy utilizing the locus of control loci of internal and external characteristics. Attribution theory, and specifically its locus of control component, will therefore be tested to understand its utility in organizing these factors. Taxonomy is a formal system for classifying multifaceted, complex phenomena according to a set of common conceptual domains and dimensions (Patton, 2002). Taxonomies promote increased clarity in defining and comparing diverse, complex interventions (Sofaer, 1999), promote clear communication, and improve the measurement and evaluation of complex concepts and environments (Bradley, Curry, & Devers, 2007). All of these benefits of creating a taxonomy will lead to a better understanding of the pre-relationship environment between the program leader and program evaluator in the planning and design phase of programs.

Focus on Locus of Control in Failure Events – The Defensive Attribution Hypothesis

Weiner (2000) documented that individuals engage in the attributional process more often after a failure than after a success. As a situation that breaches federal legislation and dismisses the advocacy of evaluation's academics, the exclusion of the

evaluator from the program planning and design phase is a failure that needs exploration. In instances of failure events, the defensive attribution hypothesis has emerged from research grounded in the locus of control component of attribution theory.

The defensive attribution hypothesis states that participants in a failure event tend to explain the failure in a way that minimizes their personal responsibility (Shaver, 1970a; Gyekye & Salminen, 2004). To do so, participants externalize the event's causality and invoke alternative explanations in order to protect themselves and their self-esteem (Zuckerman, 1979). Therefore, studies on failure events that use attribution theory to ground the research describe defensive attribution as "a notion of self-protective attributional distortion through which people deny or minimize the implication of their own responsibility" (Gyekye & Salminen, 2004, p.2325).

The hypothesis originated in 1966 from a series of experiments designed by Walster (1966). Walster presented study participants with details about a hypothetical automobile accident and then asked them whether blame should be assigned and, if so, whether to assign blame to the perpetrator or the victim. Walster (1966) originally hypothesized, and found support through her experiments, that increasing the severity of the accident would increase the likelihood that study participants would assign attributions of responsibility. Shaver (1970a, 1970b) replicated Walster's experiments, but failed to replicate support for Walster's hypothesis that severity affected assignments of attribution. Shaver (1970a; 1970b) introduced different factors into her experiments, including age, personal/situational relevance and similarity, and possession of insurance to further test Walster's original hypothesis. The results of Shaver's experiments indicated support for the inclusion of personal and situational similarity in the defensive

attribution hypothesis, where participants who could envision themselves in future situations similar to those proposed in the experiment were more likely to attribute an accident to chance. In the participant's mind, doing so theoretically absolves them from blame should they experience a similar situation and outcome. Discussing these findings, Shaver (1970a, p.112) coins the term defensive attribution: "A common thread in the three experiments reported here is a tendency toward self-protection not unlike other motivated attributional errors, which may be characterized as instances of what we shall call defensive attribution."

In terms of methods used to test the defensive attribution hypothesis in its early days, Vidmar and Crinklaw (1974) noted that researchers had used inappropriate and inadequate experimental settings, resulting in no single version of the defensive attribution hypothesis receiving consistent support. Similarly, Burger (1981), in his meta-review of experiments testing the hypothesis, noted that an expansion of methods outside of hypothetical situations posed to undergraduates was necessary in order to more rigorously test and strengthen it.

More recent applications of the defensive attribution hypothesis are found in studies researching workplace accidents, the influence of personal and situational factors, and the assignation of responsibility (see Kouabenan et al., 2001; Gyekye & Salminen, 2004; Gyekye & Salminen, 2006; Gyekye & Salminen, 2007). As the methods used in the research are based on the actual experiences of workers and the perspectives of their employers, these studies directly address Vidmar and Crinklaw's (1974) and Burger's (1981) earlier criticisms. Additionally, these workplace-based studies introduced the variable of hierarchy into research on the defensive attribution hypothesis. According to

Kouabenan et al. (2001), this variable has simplified the hypothesis when applied to an accident occurring in the workplace, where there is a tendency for supervisors to attribute internal explanations that blame the employee for the accident, and for the employee to more likely attribute the accident to external factors. Indeed, Kouabenan (2001) notes that this position is supported by several previous studies that researched workplace accidents using the defensive attribution hypothesis (see: Dejoy, 1987; Hamilton, 1986; Lacroix & Dejoy, 1989; Mitchell & Wood, 1980; Niskanen, 1994; Salminen, 1992). Finally, a review of the methods sections of these studies reveals striking similarities: the researcher(s) pre-codes a set of factors into internal and external groupings, presents the list of factors to the supervisor and the employee(s) involved in an accident, and asks each to rate the influence of each factor on the resulting accident. Results are then compared in search of evidence of defensive attributions. These studies and their methods offer valuable guidelines for this study, and will be explored further in this paper's Methods section.

Purpose and Research Questions

To set a proper foundation for the purpose of this dissertation study, I have reviewed the academic literature on the historical roots, current state, and utility of evaluation in programs, as well as the historical roots, components, and current applications of attribution theory. Overall, the purpose of the study is four-fold: 1) to identify the extent to which evaluator exclusion from a program's planning and design phase occurs; 2) to identify a comprehensive set of factors that influence evaluator exclusion from the program planning and design phase; 3) to organize these factors into a taxonomy of internal (deficiencies within the evaluation practice) or external

(deficiencies outside the evaluation practice) locus of control; and 4) to explore the extent to which program leaders and program evaluators demonstrate defensive attributions when rating these factors' influence on evaluator exclusion from the program planning and design phase.

To achieve the study's purpose, the following research questions will be examined:

1. To what extent are program evaluators excluded from the program planning and design phase?
2. What is the comprehensive set of factors that influence evaluator exclusion from the program planning and design phase?
3. When asked to rate the influence of these factors on why evaluators were excluded from the program planning and design phase, do program leaders and program evaluators attribute evaluator exclusion to an internal (deficiencies within the evaluation practice) or external (deficiencies outside the evaluation practice) locus of control?
4. When making attributions about evaluator exclusion from the program planning and design phase, to what extent do program leaders and program evaluators demonstrate defensive attributions?

The results of this study will have profound implications for the fields of program management and program evaluation, their advocates, and their practitioners. This includes new insight into the extent that evaluators are excluded from the program planning and design phase. This also includes new knowledge from program leaders and program evaluators on whether they attribute evaluator exclusion from the program planning and design phase to shortcomings within the field of evaluation or to deficiencies outside the field of evaluation. This new insight and knowledge have the potential to spur new policy and changes in a variety of fields, including the training of program managers, the training of program evaluators, the language of federal agencies' funding opportunity announcements, the language of non-federal organizations' request

for proposals, and the future of evaluation literature on the topic of advocacy for evaluator inclusion in the program planning and design phase.

CHAPTER 3

METHODS

This study's research methods were built upon recent studies that have researched failure events using the defensive attribution hypothesis (see: Kouabenan et al., 2001; Gyekye & Salminen, 2004; Gyekye & Salminen, 2006). The research was conducted as an "altered replication" of these studies (Shaver, 1970b), and therefore utilized a similar quantitative survey approach, which will be described in detail below.

The research utilized two sets of respondents to answer the four research questions. The first set of respondents were current members of the American Evaluation Association (AEA). AEA is the primary professional association of evaluators in the United States, with approximately 6,000 members representing all 50 U.S. states and 80 countries. The second set of respondents were current members of the domestic-based chapters of the Program Management Institute (PMI). PMI is the world's leading project management training organization, and currently has over 600,000 members.

Respondent Set from American Evaluation Association

As a member of the American Evaluation Association since 2006, I was familiar with AEA's process for allowing its members to conduct research on evaluation topics using the organization's membership pool. On August 12, 2021, I contacted AEA's research unit in Washington, DC, and stated my interest to conduct my doctoral candidate research using current AEA members as the respondent pool. AEA replied positively, requesting that I complete an application. I completed the application, which consisted of a topic outline, IRB approval documents, all data collection instruments, and sampling criteria. Members of AEA's leadership reviewed my application. On September 29, 2021,

AEA approved my application to send the evaluator version of my questionnaire to their membership. They provided me with a random sample of 1,000 of their members' email addresses. Aside from my request that the sample include current members and not include members who were currently students or AEA interns, I had no control over the sample's characteristics or size. AEA provided me with instructions on how to send out the survey to their membership, including that it must be sent out within 30 days of approval, no more than three reminders may be sent to non-respondents, and that the respondent list must be purged from my files once data collection is completed.

In reviewing the sample, I found one blank email address and four duplicate email addresses, thereby reducing my sample to 995 AEA members. I sent out the program evaluator version of the questionnaire to all 995 of these AEA members on October 8, 2021. From this initial distribution, eight emails bounced, thereby reducing my sample to 987 AEA members. This number serves as the final sample pool for the program evaluator version of my questionnaire. Finally, I sent out follow-up reminders requesting non-respondents to complete the questionnaire on October 14, October 21, and October 28.

Respondent Set from Project Management Institute

To develop a sample of program leaders, I began my outreach by reaching out to the research unit at the global headquarters of PMI in Philadelphia, Pennsylvania on August 9, 2021. On August 20th, I was informed that PMI Global “does not collaborate on research”. In a follow-up request, I described the sampling process I was pursuing with AEA and asked if this was an acceptable approach for PMI Global. On August 21st, I was informed that PMI Global “does not do any questionnaires to their members.”

Therefore, I pivoted my approach to contacting local chapters of PMI located in the United States. PMI's local chapters are volunteer-run communities that allow program management professionals in a similar geographic area to connect with each other, develop their skills, and build their professional network. Initially, I limited my outreach to the PMI Phoenix and PMI Washington, DC chapters, as I am familiar with both of these areas having lived in both cities for an extended period of time. While I received no response from the PMI Washington, DC chapter, the PMI Phoenix chapter did agree to collaborate with me. On September 27, 2021, I presented my research topic to the chapter's Board of Directors and answered their questions, which mostly focused on ensuring the data security of members' contact information and the confidentiality of their responses. To distribute the program leader version of my questionnaire, I wrote an article for PMI Phoenix's monthly newsletter, and authored a series of messages that were distributed through their social media channels to their approximately 2,200 members. Unfortunately, these efforts yielded a low number of responses. It was therefore necessary to expand my outreach effort to all of the local PMI chapters.

Using the PMI website, I first put together a list of all the domestic-based chapters and the email addresses of their presidents. Overall, I found a total of 142 local PMI chapters in the United States. I then put together a form letter that stated the topic of my research and my interest in collaborating with the local chapter on my data collection. Over the course of three days (October 16, 22, and 25), I emailed the form letter to all the chapters, personalizing each one in an effort to garner more interest. Of the 142 chapters I contacted, 11 replied to me (8%). Five of these 11 responses stated that they were not interested in collaborating on the research opportunity. The remaining six (inclusive of

my earlier collaboration with PMI Phoenix) expressed interest in collaborating with me. I collaborated in various ways with the six chapters to gather responses to my questionnaire:

1. PMI Phoenix: The PMI Phoenix chapter has approximately 2,400 members. After board approval, I wrote an article for their monthly newsletter describing my research topic and the need for perspectives on the topic from program leaders. I also authored a series of short messages that were distributed through their social media channels. The article and social media messages each contained a direct link to my questionnaire.
2. PMI Delaware Valley: The PMI Delaware Valley chapter has approximately 3,500 members. After board approval, I wrote an email describing my research topic and the need for perspectives on the topic from program leaders. The email contained a direct link to my questionnaire. The communications director sent the email to the chapter's membership listserv.
3. PMI Kentuckiana: The PMI Kentuckiana chapter has approximately 800 members. After board approval, I gave a 10-minute virtual presentation on my research topic at the chapter's December 2021 luncheon. I provided a direct link to my questionnaire via the chat function in Zoom. I also authored a series of short messages that were distributed through their social media channels. Each social media message contained a direct link to my questionnaire.
4. PMI Buffalo: The PMI Buffalo chapter has approximately 740 members. The chapter's president put me in touch with their marketing director. Per their request, I authored a brief description of my research topic and the need for

- perspectives on the topic from program leaders. The marketing director posted this description along with a direct link to my questionnaire on the chapter's Facebook page.
5. PMI Houston: The PMI Houston chapter has approximately 4,500 members. The chapter's board requested a letter from my dissertation chair and the department chair showing support for my research. They also requested my study's IRB approval document. Upon receipt, the board approved my collaboration request. I wrote an article for their monthly newsletter describing my research topic and the need for perspectives on the topic from program leaders. I also authored a series of short messages that were distributed through their social media channels. The article and social media messages each contained a direct link to my questionnaire.
 6. PMI Inland Northwest Washington: The PMI Inland Northwest chapter has approximately 300 members. Per the chapter president's request, I authored a brief description of my research topic and the need for perspectives on the topic from program leaders. The president posted this description along with a direct link to my questionnaire on the chapter's LinkedIn site.

Overall, I initially intended to create a random sample of program leaders via a research collaboration with PMI Headquarters. I was then forced to recalibrate this approach by collaborating directly with local PMI chapters. I made this decision with an aim to still produce a purposive sample of PMI-affiliated program leaders. However, because only 6 of the 142 local PMI chapters agreed to collaborate with me, and because respondents then needed to take their own initiative to answer the questionnaire via

newsletters, social media posts, and presentations, the pool of program leader respondents should be categorized as a convenience sample.

Questionnaires

Both sets of respondents responded to similar questionnaires (see Appendices A and B), with only two main changes made in the wording of the questions: 1) According to the set of respondents, the titles ‘program leader’ or ‘program evaluator’ were used within the questionnaires; and 2) The questionnaire for program leaders asked respondents to indicate if they currently held a Project Management Professional certification, if they were currently a member of a local PMI chapter, and, if so, which chapter the respondent belonged to. I built, distributed, and collected responses to the two versions of the questionnaires using the Qualtrics survey software. Prior to distribution, I piloted the evaluator version of my questionnaire with four evaluation colleagues. I received useable feedback from two of them, and integrated their recommendations into the final versions of the questionnaire. I also piloted the program leader version of my questionnaire with two program management colleagues. Both colleagues reviewed and tested the questionnaire, and both colleagues responded that no further revisions were necessary.

Within the questionnaires, I outlined the following for the respondents prior to presenting any questions: 1) the study’s purpose; 2) the respondents’ role in the study; 3) the respondents’ expected time commitment; 4) the voluntary nature of study participation; 5) reasons why respondents should participate in the study; 6) the minimal risk that participation poses to respondents; and 7) contact information for the study’s Principal Investigators should respondents have any questions or concerns. After this

introductory study outline, I defined the following for respondents so they would have an understanding of the study's key concepts: 1) program; 2) program leader; 3) evaluation; 4) evaluator; 5) project cycle; and 6) program planning and design phase.

Next, I presented a set of questions to explore the study's first research question. Again, this dissertation's first research question aims to measure the extent to which program evaluators are excluded from the program planning and design phase. To establish the rate of program evaluator exclusion, I instructed both sets of respondents to reflect on their three most recent program experiences. I directed respondents with less than three program experiences to reflect on their one or two most recent experiences. To guide their responses, I presented respondents with guidance for interpreting the term 'excluded' for the remaining questions. The guidance stated that the term is inclusive of a program evaluator being actively excluded from a program's planning and design phase, a program evaluator simply being not included in the program planning and design phase, and/or a program evaluator being hired/contracted after the completion of the program planning and design phase. Additionally, I reminded respondents that a program's planning and design phase covers all activities that occur prior to a program's launch and implementation. I calculated a percentage exclusion rate based on responses from both groups. As the respondents' role and evaluators' exclusion status are both categorical variables, I also ran a Chi-square test to determine if there is a significant association between the two variables before the remaining data was analyzed.

Respondents from both groups who indicated that evaluator exclusion did not occur in any of their three most recent program experiences were filtered out from the remaining questions in the questionnaire. In other words, only the respondents who

indicated that evaluator exclusion occurred during the planning and design phase during any of their three most recent program experiences continued to respond to the questionnaire. To complete the data collection for the study's first research question, I asked each respondent who had indicated that evaluator exclusion did occur to identify the type of program in which it occurred. Respondents selected one of the following three choices:

1. In my three most recent experiences, I have only experienced being excluded from the planning and design phase of federal/federally-funded projects or programs.
2. In my three most recent experiences, I have only experienced being excluded from the planning and design phase of non-federal/non-federally-funded projects or programs (e.g. foundation-funded programs, non-profit programs, for-profit programs).
3. In my three most recent experiences, I have experienced being excluded from the planning and design phase of both federal/federally-funded AND non-federal/federally-funded projects or programs.

I then presented a set of questions to explore the study's second research question. Again, this dissertation's second research question aims to identify the comprehensive set of factors that influence evaluator exclusion from the program planning and design phase. For this question, I included respondents from both groups only if they had indicated that the evaluator had been excluded from the planning and design phase of at least one of their three most recent program experiences. In other words, program leader respondents

who indicated they had included a program evaluator in all three of their most recent program leadership experiences and program evaluator respondents who indicated they had been included in all three of their most recent program evaluation experiences, were filtered out from the remainder of the questionnaire.

I then presented this subset of program leader and program evaluator respondents with a list of 16 pre-coded factors and asked them to rate the influence of the factors on the decision to exclude an evaluator during one of their most recent experiences. I used a 5-point scale for respondents to rate each factor, ranging from 1 (no influence at all) to 5 (very influential). I generated the list of factors from the existing literature that described potential reasons why evaluators may be excluded from a program's planning and design phase. Utilizing the locus of control component of Heider's attribution theory, I coded all the factors as either internal (deficiencies within the evaluation practice) or external (deficiencies outside the evaluation practice). Factor codes were not visible to the respondents.

All factors were presented to each respondent in a randomized order. The list of five internal attributions I presented to respondents in the questionnaire were:

1. Evaluator too expensive to include in budget
2. Program leadership skeptical of the benefits of evaluation
3. Inadequate skillset of evaluators to provide benefits during the planning and design phase of programs
4. Program leadership had negative previous experience with evaluators
5. Evaluator may discover negative program results

The list of eleven external attributions I presented to respondents in the questionnaire were:

6. Organization funding the program did not require evaluation activities
7. Program leadership had no knowledge of or training on the benefits of evaluation
8. Program leadership had belief/understanding that evaluation only needs to occur at the end of a project/program
9. Program leadership preferred to have program staff (who are not evaluators) manage evaluation activities
10. Previous programs/projects had been successful without an evaluator, so there was no need to change
11. Program's hiring process too difficult/lengthy to consider hiring an evaluator
12. Program leadership followed a supervisor's (or someone else higher up) decision to not hire an evaluator
13. Adding/Managing an evaluator would have added to an already heavy workload
14. Evaluation findings would not be used in program decision-making, marketing, and/or funding decisions
15. Inadequate skillset of program leadership to interpret and/or apply evaluation results/findings
16. Paperwork Reduction Act (PRA) processes mandated by Office of Management and Budget (OMB) too burdensome for designing/implementing evaluations

Additionally, in order to compile a comprehensive taxonomy of internal and external factors, I provided respondents with the opportunity to add up to two of their own factors and rate their influence. For any factors added by respondents, I conducted a qualitative

analysis by coding and theming the additions. I then reviewed these themes for potential addition to the existing set of sixteen factors.

Finally, I directed respondents to provide their demographic information, including gender, highest level of education, and total number of years of experience in their field.

Data Analysis

The approach to analyzing my study's first and second research question were described in the previous section. To answer the third and fourth research questions, I applied four sets of analyses to the ratings provided for each factor by both respondent groups:

1. Using SPSS, I created a new variable that calculated a mean score for each group's responses to the internal set of five exclusion factors. I created another new variable that calculated a mean score for each group's responses to the external set of thirteen exclusion factors. I then applied two paired t-tests: 1) the first paired t-test compared program evaluators' mean scores of the internal set of factors against their mean scores for the external set of factors; and 2) the second paired t-test compared the program leaders' mean scores of the internal set of factors against their mean scores for the external set of factors. A comparison of the means, and whether the difference is statistically significant, indicated whether each respondent group attributes evaluator exclusion during the program planning and design phase to internal factors (deficiencies within the evaluation practice) or to external factors (deficiencies outside the evaluation practice).
2. Using SPSS, I completed a factor-by-factor comparison of the program leaders' and program evaluators' causal attributions of internal and external factors by

- administering independent sample t-tests to calculate the mean scores for each individual factor and assessing the statistical significance of any difference between groups for each individual factor.
3. Using SPSS, I looked deeper into the factor ratings provided by program evaluators in two different types of programs: 1) federal/federally-funded programs; and 2) non-federal/non-federally funded programs. I completed this analysis to determine if there were any significant differences in the factor ratings given by the two groups of evaluators. To do so, I completed an independent samples t-test to determine the means and statistical significance of the individual factor ratings each evaluator group provided. To complete the analysis, I created a new variable to separate program evaluators into two new groups: 1) A group of program evaluators who indicated they had been excluded from the planning and design phase of non-federal/non-federally funded programs was retained with no changes; and 2) A group of program evaluators who indicated they had been excluded from the planning and design phase of federal/federally-funded programs was merged with the group of program evaluators who indicated they had been excluded from the planning and design phase of both federal/federally-funded and non-federal/non-federally-funded programs. This merge was able to be made as respondents who indicated that evaluator exclusion had occurred in more than one program type were directed to focus on their experience in a federal/federally-funded program when making their factor ratings.
 4. Similar to analysis set #3, I used SPSS to look deeper into the factor ratings provided by program leaders in two different types of programs: 1)

federal/federally-funded programs; and 2) non-federal/non-federally funded programs. I completed this analysis to determine if there were any significant differences in the factor ratings given by the two groups of program leaders. To do so, I followed the same steps detailed in analysis set #3, but using the individual factor ratings provided by program leaders.

CHAPTER 4

RESULTS

The results from this study are presented in the following nine sub-sections:

1. Questionnaire Response Rates
2. Respondent Demographics
3. Evaluator Exclusion Rate
4. Evaluator Exclusion - Program Type
5. Taxonomy of Evaluator Exclusion Factors
6. Analysis of Exclusion Factor Responses – Internal and External Factor Sets
7. Analysis of Exclusion Factor Responses – Individual Factors
8. Analysis of Exclusion Factor Responses – Program Evaluator Type
9. Analysis of Exclusion Factor Responses – Program Leader Type

Questionnaire Response Rates

AEA's random sample of 1,000 members' email addresses was reduced to 987 due to 4 duplicate email addresses, 1 blank email address, and 8 bounced emails. From the sample of 987 AEA members, 292 members opened the program evaluator version of the questionnaire (29.6%). Of those, 32 members did not respond. Overall, from the sample of 987 AEA members, there are a total of 260 responses that can be analyzed, for a response rate of 26.3%.

After a broad variety of outreach measures were used, a total of 97 PMI members opened the program leader version of the questionnaire. Of those, 36 members did not respond. Overall, there are a total of 61 responses that can be analyzed. As the pool of PMI respondents are the result of a convenience sample, no response rate can be calculated.

Respondent Demographics

For all demographic questions, respondents who indicated that their programs had included program evaluators in the planning and design phase were filtered out of the

questionnaires. Therefore, demographic question were only presented to respondents who indicated that their programs had excluded program evaluators from the planning and design phase. This approach aligns with the study’s main purpose: to understand the extent of, and factors contributing to, evaluator exclusion in the program planning and design phase.

Table 1 presents the respondents’ demographics. For gender, there was an approximately 2-to-1 ratio for female to male respondents for both groups. For highest level of education achieved, program evaluators were more highly educated, with approximately half achieving Doctorate degrees and half achieving Master’s degrees. For program leaders, approximately three-quarters achieved a Master’s degree. For years of experience in their respective fields, there was good representation across all experience categories.

Table 1: Respondent Demographics		
Gender	Program Evaluator (n=135)	Program Leader (n=35)
Male	31.3%	37.1%
Female	66.4%	60.0%
Non-binary / Third gender	2.3%	0.0%
Prefer not to say	0.0%	2.9%
Highest Level of Education	Program Evaluator (n=128)	Program Leader (n=35)
Secondary School	0.0%	0.0%
Associate's Degree	0.0%	5.7%
Bachelor's Degree	2.3%	11.4%
Master's Degree	45.3%	74.3%
Doctorate Degree	52.3%	8.6%
Years of Experience	Program Evaluator (n=128)	Program Leader (n=33)
0 - 5 years	16.4%	21.2%
6 - 10 years	24.2%	33.3%
11 - 15 years	18.8%	15.2%
16 - 20 years	15.6%	15.2%
Over 20 years	25.0%	15.2%

Evaluator Exclusion Rate

Table 2 presents the rate of evaluator exclusion from a program’s planning and design phase in respondents’ three most recent program experiences. For program evaluators, 71.9% were excluded in at least one of their experiences. For program leaders, 68.9% had excluded a program evaluator in at least one of their experiences.

Table 2: Evaluator Exclusion Rate		
	Program Evaluator (n=260)	Program Leader (n=61)
Yes, excluded	71.9%	68.9%
No, not excluded	23.8%	26.2%
Unsure	4.2%	4.9%

A chi-square test tested whether respondents’ perspectives on a program evaluator’s exclusion from the program planning and design phase is associated with whether the respondent was a program evaluator or program leader⁴. In all cases, the expected counts were quite close to the observed counts. The chi-square value was 0.179, which gives a p-value of 0.672. This p-value is not significant, indicating that a respondents’ indication of whether a program evaluator was excluded from the program planning and design phase is not associated with whether the respondent was a program evaluator or program leader.

Evaluator Exclusion - Program Type

Table 3 presents a breakdown of the cases of evaluator exclusion as reported by both respondent groups from the three different program types. Of the cases of evaluator exclusion cases reported, approximately 30% of cases were reported by those in non-

⁴ Chi-square test result tables are available in Appendix C.

federal or non-federally-funded programs. Approximately half of the evaluator exclusion cases were reported by respondent groups from a combination of both program types, with reported cases slightly higher among program leaders (55% to 43%). For cases of evaluator exclusion in federal or federal-funded programs, the percentage reported was slightly higher for program evaluators (27% to 15%).

Table 3: Evaluator Exclusion Breakdown by Program Type		
	Program Evaluator (n=174)	Program Leader (n=33)
Federal / Federally-funded program	27.0%	15.2%
Non-federal / Non-federally funded program	30.5%	30.3%
Both program types	42.5%	54.5%
TOTALS:	100.0%	100.0%

Taxonomy of Evaluator Exclusion Factors

When rating the influence of each factor on the exclusion of the evaluator from the planning and design phase, 43 program evaluator respondents added and rated 60 more exclusion factors, and 6 program leader respondents added and rated 8 more exclusion factors. These 68 additional exclusion factors were reviewed and coded for potential inclusion in the final taxonomy, with following results:

- For 39 of the additions, the respondent rephrased an already existing exclusion factor. Therefore, the addition was not integrated into the existing taxonomy. However, based on the review, two of the existing exclusion factors were slightly revised to improve clarity:
 - The exclusion factor ‘Program leadership had no knowledge of or training on the benefits of evaluation’ was revised to ‘Program

leadership had no knowledge of or training on the benefits of evaluation in the planning and design phase’

- The exclusion factor ‘Program leadership preferred to have program staff (who are not evaluators) manage evaluation activities’ has been revised to ‘Program leadership had program staff (who are not evaluators) plan and/or manage evaluation activities’
- For 13 of the additions, the respondent did not present the addition as an exclusion factor. In other words, the addition did not provide a reasoning for why the evaluator was excluded from the program planning and design phase.
- For 3 of the additions, the respondent provided an added exclusion factor that was not generalizable. In all three cases, the factor was unique to the situation described.
- There were 8 additions that noted the short turnaround time that exists between when a funding organization issues either a Federal Opportunity Announcement or a Request for Proposal and the due date when a completed proposal must be submitted. Respondents explained that this short turnaround time precludes them from including a program evaluator during the planning and design of the program. This exclusion factor was added to the taxonomy as a factor external to the field of evaluation.
- There were 5 additions that noted that an external evaluator cannot be paid for any work they contribute to the planning and design phase of a

program until the program actually receives funding. This factor is unique to organizations that do not have any internal evaluators on staff and must rely on outside evaluation expertise. However, for the organizations that fit this description, their decision on whether to include an evaluator during program planning and design is driven by this reasoning.

Therefore, this exclusion factor was added to the taxonomy as a factor external to the field of evaluation.

When adding these two additional factors to the original set of sixteen factors, a comprehensive set of eighteen factors influence whether an evaluator is excluded from a program’s planning and design phase. Using the attribution theory’s internal and external loci of control, the factors can be organized into a taxonomy. This taxonomy, presented in Table 4, serves two main functions for this study. First, it promotes increased clarity and understanding when exploring, discussing, and comparing the factors contributing to program evaluator exclusion. Second, it provides a foundation for testing the defensive attribution hypothesis as it applies to the failure event of evaluator exclusion during the program planning and design phase.

Table 4. Taxonomy of Factors Influencing Evaluator Exclusion during the Program Planning and Design Phase	
Factors Internal to the Evaluation Field	
1.	Evaluator too expensive to include in budget
2.	Program leadership skeptical of the benefits of evaluation
3.	Inadequate skillset of evaluators to provide benefits during the program planning and design phase
4.	Program leadership had negative previous experience with evaluators
5.	Evaluator may discover negative program results
Factors External to the Evaluation Field	
6.	Organization funding the program did not require evaluation activities

7. Program leadership had no knowledge of or training on the benefits of evaluation in the planning and design phase
8. Program leadership had belief/understanding that evaluation only needs to occur at the end of a program
9. Program leadership had program staff (who are not evaluators) plan and/or manage evaluation activities
10. Previous programs had been successful without an evaluator, so there was no need to change
11. Program's hiring process too difficult/lengthy to consider hiring an evaluator
12. Program leadership followed a supervisor's (or someone else higher up) decision to not hire an evaluator
13. Adding/Managing an evaluator would have added to an already heavy workload
14. Evaluation findings would not be used in program decision-making, marketing, and/or funding decisions
15. Inadequate skillset of program leadership to interpret and/or apply evaluation results/findings
16. Paperwork Reduction Act (PRA) processes mandated by Office of Management and Budget (OMB) too burdensome for designing/implementing evaluations
17. Short turnaround time from when funding organization issues an RFP/FOA to when proposal must be submitted
18. External evaluator cannot be paid for work completed in planning and design phase until program is awarded funding

Analysis of Exclusion Factor Responses – Internal and External Factor Sets

A test of normality on the internal factor means initially indicates the data is not normally distributed (Kolmogorov-Smirnova test statistic = 0.097, $p=0.001$; Shapiro-Wilk test statistic = 0.945, $p=0.000$). However, when viewing the Q-Q plot of the data, the observed values track very closely to the expected normal values. Overall, the internal factor means follow a normal distribution⁵. Similarly, a test of normality on the external factor means initially indicates the data is not normally distributed (Kolmogorov-Smirnova test statistic = 0.065, $p=0.075$; Shapiro-Wilk test statistic = 0.982, $p=0.022$). However, when viewing the Q-Q plot of the data, the observed values again track very

⁵ Tests of normality tables and Q-Q plots for internal and external factor sets are available in Appendix D.

closely to the expected normal values. Overall, the external factor means also follow a normal distribution.

The first set of analysis of the factor data aimed to determine whether each respondent group attributes evaluator exclusion during the program planning and design phase to internal factors (deficiencies within the evaluation practice) or to external factors (deficiencies outside the evaluation practice). This was determined through two paired t-tests⁶.

For the first paired t-test, program evaluators had an overall mean score of 2.25 for the internal exclusion factor set and an overall mean score of 2.61 for the external exclusion factor set. A comparison of these means indicate that the difference is statistically significant at an alpha of 0.05 ($p=0.00$). These results indicate that program evaluators more strongly attribute evaluator exclusion during the program planning and design phase to deficiencies external to the evaluation field.

For the second paired t-test, program leaders had an overall mean score of 2.48 for the internal exclusion factors and an overall mean score of 3.17 for the external exclusion factors. A comparison of these means indicate that the difference is statistically significant at an alpha of 0.05 ($p=0.00$). These results indicate that program leaders also more strongly attribute evaluator exclusion during the program planning and design phase to deficiencies external to the evaluation field.

Summary results are presented in Table 5. For the defensive attribution hypothesis to be supported in the instance of evaluator exclusion during the program planning and design phase, two results were necessary. First, program evaluators must more strongly

⁶ Paired t-test result tables are available in Appendix E.

attribute their exclusion to factors external to the field of evaluation. Second, program leaders must more strongly attribute their exclusion decision to factors internal to the field of evaluation. As both respondent groups more strongly attribute evaluator exclusion during the program planning and design phase to deficiencies external to the evaluation field, the defensive attribution hypothesis is not supported.

Table 5: Paired T-Test Results of Factor Sets by Respondent Group					
	Mean for Internal Factor Set	Mean for External Factor Set	Mean Difference	p-value	Statistically Significant? ($\alpha=0.05$)
Program Evaluator (n=131)	2.25	2.61	-0.36	0.00	Yes
Program Leader (n=31)	2.48	3.17	-0.68	0.00	Yes

Analysis of Exclusion Factor Responses – Individual Factors

The next set of analysis determined the means and statistical significance of the responses provided by both sets of respondents for each individual exclusion factor. This was determined through a series of independent sample t-tests⁷. Table 6 presents a simplified summary of the means for each exclusion factor by respondent group, the unweighted mean for each factor (and the standard deviation), the mean difference between groups, the p-value, and whether the p-value is statistically significant at an alpha of 0.05. Mean results are conditionally formatted with color-coding to show the highest (red) to lowest (green) means. Factors are sorted by their overall unweighted means. The one internal exclusion factor and seven external exclusion factors found to be

⁷ Independent sample t-test result tables are available in Appendix F.

statistically significant are highlighted in light green. Results for the two factors added to the taxonomy are not included in this summary due to their low n.

Table 6. Independent Sample T-Test Results for Individual Exclusion Factors by Respondent Group	Program Evaluator Mean	Program Leader Mean	Unweighted Mean (Std.Dev.)	Mean Difference	p-value	Statistically Significant? ($\alpha = 0.05$)
Factors Internal to the Evaluation Field						
Program leadership skeptical of the benefits of evaluation	2.62	2.77	2.70 (1.41)	-0.15	0.631	No
Evaluator too expensive to include in budget	2.38	2.89	2.64 (1.46)	-0.51	0.098	No
Inadequate skillset of evaluators to provide benefits during the program planning and design phase	1.71	2.62	2.17 (1.24)	-0.91	0.002	Yes
Evaluator may discover negative program results	2.37	1.96	2.17 (1.44)	0.41	0.202	No
Program leadership had negative previous experience with evaluators	1.87	1.96	1.92 (1.22)	-0.09	0.753	No
Factors External to the Evaluation Field						
Program leadership had program staff (who are not evaluators) plan and/or manage evaluation activities	2.68	3.75	3.22 (1.57)	-1.07	0.001	Yes
Program leadership had no knowledge of or training on the benefits of evaluation in the planning and design phase	3.33	3.07	3.20 (1.49)	0.26	0.405	No
Program leadership had belief/understanding that evaluation only needs to occur at the end of a program	3.52	2.79	3.16 (1.48)	0.73	0.017	Yes
Organization funding the program did not require evaluation activities	2.35	3.90	3.13 (1.53)	-1.55	0.000	Yes
Previous programs had been successful without	2.49	3.67	3.08 (1.44)	-1.18	0.000	Yes

an evaluator, so there was no need to change						
Inadequate skillset of program leadership to interpret and/or apply evaluation results/findings	2.82	3.17	3.00 (1.53)	-0.35	0.267	No
Program leadership followed a supervisor's (or someone else higher up) decision to not hire an evaluator	2.11	3.36	2.74 (1.52)	-1.25	0.000	Yes
Evaluation findings would not be used in program decision-making, marketing, and/or funding decisions	2.31	3.04	2.68 (1.42)	-0.73	0.018	Yes
Adding/Managing an evaluator would have added to an already heavy workload	2.47	2.76	2.62 (1.39)	-0.29	0.325	No
Program's hiring process too difficult/lengthy to consider hiring an evaluator	2.08	2.52	2.30 (1.41)	-0.44	0.159	No
PRA processes mandated by OMB too burdensome for designing/implementing evaluations	1.46	2.00	1.73 (1.06)	-0.54	0.043	Yes
Short turnaround time from when funding organization issues an RFP/FOA to when proposal must be submitted	4.71	5.00				
External evaluator cannot be paid for work completed in planning and design phase until program is awarded funding	4.20	0.00				

Analysis of Exclusion Factor Responses – Program Evaluator Type

The next set of analysis looked deeper into the factor ratings of program evaluators in two different types of programs: 1) federal/federally-funded programs; and

2) non-federal/non-federally funded programs. This was determined through a series of independent sample t-tests⁸. Table 7 presents a simplified summary of the means for each exclusion factor by evaluator program type, the overall mean for each factor (and the standard deviation), the mean difference between groups, the p-value, and whether the p-value is statistically significant at an alpha of 0.05. Mean results are conditionally formatted with color-coding to show the highest (red) to lowest (green) means. Factors are sorted by their overall means. The one external exclusion factor found to be statistically significant is highlighted in light green. With only one exclusion factor found to be statistically significant, fifteen of the sixteen factors had no significant differences between the factor means of the two program evaluator group types. Results for the two factors added to the taxonomy are not included in this summary due to their low n.

Table 7. Independent Sample T-Test Results for Individual Exclusion Factors by Evaluator Program Type	Federal Program Evaluator Mean	Non-Federal Program Evaluator Mean	Overall Mean (Std.Dev.)	Mean Difference	p-value	Statistically Significant? ($\alpha = 0.05$)
Factors Internal to the Evaluation Field						
Program leadership skeptical of the benefits of evaluation	2.62	2.63	2.62 (1.43)	-0.01	0.982	No
Evaluator too expensive to include in budget	2.38	2.35	2.38 (1.47)	0.03	0.936	No
Evaluator may discover negative program results	2.25	2.64	2.37 (1.43)	-0.39	0.177	No
Program leadership had negative previous experience with evaluators	1.89	1.86	1.87 (1.22)	0.03	0.915	No
Inadequate skillset of evaluators to provide benefits during the planning and design phase of programs	1.64	1.90	1.71 (1.20)	-0.26	0.306	No
Factors External to the Evaluation Field						

⁸ Independent sample t-test result tables are available in Appendix G.

Program leadership had belief/understanding that evaluation only needs to occur at the end of a program	3.45	3.71	3.52 (1.46)	-0.26	0.377	No
Program leadership had no knowledge of or training on the benefits of evaluation in the planning and design phase	3.33	3.29	3.33 (1.48)	0.04	0.874	No
Inadequate skillset of program leadership to interpret and/or apply evaluation results/findings	2.80	2.78	2.82 (1.57)	0.02	0.942	No
Program leadership had program staff (who are not evaluators) plan and/or manage evaluation activities	2.78	2.51	2.68 (1.51)	0.27	0.385	No
Previous programs had been successful without an evaluator, so there was no need to change	2.49	2.52	2.49 (1.37)	-0.03	0.922	No
Adding/Managing an evaluator would have added to an already heavy workload	2.46	2.57	2.47 (1.38)	-0.11	0.702	No
Organization funding the program did not require evaluation activities	2.30	2.45	2.35 (1.44)	-0.15	0.623	No
Evaluation findings would not be used in program decision-making, marketing, and/or funding decisions	2.29	2.35	2.31 (1.40)	-0.06	0.831	No
Program leadership followed a supervisor's (or someone else higher up) decision to not hire an evaluator	2.20	1.97	2.11 (1.41)	0.23	0.463	No
Program's hiring process too difficult/lengthy to consider hiring an evaluator	1.97	2.37	2.08 (1.36)	-0.40	0.185	No
PRA processes mandated by OMB too burdensome for designing/implementing evaluations	1.55	1.14	1.46 (0.99)	0.41	0.045	Yes
Short turnaround time from when funding organization issues an RFP/FOA to when	4.20	4.67				

proposal must be submitted			
External evaluator cannot be paid for work completed in planning and design phase until program is awarded funding	4.75	0.00	

Analysis of Exclusion Factor Responses – Program Leader Type

The next set of analysis looked deeper into the factor ratings of program leaders in two different types of programs: 1) federal/federally-funded programs; and 2) non-federal/non-federally funded programs. This was determined through a series of independent sample t-tests⁹. Table 8 presents a simplified summary of the means for each exclusion factor by program leader program type, the overall mean for each factor (and the standard deviation), the mean difference between groups, the p-value, and whether the p-value is statistically significant at an alpha of 0.05. Mean results are conditionally formatted with color-coding to show the highest (red) to lowest (green) means. Factors are sorted by their overall means. No exclusion factors were found to be statistically significant, indicating that there are no significant differences between the factor means of the two program leader group types. Results for the two factors added to the taxonomy are not included in this summary due to their low n.

Table 8. Independent Sample T-Test Results for Individual Exclusion Factors by Program Leader Program Type	Federal Program Leader Mean	Non-Federal Program Leader Mean	Overall Mean (Std.Dev.)	Mean Difference	p-value	Statistically Significant? ($\alpha = 0.05$)
Factors Internal to the Evaluation Field						
Evaluator too expensive to include in budget	2.63	3.17	2.89 (1.34)	-0.54	0.402	No

⁹ Independent sample t-test result tables are available in Appendix H.

Program leadership skeptical of the benefits of evaluation	2.44	3.17	2.77 (1.37)	-0.73	0.252	No
Inadequate skillset of evaluators to provide benefits during the planning and design phase of programs	2.85	2.14	2.62 (1.20)	0.71	0.233	No
Program leadership had negative previous experience with evaluators	1.88	2.17	1.96 (1.27)	-0.29	0.726	No
Evaluator may discover negative program results	1.84	1.25	1.96 (1.27)	0.59	0.416	No
Factors External to the Evaluation Field						
Organization funding the program did not require evaluation activities	3.84	4.14	3.90 (1.21)	-0.30	0.465	No
Program leadership had program staff (who are not evaluators) plan and/or manage evaluation activities	3.74	4.17	3.75 (1.53)	-0.43	0.557	No
Previous programs had been successful without an evaluator, so there was no need to change	3.50	4.29	3.67 (1.30)	-0.79	0.151	No
Program leadership followed a supervisor's (or someone else higher up) decision to not hire an evaluator	3.53	2.40	3.36 (1.52)	1.13	0.165	No
Inadequate skillset of program leadership to interpret and/or apply evaluation results/findings	3.05	3.14	3.17 (1.39)	-0.09	0.889	No
Program leadership had no knowledge of or training on the benefits of evaluation in the planning and design phase	2.95	3.20	3.07 (1.56)	-0.25	0.759	No
Evaluation findings would not be used in program decision-making, marketing, and/or funding decisions	2.74	3.80	3.04 (1.40)	-1.06	0.133	No
Program leadership had belief/understanding that evaluation only needs to occur at the end of a program	2.60	3.17	2.79 (1.45)	-0.57	0.415	No
Adding/Managing an evaluator would have added to an already heavy workload	2.63	3.14	2.76 (1.41)	-0.51	0.428	No

Program's hiring process too difficult/lengthy to consider hiring an evaluator	2.33	3.00	2.52 (1.56)	-0.67	0.393	No
PRA processes mandated by OMB too burdensome for designing/implementing evaluations	2.00	1.83	2.00 (1.25)	0.17	0.800	No
Short turnaround time from when funding organization issues an RFP/FOA to when proposal must be submitted	5	N/A				
External evaluator cannot be paid for work completed in planning and design phase until program is awarded funding	N/A	N/A				

CHAPTER 5

DISCUSSION

The purpose of this study was to identify the extent to which evaluator exclusion from a program's planning and design phase occurs; to organize a taxonomy of internal and external factors that influence this exclusion; and to explore the extent to which program leaders and program evaluators demonstrate defensive attributions when rating these factors' influence on this exclusion. In this section, I will interpret the study's results, explore how they relate to the current literature, identify why they are valuable, and discuss their implications for the program leadership field, program evaluation field, and funding organizations.

Evaluator Exclusion Rate

A review of the current literature presented several recent developments that promote evaluator involvement in the program planning and design phase. Evaluation researchers and academics identified the benefits of including evaluators in program planning and design and have advocated for the practice. Three pieces of federal legislation, most recently the Evidence Act of 2019, mandate federal agencies to utilize evaluators to collect rigorous evidence and data prior to commencing any internal decision-making processes, and then use that information to craft policies, and to design and fund programs, that will achieve expected results. This advocacy has led to an expansion of the program evaluator role in both federal and non-federal programs. Building on this momentum, a new approach to evaluation emerged, named developmental evaluation, which relies on an evaluator's close involvement with a program staff's planning and design activities.

Despite these developments, the extent to which evaluators were actually being included in a program's planning and design phase was unknown. Based on responses from 260 members of the American Evaluation Association and 61 members of the Project Management Institute, this study established that approximately 70% of evaluators were not included in program planning and design in at least one of their three most recent program experiences. A chi-square test confirmed that there were no significant differences between responses from program evaluators and program leaders on this data point, further confirming this finding.

In light of the recent developments that both promote and mandate evaluator involvement in the program planning and design phase, a 70% exclusion rate can only be perceived as quite high. As this is the first study to identify an evaluator exclusion rate, it is unknown where this rate falls in the overall trajectory of evaluator exclusion from program planning and design efforts. The rate may represent a decrease in evaluator exclusion due to recent inclusion advocacy efforts and the mandates of the Evidence Act. The rate may represent an increase in evaluator exclusion due to tighter organizational budgets where program leaders are asking existing program staff to take on evaluation activities, program leaders are more actively excluding evaluators due to a continued trend of data overload, and many other possible reasons. The rate may also represent a steady, unwavering data point due to program leaders' ongoing lack of knowledge about the benefits of evaluation during the program planning and design phase. Despite these unknowns, an evaluator exclusion rate of 70% during program planning and design sets a baseline data point for future research studies to build upon. Future studies may also

clarify which (or what combination) of these developments have had the strongest effect on the evaluator exclusion rate, and in what direction.

Overall, the rate of exclusion casts doubt on whether the professional responsibilities of an evaluator are even possible to accomplish. If such a high rate of exclusion from program planning and design activities continues, the program evaluation and program leadership fields should expect that: 1) evaluators will be unable to deliver the full spectrum of the benefits associated with the evaluation profession; 2) evaluators will continue to fall short of program leaders' expectations; 3) evaluators will be unable to meet the Evidence Act's mandate to use rigorous evidence to plan effective policies and programs; and 4) the field of developmental evaluation will remain a niche approach that is only available to program teams with an embedded evaluator.

A broad recommendation for addressing the evaluator exclusion rate is for the American Evaluation Association to collaborate with the federal government to clearly define, and provide actionable guidelines on, what type of programs require evaluator inclusion at all, and which subset of these programs require evaluator inclusion during the planning and design phase. Restricting measurement of an evaluator exclusion rate to only a pool of programs that require evaluator involvement is likely to decrease the rate. For more specific and intentional recommendations, a deeper look at the influence that the individual factors identified through this study have on evaluator exclusion will be presented in the ensuing sections.

Taxonomy

A review of the current literature uncovered numerous ideas on why a program leader may decide to include or exclude an evaluator from the program planning and design

phase. From these previously scattered ideas, this study organized a list of factors for program leaders and program evaluators to rate their influence on this decision. The list was given additional structure by organizing the factors into internal (deficiencies within the evaluation field) and external (deficiencies outside the evaluation field) factors. As a taxonomy differentiates itself from a classification by being comprehensive, study participants were provided an opportunity to add and rate additional exclusion factors to the initial list. Based on an analysis of the added factors, two more exclusion factors not found in the current literature were added to the list. These factors are both outcomes of the current funding process, which pressures program leaders to quickly respond to requests for funding without program evaluator input. The two added factors, categorized as deficiencies that are external to the field of evaluation, are:

1. Short turnaround time from when a funding organization issues an RFP/FOA to when proposal must be submitted; and
2. External evaluator cannot be paid for work completed in planning and design phase until program is awarded funding

With the addition of these two factors, funding organizations and the fields of evaluation and program management now have a comprehensive and organized taxonomy of the factors that influence evaluator exclusion during the program planning and design phase. This taxonomy provides numerous benefits to funding providers and recipients. It provides a foundational set of terms and concepts when discussing the topic of evaluator involvement in program planning and design. The comprehensive set of exclusion factors provides a reference point for identifying policies and processes to both directly address evaluator exclusion as well as encourage the practice of evaluator

inclusion when deemed necessary. Finally, a comprehensive set of exclusion factors lays the groundwork for future efforts to measure the effectiveness of policies and processes aimed at promoting evaluator inclusion in the program planning and design phase.

While over 150 respondents rated the influence of at least one factor in the taxonomy, this research study marks the first time that evaluator exclusion factors have been organized and analyzed. As the program evaluation and program management fields continue to mature and evolve in reaction to federal legislation, additional exclusion factors are likely to arise that will necessitate amendments to the current taxonomy. Therefore, future studies that explore the completeness of the taxonomy are enthusiastically encouraged.

Additionally, this study organized the taxonomy by locus of control in order to study the research topic through the lens of attribution theory. However, there may be more appropriate ways to organize the exclusion factors. Funding organizations that are interested in promoting evaluator inclusion may find it more useful to organize the factors by those which can be addressed by policy, process, practice, or training. Professional organizations, such as AEA or PMI, may find it within their interest to organize the factors by those that can be addressed through research, training, consultation, collaboration, or advocacy. Overall, alternate approaches to organizing the current list may uncover missing, redundant, or unclear factors that will contribute to the creation of a more complete and comprehensive taxonomy.

Evaluator Exclusion - Internal and External Factor Sets

The defensive attribution hypothesis states that participants in a failure event tend to explain the failure in a way that minimizes their personal responsibility. They do so by

externalizing the event's causality and searching for alternative explanations in order to protect themselves and their self-esteem. In the case of evaluator exclusion in the program planning and design phase, the hypothesis would be upheld if the following two statements were supported by the data: 1) Program evaluators more strongly attribute their lack of inclusion to deficiencies outside of the evaluation practice; and 2) Program leaders more strongly attribute evaluator exclusion to deficiencies within the evaluation practice.

Based on this study's results, the defensive attribution hypothesis is not supported in the case of evaluator exclusion during the program planning and design phase. Program evaluators did more strongly attribute their lack of inclusion to deficiencies outside of the evaluation practice. However, program leaders did not more strongly attribute evaluator exclusion to deficiencies within the evaluation practice. Instead, program leaders joined their program evaluator colleagues by more strongly attributing evaluator exclusion to deficiencies outside of the evaluation practice. Additionally, these results were found to be statistically significant for both respondent groups. This is critically important information for all who are interested in achieving higher rates of evaluator inclusion in program planning and design, as this agreement between the two respondent groups point to a set of factors where changes to policy, process, and training are needed.

To better understand this information, it is necessary to look deeper into the approach that yielded this result. First, it must be noted that the number of exclusion factors included in each factor set are not balanced. There are five exclusion factors identified as internal to the evaluation field, and there are thirteen exclusion factors identified as external to the evaluation field. As the current taxonomy of exclusion factors is

considered to be comprehensive, additional internal factors cannot simply be manufactured to achieve a better balance. Similarly, external factors cannot simply be removed. Overall, this imbalance accurately represents a situation where program leaders have more options to make a decision to exclude an evaluator based on deficiencies found within the program management field and in funding organization processes, and have fewer options to make a decision to exclude evaluator based on deficiencies found within the evaluation practice.

This imbalance was addressed by a few study elements. First, due to their low n, the two factors added to the external factor set after data collection was completed were not included when comparing the factor sets. This reduces the number of factors in the external factor set to eleven, while keeping the number of factors in the internal factor set at five. Next, calculating the means for each respondent's ratings for factors within each overall set mitigates the existing factor imbalance. Then, comparing these means for each factor set within each respondent group provides a standard approach for ascertaining which factor set each group more strongly attributes to influencing evaluator exclusion during the program planning and design phase. Finally, the possibility remains that a higher mean for one individual factor has the ability to drive up the overall mean of the internal factor set more than it does in the external factor set. However, in the case of this study, this situation did not occur. Moreover, the seven individual factors with the highest overall mean are all found in the external factor set (see Table 6), strengthening the finding that both respondent groups more strongly attribute evaluator exclusion during the program planning and design phase to factors that are external to the evaluation field.

While this conclusion is accurate at this time, it is again imperative to note that the taxonomy on which these findings are based is subject to change. This is especially true since the implications of the Evidence Act on the program management and program evaluation fields are only just beginning to emerge. Any revisions made to the exclusion factors will affect the structure of the factor sets, which will, in turn, affect the results of the relational tests applied to respondents' factor ratings. Additionally, any future efforts made by program management, program evaluation, or funding organizations to address the exclusion factors will necessitate a reiteration of this study's data collection activities to assess their affects.

Evaluator Exclusion – Individual Factors

In previous sections, the discussion explored the evaluator exclusion rate, the taxonomy of evaluator exclusion factors, and the direction of attribution from respondent groups. As the defensive attribution hypothesis was shown to not be supported in the case of evaluator exclusion during the program planning and design phase, the discussion in this section will focus on the extent to which each individual factor influenced this finding. To do so, the discussion will focus on the top five of the sixteen factors found to influence evaluator exclusion, as measured by their unweighted means¹⁰. Additionally, the discussion will include factor results found to be statistically significant as well as those found not to be statistically significant, in terms of their mean differences between the program leader and program evaluator respondent groups. In this study, the factors found to be statistically significant are just as important as the factors found to not be statistically significant because they tell us two different, but helpful bits of information.

¹⁰ The unweighted mean = (factor mean of program leaders + factor mean of program evaluators) / 2

First, for factors found not to be statistically significant, the two respondent groups are assigning the factor a similar level of influence on the decision to exclude an evaluator. If these factors are also rated at a high level of influence, then both respondent groups are agreeing that the factor is highly influential on the decision to exclude an evaluator from program planning and design activities. Second, for factors found to be statistically significant, the two respondent groups are assigning the factor different levels of influence on the decision to exclude an evaluator. One respondent group is therefore indicating that a factor more strongly contributes to this decision than the other group believes it does. This dichotomy illustrates how each respondent group interprets the factors contributing to evaluator exclusion differently, and attributes their own different level of influence to the factor. Overall, each of these situations provides a revealing glimpse into how program leaders and program evaluators attribute level of influence to the factors affecting evaluator exclusion in program planning and design.

1. Existing Program Staff Plan and Manage Evaluation Activities

The exclusion factor with the highest overall unweighted mean is: Program leaders have program staff, who are not evaluators, plan and manage evaluation activities. Program leaders rated this as their second most influential exclusion factor, with a mean score of 3.75 out of 5.00. Program evaluators rated this as their fourth most influential exclusion factor, with a mean score of 2.68 out of 5.00. The differences between these means scores was found to be statistically significant ($p=0.001$).

The literature review noted that evaluation was a formalized part of the program planning and design process beginning in the mid-1960s, where a needs assessment and evaluation plan were expected to be present in every funded program (Patton, 1978). As

federal funding for social service programs declined through the 1980s, the emphasis on evaluative thought and practice began to decline (Chelimsky, Cordray, & Datta, 1989; Carman et al., 2008). According to this study's results, the most influential factor on evaluator exclusion during program planning and design is that program leaders have offloaded the planning, management, and implementation of evaluation activities to program staff who are not professional evaluators. Funding organizations aiming to instill a renewed emphasis on program evaluation activities in their FOAs and RFPs are therefore encountering program leaders who are willing to delegate these activities to non-evaluator staff.

It is possible that program leaders do not have the budget or time to hire and assign evaluation activities to a professional evaluator. However, both program leaders and program evaluators rated the factors of evaluator expense and hiring timeframe as less influential than assigning evaluation activities to non-evaluation staff. Moreover, as the most influential factor on evaluator exclusion in the program planning and design phase, program leaders are expressing a preference to do so over all other factors. As most funding organizations do not mandate or enforce that a professional evaluator plan, implement, and manage evaluation activities, program leaders have the flexibility to delegate these activities as they see fit, and are taking advantage of that opportunity. Overall, these actions indicate that program leaders see evaluation activities as additional responsibilities that existing program staff can plan, implement, and manage on their own. Conversely, program leaders do not see evaluative activities as a set of competencies only reserved for professionally trained evaluators.

In cases where a funding organization does not require an experimental design, quasi-experimental design, or other rigorous, academic program assessment, this approach may adequately fulfill expected evaluation activities. However, in all cases where untrained program staff plan, implement, and manage evaluation activities, the broad spectrum of benefits associated with professional program evaluation will be poorly provided, will be provided in a piecemeal fashion, or will not be provided at all. As presented in the literature review, untrained program staff do not possess the necessary tools or skills to systematically build program knowledge that will hold the program accountable, inform program-based decisions, or accurately measure program success. Program leaders (and, by proxy, funding organizations) who delegate evaluation activities to these untrained staff should not expect them to.

Two factors were added to the taxonomy after data collection was completed, and these factors likely affect how respondent groups perceive the influence of delegating evaluation activities to non-evaluation staff. With a short turnaround time to respond to a request for proposal, and an inability to pay external evaluators for planning and design work before funding is awarded, program leaders have likely grown accustomed to delegating evaluation planning and design activities to existing program staff. This is a rational response to a funding process that requires rapid mobilization of the program staff preparing a proposal, but that does not provide advanced funding to subsidize planning and design activities. With this being said, future studies will need to assess the relationship between program leaders delegating evaluation activities to non-evaluation staff and these two new factors in order to better understand the practice.

As the current funding process passively encourages the practice, and program leaders have the flexibility to implement the practice, the delegation of evaluation activities to non-evaluation staff is a reality that is likely to continue. The professional evaluation field, led by AEA, therefore has an opportunity to position itself as a proactive collaborator in this area. AEA can collaborate with funding organizations to define and delineate the programs that will benefit from having a professional evaluator, the programs that may delegate evaluation activities to program staff, and the programs that require no evaluation activities at all. For programs that require a professional evaluator, AEA can publish guidance for funding organizations on how to incorporate evaluation requirements into their RFPs/FOA, how to set program evaluation budgets with grantees, and how to score the evaluation sections of submitted proposals. As funding organizations are currently building policies and processes in response to the requirements of the Evidence Act, this collaboration between AEA and funding organizations will promote practices to involve professional evaluators in planning and design activities in programs that need their expertise.

In regards to this factor's relationship with the defensive attribution hypothesis, the willingness of program leaders to exclude an evaluator from the planning and design phase is not an indicator that evaluation activities do not occur. It is an indication that evaluators are being omitted by program leaders who prefer to delegate evaluation activities to non-evaluation staff. As the funding process allows, and may even encourage, program leaders to do so, both of the study's respondent groups appear to be attributing evaluator exclusion not only to deficiencies within the program leadership field, but to extra-external deficiencies outside of the program leadership and program

evaluation fields that currently exist within the funding process. This is a critical distinction that will be expanded upon and discussed in more detail below.

2. Program leaders have no knowledge of or training on the benefits of evaluation in the planning and design phase

3. Program leaders believe evaluation only needs to occur at the end of a program

Due to the similarities in the factors rated by respondents as the second and third most influential on evaluator exclusion, they will be discussed together.

The exclusion factor with the second highest overall unweighted mean is: Program leaders have no knowledge of or training on the benefits of evaluation in the planning and design phase. Program leaders rated this as their sixth most influential exclusion factor, with a mean score of 3.07 out of 5.00. Program evaluators rated this as their second most influential exclusion factor, with a mean score of 3.33 out of 5.00. The differences between these means scores was not found to be statistically significant ($p=0.405$).

The exclusion factor with the third highest overall unweighted mean is: Program leaders have the belief/understanding that evaluation activities only need to occur at the end of the program. Program leaders rated this as their ninth most influential exclusion factor, with a mean score of 2.79 out of 5.00. Program evaluators rated this as their most influential exclusion factor, with a mean score of 3.52 out of 5.00. The differences between these means scores was found to be statistically significant ($p=0.017$).

Program leaders' level of knowledge and training about evaluation in program planning and design has been proposed as a potential exclusion factor in the academic literature. As presented in the literature review, Sanders (2003), Carman (2007), and Preskill (2014) noted that program leaders typically have little or no training on the

benefits that evaluation can provide to programs in their planning and design phase. For program leaders that have received formal training on the benefits of evaluation, McLaughlin and Jordan (2004), Rossi and Freeman (1993), Preskill and Torres (2001), and Scheirer (2012) noted that this training typically presents evaluation as an activity that occurs at the end of the program. Through this study, program leaders' limited knowledge and training on evaluation were found to be two of the leading factors that influence evaluator exclusion during the program planning and design phase. Moreover, program evaluators rated these factors as the top two influences on evaluator exclusion, with both factors the only ones receiving an average influence rating above 3.00 from the respondent group. In doing so, program evaluators lend support to the defensive attribution hypothesis by attributing their exclusion from program planning and design activities to program leaders' level of evaluation knowledge. However, as described earlier, the defensive attribution hypothesis needs each respondent group to attribute failure to the other group, and this dual attribution did not occur in this study.

Program leaders did not rate their own level of evaluation knowledge as having as strong of an influence on evaluator exclusion as program evaluators did. While these differences were only found to be statistically significant in the case of evaluation activities occurring at the end of a program, future studies should explore why program evaluators so strongly attribute their exclusion to program leaders' level of evaluation knowledge. Additionally, future studies should explore the relationship between program leaders' level of evaluation knowledge and their willingness to delegate evaluation activities to program staff who are not evaluators.

Based on this study's findings, program leaders' lack of knowledge that evaluators can be included in program planning and design activities, and the benefits they can provide when included in this phase, is an indicator that program leaders may not be consciously excluding evaluators during this phase. It is an indication that program leaders are unknowingly omitting evaluators during the program planning and design phase. Overall, program evaluators and, to a lesser extent, program leaders perceive the level and type of evaluation knowledge as two of the main factors influencing evaluator exclusion during the program planning and design phase. Again, AEA has an opportunity to position itself as a proactive collaborator in this area as well. AEA has the expertise to collaborate with professional program management training organizations, such as PMI, to design a basic evaluation training module that can be integrated into current curricula. At a minimum, this training should include the general concepts of evaluation, how to identify program situations where evaluation expertise would be most helpful, how to establish and maintain a collaborative relationship with a program evaluator, and the benefits that evaluation and evaluative thinking can provide to a program. Doing so will provide program staff who pursue professional training with the basic knowledge and skills needed to plan, implement, and manage the evaluation activities delegated to them by program leaders, as well as the knowledge of when program evaluation activities are best applied by professional evaluators.

While this suggestion is driven by the study's findings, it is one with precedent. Writing in 1985 for the *Project Management Journal*, a publication of PMI, Loo (1985, p.41) offered that the Institute's program management approach is "easily amenable to comprehensive evaluation." Further, Loo proposed that evaluation concepts and methods

be a required subject of study in program management education and be integrated into PMI's curriculum to foster positive attitudes towards evaluation. Perhaps to build consensus and momentum, Cleland (1985) offered a similar perspective in the very same journal publication. Cleland observed that a project manager who neglects an ongoing evaluation is 'at sea without a compass,' for without an evaluator there is no assessment of whether project activities have been effectively accomplished. While the authors' advice and guidance was not heeded then, a renewed collaborative effort between AEA and PMI would mitigate the influence of program leaders' level of evaluation knowledge on evaluator exclusion.

4. Funding organization did not require evaluation activities

The exclusion factor with the fourth highest overall unweighted mean is: The organization funding the program did not require evaluation activities. Program leaders rated this as their most influential exclusion factor, with a mean score of 3.90 out of 5.00. Program evaluators rated this as their tenth most influential exclusion factor, with a mean score of 2.35 out of 5.00. The differences between these means scores was found to be statistically significant ($p=0.000$).

As presented in the literature review, no research has reviewed the evaluative components that funding organization include in their FOAs or RFPs. Therefore, the extent to which funding organizations are not including evaluation activities in these documents is unknown. From this study's data, program leaders seem to indicate that evaluation activities are not typically a requirement tied to program funding, and this factor most strongly influenced their decisions to not include an evaluator in program planning and design. Program evaluators seem to indicate that a lack of required

evaluation activities does not strongly influence their exclusion from a program's planning and design phase. However, as this study asked both respondent groups to reflect on their program experiences when rating the influence of the exclusion factors, the results may simply be a quirk of the data collection process. Briefly, program leaders can have a program experience without including an evaluator in any program phase. Therefore, program leaders have the flexibility to exclude an evaluator from a program entirely if a funding organization did not require any evaluation activities. This allowed program leaders to respond to the study's questionnaire regardless of whether an evaluator was included in their program experience or not. However, for a program evaluator to have a program experience, they must have been included at some point in the program. Program evaluators were allowed to respond to the study's questionnaire only if they had experienced exclusion during the planning and design phase of a recent program but had then been included in subsequent program phases. Overall, program leader respondents had more program experiences where funding organizations did not require evaluation activities than their program evaluator colleagues. The perceived influence of the factor on evaluator exclusion was therefore higher in program leaders' experiences than in program evaluators' experiences, and this difference was statistically significant.

The results for this factor best highlight a trend in program leaders' ratings of factor influence on evaluator exclusion. To test the defensive attribution hypothesis in the area of evaluator exclusion in the program planning and design phase, this study categorized exclusion factors as either internal to the evaluation field or external to the evaluation field. However, not all exclusion factors that are external to evaluators are, by

default, internal to program leaders. The factors discussed above contain some extra-external elements: 1) funding organizations, which are external to the program leadership field, allow program leaders to delegate evaluation activities to non-evaluation staff; and 2) professional program management training organizations, which are also external to the program management field, either do not train program leaders on evaluation at all, or train them that evaluation only needs to occur at the end of a program. For the factor currently being discussed, program leaders seem to be defensively attributing evaluator exclusion to funding organizations. A key takeaway here is that evaluator exclusion during the program planning and design phase does not take place in a closed system. Program leaders do not only have the options of attributing evaluator exclusion to themselves or to program evaluators. Conversely, program evaluators do not only have the options of attributing their own exclusion to themselves or to program leaders. The context of the system in which this exclusion occurs matters, and this context includes the practices of funding organizations and professional training organizations, and the influence of these practices on the decision to exclude. This reality presents complicating factors for the defensive attribution hypothesis to navigate. It is an indicator that the defensive attribution hypothesis may be best suited for application within a closed system of relationships. Research studies that explore the utility of the defensive attribution hypothesis in closed and open systems will be necessary to draw an informed conclusion.

The opportunity for program leaders to attribute evaluator exclusion to factors external to the program leadership and program evaluation fields does not absolve them of their failure to include evaluators. With an evaluator exclusion rate of approximately 70% during the program planning and design phase, program leaders do not appear

inclined to proactively include program evaluators when a funding organization does not require evaluation activities. Similar to the discussion of the most influential exclusion factor, where program leaders expressed a preference to delegate evaluation activities to non-evaluation staff, program leaders here are expressing a preference to not include evaluators if a funding organization does not require them to do so.

In an effort to gather good evidence on which to base data-driven programmatic decisions, the Evidence Act is actively legislating against the continuation of this practice. As presented in the discussion of the most influential exclusion factor, it is once again recommended that AEA collaborate with funding organizations to define programs that will benefit from having a professional evaluator, and publish guidance for funding organizations on how to incorporate evaluation requirements into the RFPs/FOAs of these programs. Building on this suggestion, the guidance should focus on the evaluative components that funding organizations should include in their RFPs/FOAs. As presented in the literature review, these components would include a program needs assessment, a theory of change, and an evaluation plan, all agreed upon by the potential program leader and evaluator.

5. Previous programs had been successful without an evaluator

The exclusion factor with the fifth highest overall unweighted mean is: Previous programs had been successful without an evaluator, so there was no need to change. Program leaders rated this as their third most influential exclusion factor, with a mean score of 3.67 out of 5.00. Program evaluators rated this as their sixth most influential exclusion factor, with a mean score of 2.49 out of 5.00. The differences between these means scores was found to be statistically significant ($p=0.000$).

As presented in the review, the current literature presents several reasons why program leaders knowingly choose to exclude evaluators from their programs. These include skepticism of the benefits of evaluation, inadequate skillset to interpret and apply evaluation findings, non-use of evaluation findings in program decision-making, and poor previous experience with an evaluator. From the perspective of program leaders, all of these factors were rated as less influential than the factor where program leaders had achieved program success without evaluators, so there was no need to include one.

Saxena and Lamest (2018) noted that this type of organizational or program inertia is often due to information overload experienced by program leaders, as data is more easily generated and available than ever before. It may be that the data needs of this study's program leader respondents are being met without the inclusion of evaluators. Future studies are necessary to understand the extent to which this is the case. Additionally, future studies should explore how program leaders define program success, and, more importantly for the evaluation field, how program leaders know their programs have achieved success when a program evaluator was not included to substantiate the claim.

In relation to previously discussed factors, we have learned that program leaders express a preference to not include evaluators in program planning and design activities if a funding organization either did not require them to do so, or allowed them to delegate evaluation activities to non-evaluation staff. Similarly, program leaders are also expressing a preference to not include evaluators in program planning and design activities because previous programs had been successful without doing so. However, whereas the previous factors allowed program leaders to defensively attribute evaluator exclusion to funding organizations, program leaders have abandoned this defensive

stance when rating the influence of previous program success on their decision-making. By acknowledging the strong influence of previous program success when evaluators were not included, program leaders are attributing evaluator exclusion directly to their own program staffing decisions. Where previously discussed external exclusion factors had extra-external components that allowed program leaders to maintain a partially defensive stance on the topic of evaluator exclusion, this factor demonstrates program leaders' willingness to accept responsibility for evaluator exclusion. When viewed alongside program leaders' low influence ratings for factors internal to the evaluation field, their willingness to accept responsibility highlights why the defensive attribution hypothesis is not supported in this study.

Program leaders' preference to continue applying an approach that has led to previous program success is a reasonable one. While the literature review established that program evaluators possess a unique skillset that can benefit all program types, it is expected that program types exist where program leaders are able to achieve program objectives and credibly support claims of program success without the inclusion of an evaluator. This allows for an improvement upon an earlier recommendation that AEA, in collaboration with funding organization, define the types of programs that require, would benefit from, or do not require evaluation activities. Based on this study's findings, AEA, in collaboration with funding organizations, should define the types of programs that require, would benefit from, or do not require evaluation activities to achieve and substantiate program success. Moreover, identifying the pool of programs that require an evaluator or would benefit from an evaluator will further refine the evaluator exclusion rate established by this study. It will also allow future studies to hone our understanding

of why evaluator exclusion during the program planning and design phase occurs in programs that require, or would benefit from, the inclusion of evaluators in planning and design activities to achieve and substantiate program success.

Evaluator Exclusion – Individual Factors Summary

The discussion above focused on the five factors that most influence evaluator exclusion during the program planning and design phase. All five factors are external to the field of evaluation, thereby reinforcing the finding that the defensive attribution hypothesis is not supported in this area.

Program evaluators most strongly attribute their exclusion to program leaders' insufficient training on the role of evaluation during the program planning and design phase. This includes program leaders' lack of knowledge about the benefits of evaluation during program planning and design activities, as well as program leaders' inaccurate belief that evaluation activities only need to occur at the end of a program. In this regard, program evaluators' attributions are wholly defensive. However, program evaluators' defensive attributions can be assigned to both the program leaders' own lack of knowledge as well as the professional program management training organizations that neglected to provide them with this knowledge. The latter is an extra-external component of the program leader and program evaluator relationship, which complicates the application of the defensive attribution hypothesis. The lack of evaluation content in program management training curriculum indicates that program leaders may not be consciously excluding evaluators in all situations, but simply omitting evaluators from the program planning and design phase due to a lack of knowledge otherwise.

Program leaders most strongly attribute evaluator exclusion to two main factor areas. The first area is program leaders' own staffing decisions, where they either decide to not include evaluators due to achieving previous program success without them, or independently decide to assign evaluation activities to non-evaluation staff. These decisions demonstrate program leaders' preference to exclude evaluators from program planning and design activities when afforded the opportunity to do so. As program leaders provided low influence ratings for factors internal to the evaluation field, their willingness to accept responsibility for evaluator exclusion underscores why the defensive attribution hypothesis is not supported in this study.

The second area is a deficient funding process, where funding organizations either do not require program evaluation activities at all, or allow program leaders to assign evaluation activities to non-evaluation staff. Here, program leaders are attributing evaluator exclusion to deficiencies that are external to the program evaluation, but that are not entirely internal to the program leadership field. While program leaders still retain the power to make a final decision on whether to include a program evaluator or not, a funding organization guides this decision based on their own funding requirements. By not requiring program evaluation activities or by allowing program leaders to delegate evaluation activities to non-evaluation staff, funding organizations become an extra-external deficiency outside of the program leadership and program evaluation fields to which evaluator exclusion can be partially attributed. Evaluator exclusion during the program planning and design phase is a failure event that does not exist within a closed system: extra-external elements to this event exert influence on program leaders' decision to exclude. This open system presents a complicated set of influences for the defensive

attribution hypothesis to navigate, and provides another indication of why the defensive attribution hypothesis is not supported in this study.

CHAPTER 6

CONCLUSION

This chapter will conclude the study by summarizing the main results and findings in relation to the study's overall purpose, discussing the study's value and contribution, reviewing the study's limitations, and proposing opportunities for future research.

Results and Findings

This study had four main purposes, each of which was framed by a research question. First, the study aimed to identify the extent to which evaluator exclusion from a program's planning and design phase occurs. This purpose emerged from a collective frustration within the program evaluation field, where federal legislation mandates, and evaluation academics advocate for, evaluator inclusion in the program planning and design phase, but evaluators still find themselves excluded from these activities. This study identified that 72% of program evaluators (n=260) had been excluded from the program planning and design phase in at least one of their three most recent experiences. Similarly, 69% of program leaders (n=61) acknowledged that they had excluded program evaluators from the program planning and design phase in at least one of their three most recent program experiences. Taken together, this study found that approximately 70% of program evaluators have experienced exclusion from the planning and design phase in their recent program experiences.

The second and third purposes of the study were to identify a comprehensive set of factors that influence evaluator exclusion during the program planning and design phase, and to organize these factors into a taxonomy of internal (deficiencies within the evaluation practice) or external (deficiencies outside the evaluation practice) locus of

control. The academic literature proposes multiple possibilities of why program evaluators are being excluded from programs, and this study organized these previously scattered ideas into a taxonomy of sixteen factors (five internal factors and eleven external factors). Based on an analysis of respondents' additions, two additional external factors not found in the literature were added, resulting in a comprehensive taxonomy of eighteen evaluator exclusion factors (five internal factors and thirteen external factors).

The fourth and final purpose of the study was to explore the extent to which program leaders and program evaluators demonstrated defensive attributions when rating these factors' influence on evaluator exclusion during the program planning and design phase. This study utilized its taxonomy of eighteen exclusion factors to provide a foundation and structure for doing so. Based on the study results, the defensive attribution hypothesis is not supported in the case of evaluator exclusion during the program planning and design phase. While program evaluators more strongly attributed their lack of inclusion to deficiencies outside of the evaluation practice, program leaders also more strongly attributed evaluator exclusion to deficiencies outside of the evaluation practice. These results were statistically significant for both groups.

By applying a deeper analysis to the factors that influence evaluator exclusion during the program planning and design phase, this study found program evaluators most strongly attribute their exclusion to program leaders' insufficient training on the role of evaluation during the program planning and design phase. This included program leaders' lack of knowledge about the benefits of evaluation during program planning and design activities, as well as program leaders' inaccurate belief that evaluation activities only need to occur at the end of a program. Program leaders most strongly attribute evaluator

exclusion to two main areas: 1) Program leaders' own staffing decisions, where they either decide to not include evaluators due to achieving previous program success without them, or independently decide to assign evaluation activities to non-evaluation staff; and 2) A deficient funding process, where funding organizations either do not require program evaluation activities at all, or allow program leaders to assign evaluation activities to non-evaluation staff.

Overall, the study achieved its four main purposes.

Value and Contribution

The results of this study have profound implications for funding organizations, the program management and program evaluation fields, their advocates, and their practitioners. This includes new knowledge on the extent to which evaluators are excluded from the program planning and design phase, and that this exclusion is attributed to deficiencies external to the evaluation field. This knowledge will provide a collective atonement and exculpation to program evaluators frustrated by their exclusion, as the study found that the phenomenon is not an indictment of individual evaluator's skillsets or of the general practices of the evaluation field. However, the knowledge will also be alarming to a professional evaluation field that is dedicated to providing programmatic benefits across all program phases, as it is an indicator that these efforts are being hampered by external deficiencies. Such a high rate of exclusion will move program evaluators to question whether their professional responsibilities are possible to accomplish in the current system of program funding and management.

The requirements specified in GPRA, the GPRA Modernization Act, and the Evidence Act provide the guardrails for the current funding system, and for the program

evaluators working within that system. As the most recent legislation, the Evidence Act directs policymakers and program planners at every federal agency to collect rigorous evidence and data before any internal decision-making processes, and to use that information to craft policies and to design and fund programs that will achieve expected results. Inclusion of evaluators during the program planning and design phase is key to realizing the intended benefits of the Evidence Act. The extent of evaluator exclusion uncovered through this study endangers society's ability to achieve those benefits.

As this study also explored the factors that influence evaluator exclusion, several practical policy and process changes arise that will directly address them. Broadly, the taxonomy of exclusion factors developed through this study provides a foundational set of terms and concepts to discuss and address the topic of evaluator involvement in program planning and design. The taxonomy also provides a reference point for identifying policies and processes to directly address the factors influencing evaluator exclusion, and lays the groundwork for future efforts to measure the effectiveness of these policies and processes in mitigating evaluator exclusion in the program planning and design phase.

More specifically, the taxonomy's factors identified as most influential to evaluator exclusion direct stakeholders to several areas of necessary policy and process changes. This comes at a critical time, as funding organizations are determining appropriate approaches to comply with Evidence Act requirements. First, this study found that program evaluators most strongly attribute their exclusion to program leaders' inaccurate, insufficient, or nonexistent training on the role of evaluation during the program planning and design phase. To address this factor, AEA has an opportunity to collaborate with

professional program management training organizations, such as PMI, to design a basic evaluation training module that can be integrated into current curricula. This training should include the general concepts of evaluation, how to identify program situations where evaluation expertise would be most helpful, how to establish and maintain a collaborative relationship with a program evaluator, and the benefits that evaluation and evaluative thinking can provide to a program.

Second, this study found that program leaders most strongly attribute evaluator exclusion to two main factor areas. The first area is program leaders' own staffing decisions, where they either decide to not include evaluators due to achieving previous program success without them, or independently decide to assign evaluation activities to non-evaluation staff. The second area is a deficient funding process, where funding organizations either do not require program evaluation activities at all, or allow program leaders to assign evaluation activities to non-evaluation staff. To address these factors, AEA should pursue collaborative partnerships with funding organizations to define and delineate the programs that will benefit from having a professional evaluator, the programs that may delegate evaluation activities to program staff, and the programs that require no evaluation activities in order to achieve and substantiate their success. For programs that require a professional evaluator, AEA should publish guidance on how funding organizations can incorporate evaluation requirements into their RFPs/FOA, such as a program needs assessment, a theory of change, and an evaluation plan. The guidance should also advise funding organizations on how to set program evaluation budgets with grantees and how to score the evaluation sections of submitted proposals.

Implementation of these policy and process changes should mitigate evaluator exclusion during the program planning and design phase, thereby leading to better designed program proposals, improved program evaluation plans, richer and higher quality data for program leaders to inform programmatic decisions, and more rigorous evidence for funding organizations to design and fund programs that will achieve expected results.

In terms of its academic contribution, this study fills a critical gap in evaluation research and literature, as no previous studies had explored the factors that influence program leaders' decision whether to include or exclude an evaluator in the program planning and design phase. As a practice-based field, most academic literature on evaluation explores and instructs practitioners how to better conduct the numerous types of evaluation in different contexts and capacities. This study focused on exploring the space and time prior to implementation of evaluation activities to better understand how external decisions are made that profoundly affect the field of evaluation. The study's findings suggest that the evaluation literature needs to place more stress on how to get evaluators included in programs rather than instructing them on what they should do once they are actually included. Additionally, this study's findings indicate that it is time for evaluation literature to move beyond simple advocacy for evaluator inclusion in all phases of a program's lifecycle. The literature should provide practical guidelines, practices, activities, and descriptive case studies for program leaders and program evaluators to do so. Furthermore, this literature should not be solely directed towards program evaluators, who have generally embraced the advocacy offered, but towards program leaders, who still need training on guidance on why and how to include

evaluators in planning and designing programs, beyond the reason of requirement or mandate.

This study also represents the first time that attribution theory and the defensive attribution hypothesis were applied to the field of program evaluation. As program leaders partially attributed evaluator exclusion to their own staffing decisions, the defensive attribution hypothesis was not supported. However, program evaluators partially attributed their exclusion to program management training organizations and program leaders partially attributed evaluator exclusion to the practices of funding organizations. Both are factors that are external to the program management and program evaluation fields. This indicates that the phenomenon of evaluator exclusion does not occur within a closed system. The context within which evaluator exclusion occurs exerts influence on program leaders' staffing decisions. This network of influences presents complicating circumstances for the defensive attribution hypothesis to navigate, and may be an indicator that the hypothesis is best suited for application within a closed system of relationships.

Limitations

For program evaluator respondents, this study utilized a sample of 1,000 of the approximately 6,000 members of the American Evaluation Association, and achieved a 26% response rate. The respondent group was majority female (66%) with a doctorate degree (52%). As AEA does not require its members to complete a demographic profile, it is unknown if the demographics of the program evaluator respondents constitute a representative sample of AEA membership. Furthermore, as not all professional program

evaluators are members of AEA, study results are not generalizable to the wider program evaluation field.

The 61 program leader respondents in this study represented six of the 142 local Project Management Institute chapters in the United States. To respond to the study questionnaire, respondents needed to take their own initiative to follow a link provided via PMI chapter newsletters, social media posts, and presentations. Therefore, this study's pool of program leader respondents is categorized as a convenience sample. Due to this, there may be bias in the set of program leader responses, and study results should not be generalized to the wider program management field. Additionally, due to the low number of responses and potential response bias from program leaders, future researchers may encounter challenges when attempting to replicate the results of this study.

Finally, two exclusion factors were generated by study respondents and then added to the taxonomy after data collection was complete. Therefore, respondents did not have an opportunity to rate the influence of these two exclusion factors in relation to the set of sixteen factors presented in the questionnaires. If these two factors had been included in the study's questionnaires prior to data collection, respondents may have altered their ratings of influence on evaluator exclusion in the other sixteen factor areas.

Future Opportunities

As this study is the first to explore evaluator exclusion in the program planning and design phase, much work remains to corroborate and build upon its results and findings. First, a 70% evaluator exclusion rate in the program planning and design sets a baseline data point for future longitudinal studies to determine an overall trend. Additionally, efforts to identify the pool of programs that require an evaluator, or would benefit from

an evaluator, will further refine the evaluator exclusion rate established by this study. These efforts will also provide an opportunity for future studies to hone our understanding of why evaluator exclusion occurs in programs that require, or would benefit from, the inclusion of evaluators to achieve and substantiate program success.

Second, future research studies will need to test the comprehensiveness of the taxonomy of evaluator exclusion factors. As the practices of the program evaluation and program management fields continue to evolve in reaction to federal legislation, additional exclusion factors may arise that will necessitate amendments to the taxonomy. Researchers will also want to have new respondents rate the influence of all eighteen factors (inclusive of the two added factors) to see if this expanded taxonomy alters their ratings in any way. Researchers also have an opportunity to explore alternate approaches to organizing the exclusion factors that are more appropriate for the program management field and funding organizations. These efforts may uncover missing, redundant, or unclear factors that will further contribute to the creation of a more complete and comprehensive taxonomy.

Third, deeper exploration into the factors influencing evaluator exclusion provides numerous research opportunities to better understand the phenomenon. Future studies can examine why program evaluators so strongly attribute their exclusion to program leaders' level of evaluation knowledge while program leaders do not. Additionally, future studies can investigate the relationship between program leaders' level of evaluation knowledge and their willingness to delegate evaluation activities to program staff who are not evaluators or to omit evaluators altogether. These studies will lead to a more nuanced understanding of the program leader/program evaluator relationship, which may lead to

more appropriate policies that encourage evaluator inclusion in program planning and design activities. Any future efforts made by program management, program evaluation, or funding organizations to improve current policies that address evaluator exclusion will necessitate a reiteration of this study's data collection activities to assess their affects. Furthermore, future studies may wish to explore which, or what combination, of these policy changes have the strongest effect on the evaluator exclusion rate, and in what direction.

Fourth, this study proposed the possibility that the defensive attribution hypothesis is best suited for application within a closed system of relationships. Research studies should investigate this possibility in order to draw an informed conclusion on the scope and applicability of the hypothesis.

To ensure a spectrum of perspectives, researchers across the program evaluation, program management, and funding organization fields should pursue studies into each of these four research areas. Similarly, researchers from these fields should disseminate their findings via a diverse set of publications to ensure that program leaders, program evaluators, and funding organizations all have the opportunity to learn the knowledge, skills, and behaviors necessary to establish mutually beneficial relationships throughout all phases of the project cycle.

REFERENCES

- Alkin, M. C., & King, J. A. (2016). The Historical Development of Evaluation Use. *American Journal of Evaluation*, 37(4), 568–579.
- Alkin, M. C., & King, J. A. (2017). Definitions of Evaluation Use and Misuse, Evaluation Influence, and Factors Affecting Use. *American Journal of Evaluation*, 38(3), 434–450.
- Alkin, M., & Christie, C. (2004). An evaluation theory tree. In C. Alkin (Ed.), *Evaluation Roots*, 12–65.
- American Evaluation Association. (2010). An evaluation roadmap for a more effective government. Accessed here: http://comm.eval.org/coffee_break_webinars/viewdocument/an-evaluation-roadma
- American Evaluation Association. (2018). AEA Competencies. Available here: <https://www.eval.org/page/competencies>
- Bartle, P. (2007). The Nature of Monitoring and Evaluation: Definition and Purpose. Retrieved from: <http://cec.vcn.bc.ca/cmp/modules/mon-wht.htm>
- Baum, W. (1970). The project cycle. *Finance and Development*, 7(2), 2–13.
- Baum, W. (1978). World Bank Project Cycle. *Finance & Development*, 10–17.
- Biggs, S., & Smith, S. (2003). A Paradox of learning in project cycle management and the role of organizational culture. *World Development*, 31(10), 1743–1757.
- Bozzo, S. L. (2000). Evaluation Resources for Nonprofit Organizations. *Nonprofit Management & Leadership*, 10(4), 463–472.
- Bradley, E., Curry, L., & Devers, K. (2007). Qualitative Data Analysis for Health Services Research: Developing Taxonomy, Themes, and Theory. *Health Services Research*, 42(4), 1758-1772.
- Braithwaite, D. and Schrodt, P. (2014). *Engaging Theories in Interpersonal Communication: Multiple Perspectives*. Sage Publications.
- Burger, J. M. (1981). Motivational biases in the attribution of responsibility for an accident: A meta-analysis of the defensive attribution hypothesis. *Psychological Bulletin*, 90, 496-512.
- Camilleri, A. (2017). The Presentation Format of Review Score Information Influences Consumer Preferences through the Attribution of Outlier Reviews. *Journal of Interactive Marketing*, 39, 1-14.

- Carless, S., & Waterworth, R. (2012). The importance of ability and effort in recruiters' hireability decisions: An empirical examination of attribution theory. *Australian Psychologist*, 47, 232-237.
- Carman, J. G. (2007). Evaluation practice among community-based organizations. *American Journal of Evaluation*, 28(1), 60.
- Carman, J.G. (2009). Nonprofits, Funders, and Evaluation: Accountability in Action. *American Review of Public Administration*, 39(4), 374–390.
- Carman, J. G., Fredericks, K. A., & Introcaso, D. (2008). Government and Accountability: Paving the way for nonprofits and evaluation. In J. G. Carman & K. A. Fredericks (Eds.), *Nonprofits and evaluation. New Directions for Evaluation*, 119, 5–12.
- Chelimsky, E. (1997). The coming transformations in evaluation. In Chelimsky, E. and Shadish, W. (Eds.), *Evaluation for the 21st Century: A handbook*. Thousand Oaks, CA: SAGE.
- Chelimsky, E. (2013). Balancing Evaluation Theory and Practice in the Real World. *American Journal of Evaluation*, 34(1), 91–98.
- Chelimsky, E., Cordray, D., & Datta, L. (1989). Federal Evaluation: The Pendulum has Swung Too Far. *Evaluation Practice*, 10(2), 25–30.
- Chen, D., & Wang, Z. (2014). The effects of human resource attributions on employee outcomes during organizational change. *Social Behavior and Personality: An International Journal*, 42, 1431–1443.
- Chouinard, J. (2013). The Case for Participatory Evaluation in an Era of Accountability. *American Journal of Evaluation*, 34(2), 237-253.
- Cleland, D. I. (1985). A strategy for ongoing project evaluation. *Project Management Journal*, 16(3), 11–17.
- Cole, N. D. (2008). The effects of differences in explanations, employee attributions, type of infraction, and discipline severity on perceived fairness of employee discipline. *Canadian Journal of Administrative Sciences*, 25, 107.
- Cooke, B., & Kothari, U. (Eds.). (2000). *Participation: the new tyranny*. London: Zed Books.
- Coryn, C. L. S., Hattie, J. A., Scriven, M., & Hartmann, D. J. (2007). Models and mechanisms for evaluating government-funded research: An international comparison. *American Journal of Evaluation*, 28(4), 437–457.
- Coryn, C., & Scriven, M. (2008). Editor's Notes. *New Directions for Evaluation*, 118, 1–5.

- Cousins, J., & Earl, L. (Eds.). (1995). *Participatory evaluation in education: Studies in evaluation use and organizational learning*. London: Falmer.
- Cozzens, S. E. (2000). Assessing federally-supported academic research in the United States. *Research Evaluation*, 8(1), 5–10.
- Cresswell, J. W. (2014). *Research Design*, 4th edition. Thousand Oaks: Sage Publications.
- Cronbach, Lee J. & Associates (1980). *Toward reform of program evaluation*. San Francisco: Jossey-Bass.
- Dejoy, D. M. (1987). Supervisor attributions and responses for multicausal workplace accidents. *Journal of Occupational Accidents*, 9, 213-223.
- Fierro, L. and Christie, C. (2017). Evaluator and Program Manager Perceptions of Evaluation Capacity and Evaluation Practice. *American Journal of Evaluation*, 38(3), 376–392.
- Fitzpatrick, J. L. (1988). Roles of the Evaluator: A Formative Evaluation. *Evaluation Review*, 12(4), 449–461.
- Fontinha, R., José Chambel, M., & De Cuyper, N. (2012). HR attributions and the dual commitment of outsourced IT workers. *Personnel Review*, 41, 832–848.
- Forsyth, D., & McMillan, J. (1981). Attributions, Affect, and Expectations: A Test of Weiner's Three-Dimensional Model. *Journal of Educational Psychology*, 73(3), 393-403.
- Foundations for Evidence-Based Policymaking Act (2019). Public Law 115-435. Available here: <https://www.congress.gov/115/plaws/publ435/PLAW-115publ435.pdf>.
- Freeman, H. (1982). *A Federal Evaluation Agenda for the 1980's: Some Speculations and Suggestions*. Washington, DC.
- Forbes, D. P. (1998). Measuring the Unmeasurable: Empirical Studies of Nonprofit Organization Effectiveness from 1977 to 1997. *Nonprofit and Voluntary Sector Quarterly*, 27(2), 183–202.
- GAO (2011). *Government Performance: GPRA Modernization Act Provides Opportunities to Help Address Fiscal, Performance, and Management Challenges*. Available here: <https://www.gao.gov/assets/130/125777.pdf>
- Gonçalves, S. M. P., da Silva, S. A., Lima, M. L., & Meliá, J. L. (2008). The impact of work accidents experience on causal attributions and worker behaviour. *Safety Science*, 46, 992-1001.

- Government Performance and Results Act (1993). Public Law 103-62. Available here: <https://www.govinfo.gov/content/pkg/STATUTE-107/pdf/STATUTE-107-Pg285.pdf>.
- GPRA Modernization Act (2010). Public Law 111-352. Available here: <https://www.congress.gov/111/plaws/publ352/PLAW-111publ352.pdf>.
- Grants.gov (2020). Grant-making Agencies. Accessed here: <https://www.grants.gov/web/grants/learn-grants/grant-making-agencies.html>
- Grob, G. (2017). Evaluation Practice: Proof, Truth, Client Relationships, and Professional Growth. *American Journal of Evaluation*, 39(1), 123–132.
- Guba, E. G., & Stufflebeam, D. L. (1968) Evaluation: The process of stimulating, aiding, and abetting insightful action. In R. Ingle & W. Gephart (Eds.), *Problems in the training of educational researchers*. Bloomington, IN.
- Gyekye, S. A. (2010). Occupational safety management: The role of causal attribution. *International Journal of Psychology*, 45, 405–416.
- Gyekye, S.A. and Salminen, S. (2004). Causal Attributions of Ghanaian industrial Workers for Accident Occurrence. *Journal of Applied Social Psychology*, 34(11), 2324-2342.
- Gyekye, S.A. and Salminen, S. (2006). The self-defensive attribution hypothesis in the work environment: Co-workers' perspectives. *Safety Science*, 44, 157-168.
- Gyekye, S.A. and Salminen, S. (2007). Religious Beliefs and Responsibility Attributions for Industrial Accidents among Ghanaian Workers. *Journal for the Study of Religion*, 20(1), 73–86.
- Hamilton, V. L. (1986). Chains of command: Responsibility attribution in hierarchies. *Journal of Applied Social Psychology*, 16, 118-138.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Henry, G. T., & Mark, M. M. (2003). Toward an agenda for research on evaluation. In C. A. Christie (Ed.), *The practice-theory relationship in evaluation: New directions for evaluation* (Vol. 97, 69-80). San Francisco: Jossey-Bass.
- Hewett, R., Shantz, A., Mundy, J., & Alfes, K. (2018). Attribution theories in Human Resource Management research: a review and research agenda. *The International Journal of Human Resource Management*, 2991, 87-126.
- Howell, E. M., & Yemane, A. (2006). An Assessment of Evaluation Designs: Case Studies of 12 Large Federal Evaluations. *American Journal of Evaluation*, 27(2), 219–236.

- Huberty, C. J. (1988). Another perspective on the role of an internal evaluator. *Evaluation Practice*, 9(4), 25–32.
- Kane, M. & Trochim, W. M. K. (2007). *Concept Mapping for Planning and Evaluation*. Thousand Oaks, CA: Sage.
- Kaushik, V. & Walsh, C. (2019). Pragmatism as a Research Paradigm and Its Implications for Social Work Research. *Social Sciences*, 8(9), 1-17.
- Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (ed.), *Nebraska Symposium on Motivation* (Volume 15, pp. 192-238). Lincoln: University of Nebraska Press.
- Kelley, H.H. (1973). The processes of causal attribution. *American Psychologist*, 28(2), 107-128.
- Kelley, H. H., & Michela, J. L. (1980). Attribution theory and research. *Annual Review of Psychology*, 31, 457-501.
- Knorr, K.D. (1977). Policymakers' use of social science knowledge: Symbolic or instrumental? In C. H. Weiss (Ed.), *Using social research in public policy making* (pp.165-182). Lexington, MA: Lexington Books.
- Knorr, K. D. (1980). The gap between knowledge and policy. In S. Nagel (Ed.), *Improving policy analysis* (219–233). Beverly Hills, CA: Sage.
- Kouabenan, D. R., Gilibert, D., Medina, M., & Bouzon, F. (2001). Hierarchical position, gender, accident severity, and causal attribution. *Journal of Applied Social Psychology*, 31, 553-575.
- Lacroix, D. V., & Dejoy, D. M. (1989). Causal attributions to effort and supervisory response to workplace accidents. *Journal of Occupational Accidents*, 11, 97-109.
- Levy, P. E., Cawley, B. D., & Foti, R. J. (1998). Reactions to appraisal discrepancies: Performance ratings and attributions. *Journal of Business and Psychology*, 12, 437–455.
- Loo, R. (1985). Evaluation in the project management environment. *Project Management Journal*, 16(3), 36–41.
- Manusov, V. and Spitzberg, B. (2008). Attribution Theory. In: *Engaging Theories in Interpersonal Communication: Multiple Perspectives*. Edited by: Leslie A. Baxter & Dawn O. Braithwaite.
- Mark, M. (2012). Program Lifecycle Stage as a Guide to Evaluation Decision-making: Benefits, Limits, Alternatives, and Future Decisions. *American Journal of Evaluation*, 33(2), 277–280.

- Mark, M., Henry, G., & Julnes, G. (2000). *Review of Evaluation: An Integrated Framework for Understanding, Guiding, and Improving Policies and Programs*. San Francisco: Jossey-Bass.
- Martinko, M. J. (2004). *Attribution Theory in the Organizational Sciences: Theoretical and Empirical Contributions*. Greenwich, Conn: IAP-Information Age Publishing.
- Martinko, M. & Mackey, J. (2019). Attribution theory: An introduction to the special issue. *Journal of Organizational Behavior*, 40, 523-527.
- Mayo, M., & Mallin, M. L. (2010). The impact of sales failure on attributions made by 'resource-challenged' and 'resource-secure' salespeople. *Journal of Marketing Theory and Practice*, 18, 233-248.
- McLaughlin, J. A., & Jordan, G. B. (2004). Using logic models. In J. S. Wholey, H. P. Hatry, & K. E. Newcomer (Eds.), *Handbook of practical program evaluation* (2nd ed., pp. 7-32). San Francisco: Jossey-Bass.
- Mitchell, G. (2014). Strategic Responses to Resource Dependence Among Transnational NGOs Registered in the United States. *Voluntas*, 25(1), 67-91.
- Mitchell, T. R., & Wood, R. E. (1980). Supervisor's responses to subordinate poor performance. A test of an attributional model. *Organizational Behavior and Human Performance*, 25, 125-138.
- Nature. (2006). Evaluate this, 440, 1. Retrieved from <http://dx.doi.org/10.1038/440001b>.
- Newcombe, R. D., & Rutter, D. R. (1982). Ten reasons why ANOVA theory and research fail to explain attribution processes. *Current Psychological Reviews*, 2, 95-108.
- Newcomer, K. E. (Ed.). (1997). Using performance measurement to improve public and nonprofit programs. *New Directions for Evaluation*, 75. San Francisco: Jossey-Bass.
- Newcomer, K. E., Hatry, H. P., & Wholey, J. S. (2004). Meeting the need for practical evaluation approaches: An introduction. In J. S. Wholey, H. P. Hatry & K. E. Newcomer (Eds.), *Handbook of practical program evaluation* (2nd ed., pp. 1-10). San Francisco, CA: Jossey-Bass.
- Niskanen, T. (1994). Assessing the safety environment in work organization of road maintenance jobs. *Accident Analysis and Prevention*, 26, 27-39.
- O'Reilly III, C. A. (1980). Individuals and information overload in organizations: is more necessarily better?. *Academy of Management Journal*, 23(4), 684-696.
- Orth, U., Stoeckl, A., Crouch, R., Joëlle, B., Cavicchi, A., Faraoni, M., Larreina, M., Lecat, B., Olsen, J., Rodriguez Santos, C., Santini, C., and Wilson, D. (2012). Using attribution theory to explain tourists' attachments to place-based brands. *Journal of Business Research*, 65(9), 1321-1327.

- Pardo, C., & Alfonso, W. (2017). Applying "attribution theory" to determine the factors that lead to the failure of entrepreneurial ventures in Colombia. *Journal of Small Business and Enterprise Development*, 24(3), 562-584.
- Patton, M. Q. (1978). *Utilization-focused evaluation*. Beverly Hills: Sage Publications.
- Patton, M. Q. (1986). *Utilization-focused evaluation (2nd ed.)*. Newbury Park, CA: Sage Publications.
- Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text (3rd ed.)*. Thousand Oaks, CA: Sage Publications.
- Patton, M. Q. (2002). *Qualitative Research and Evaluation Methods*, 3d Edition. Thousand Oaks, CA: Sage Publications
- Patton, M. Q. (2006). Evaluation for the Way We Work. *Nonprofit Quarterly*, 13(1), 28-33.
- Patton, M. Q. (2011). *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York: The Guilford Press.
- Peace Corps (2020). Program Manager Position Description. Available here: <https://fj.usembassy.gov/wp-content/uploads/sites/180/Suva-2020-01-Peace-Corps-Program-Manager.pdf>
- PMI (2018). Accessed here: <https://www.pmi.org/about/learn-about-pmi>.
- Preskill, H. (2014). Now for the hard stuff: Next steps in ECB research and practice. *American Journal of Evaluation*, 35, 116–119.
- Preskill, H., & Torres, R. T. (Eds.). (2001). *The Learning Dimension of Evaluation Use*. *New Directions for Evaluation*, (88). San Francisco: Jossey-Bass.
- Rippey, R.M. (1973) *Studies in transactional evaluation*. Berkeley, CA: McCutchan.
- Ross, L. D. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173-220). New York: Academic Press.
- Rossi, P. H., & Freeman, H. E. (1993). *Evaluation: A systemic approach (5th ed.)*. Newbury Park, CA: Sage.
- Rotter, J.B. (1954). *Social learning and clinical psychology*. NY: Prentice-Hall.
- Rotter, J.B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 80(1), 1-28.
- Rotter, J. B. (1975). Some problems and misconceptions related to the construct of internal versus external control of reinforcement. *Journal of Consulting and Clinical Psychology*, 43(1), 56–67.

- Salminen, S. (1992). Defensive attribution hypothesis and serious occupational accidents. *Psychological Reports, 70*, 1195-1199.
- Sanders, J. R. (2003). Mainstreaming evaluation. *New Directions for Evaluation, (99)*, 3-6. San Francisco: Jossey-Bass.
- Sanders, K., & Yang, H. (2016). The HRM process approach: The influence of employees' attribution to explain the HRM-performance relationship. *Human Resource Management, 55*, 201–217.
- Saxena, D. & Lamest, M. (2018). Information Overload and Coping Strategies in the Big Data Context: Evidence from the Hospitality Sector. *Journal of Information Science, 44(3)*, 287–297.
- Scheirer, M. A. (2012). Planning Evaluation through the Program Life Cycle. *American Journal of Evaluation, 33(2)*, 263–277.
- Scherer, C.S. (2015). *Grant Evaluation: What Private Foundations Do and Why They Do It: An Examination of Grant Assessment Motivations and Practices*. [Unpublished doctoral dissertation]. University of Pittsburgh.
- Schwandt, T. A. (2015). *Evaluation foundations revisited: Cultivating a life of the mind for practice*. Stanford, CA: Stanford University Press.
- Shaver, K. G. (1970a). Defensive attribution: Effects of severity and relevance on the responsibility assigned for an accident. *Journal of Personality and Social Psychology, 14*, 101-113.
- Shaver, K. G. (1970b). Redress and conscientiousness in the attribution of responsibility for accidents. *Journal of Experimental Psychology, 6*, 100- 110.
- Shulha, L. M., & Cousins, J. B. B. (1997). Evaluation Use: Theory, Research, and Practice since 1986. *American Journal of Evaluation, 18(3)*, 195–208.
- Smith, M.F. (1994). From the Editor. *Evaluation Practice, 15(3)*, 213.
- Sofaer, S. (1999). Qualitative Methods: What Are They and Why Use Them? *Health Services Research, 34(5)*, 1101–1118.
- Sparrow, P. (1999). Strategy and Cognition: Understanding the Role of Management Knowledge Structures, Organizational Memory and Information Overload. *Creativity and Innovation Management, 8(2)*, 140–148.
- Struthers, C. W., Weiner, B., & Allred, K. (1998). Effects of causal attributions on personnel decisions: A social motivation perspective. *Basic and Applied Social Psychology, 20*, 155–166.

- Stone, M., Bigelow, B., & Crittenden, W. (1999). Research on Strategic Management in Nonprofit Organizations: Synthesis, Analysis, and Future Directions. *Administration & Society*, 31(3), 378–423.
- Stone, M. M., & Cutcher-Gershenfeld, S. (2001). Challenge of measuring performance in nonprofit organizations. In P. Flynn & V. A. Hodgkinson (Eds.), *Measuring the Impact of the Nonprofit Sector* (pp. 33–58). New York: Kluwer Academic/Plenum.
- Stufflebeam, D. (1969) Evaluation as Enlightenment for Decision Making. In Walcott, A. Beaty (ed.) *Improving Educational Assessment and an Inventory of Measures of Affective Behavior*. Washington, D.C.: Assoc. for Supervision and Curriculum Development.
- Stufflebeam, D. (2001). Interdisciplinary Ph.D. programming in evaluation. *American Journal of Evaluation*, 22, 445-455.
- Tandung, J. C. (2016). The link between HR attributions and employees' turnover intentions. *Gajah Mada International Journal of Business*, 18, 55.
- The Evidence-Based Policy Commission Act of 2015, H.R.1831, 114th Congress (2015). Available here: <https://www.govtrack.us/congress/bills/114/hr1831/text>.
- Thomson, D. (2010). Exploring the Role of Funders' Performance Reporting Mandates in Nonprofit Performance Measurement. *Nonprofit and Voluntary Sector Quarterly*, 39(4), 611–629.
- USAID (2015). Program Officer Position Description. Available here: <https://www.usaid.gov/documents/1860/program-officer>.
- Vidmar, N., & Crinklaw, L. D. (1974). Attributing responsibility for an accident: A methodological and conceptual critique. *Canadian Journal of Behavioural Sciences*, 6, 112- 130.
- Volkov, B. B. (2011). Beyond Being an Evaluator: The multiplicity of Roles of the Internal Evaluator. *New Directions for Evaluation*, 132(1), 25–42.
- Walster, E. (1966). Assignment of responsibility for an accident. *Journal of Personality and Social Psychology*, 3, 73-79.
- Ward, J.L. (2011). *Dictionary of Project Management Terms*, 3rd Edition. Arlington: CRC Press.
- Wargo, M. J. (1994). The Impact of the President's Reinvention Plan on Evaluation. *Evaluation Practice*, 15(1), 63–72.
- Wargo, M. J. (1995). The impact of federal government reinvention on federal evaluation activity. *Evaluation Practice*, 16(3), 227–237.

- Weiner, B. (1972). Attribution Theory, Achievement Motivation, and the Educational Process. *Review of Educational Research*, 42(2), 203–215.
- Weiner, B. (1979). A Theory of Motivation for Some Classroom Experiences. *Journal of Educational Psychology*, 71(1), 3-25.
- Weiner, B. (1986). An attributional theory of motivation and emotion. New York: SpringerVerlag.
- Weiner, B. (2000). Attributional thoughts about consumer behavior. *Journal of Consumer Research*, 27, 382-387.
- Weiner, B. (2019). Wither Attribution Theory? *Journal of Organizational Behavior*, 40, 603-604.
- Weiss, C. H. (1972). Utilization of evaluation: Toward comparative study. In C. H. Weiss (Ed.), *Evaluating action programs: Readings in social action and education* (pp. 318–326). Boston, MA: Allyn and Bacon.
- Weiss, C. H. (1988). Reports on Topic Areas: Evaluation for Decisions: Is Anybody There? Does Anybody Care? *American Journal of Evaluation*, 9(1), 5–19.
- Wholey, J. (1981). Using evaluation to improve program performance. In R. A. Levine, M. A. Solomon, G. M. Hellstern, & H. Wollman (Eds.), *Evaluation Research and Practice: Comparative and International Perspectives* (pp. 92-106). Beverly Hills, CA: Sage.
- Wholey, J. (1983). *Evaluation and Effective Public Management*. Boston: Little, Brown.
- Wood, J. T. (2008). *Communication mosaics* (5th ed.). Belmont, CA: Thompson-Wadsworth.
- Wye, C. & Sonnichsen, R. (1992). Another look at the future of program evaluation in the federal government: Five views. *Evaluation Practice*, 13(3), 185-195.
- Zuckerman, M. (1979). Attribution of success and failure revisited: The motivational bias is alive and well in attribution theory. *Journal of Personality*, 47, 245-287.

APPENDIX A

QUESTIONNAIRE – PROGRAM EVALUATOR VERSION

Start of Block: Information Letter

Q1 Hello!

My name is Matthew Gallagher, and I am a PhD Candidate at Arizona State University in the Community Resources and Development program. I have also been a member of the American Evaluation Association since 2009.

PURPOSE: My research explores the topic of evaluator exclusion during the planning and design phase of projects/programs. I am conducting this survey to gather data from experts in the fields of program evaluation and program leadership (i.e. Program Managers and Program Directors). The purpose of this survey research is four-fold:

1. To identify the extent to which evaluator exclusion from a program's planning and design phase occurs;
2. To identify a comprehensive set of factors that influence evaluator exclusion from the planning and design phase of programs;
3. To organize these factors into a taxonomy of internal and external locus of control; and
4. To explore the extent to which program leaders and evaluators demonstrate defensive attributions when rating these factors' influence on evaluator exclusion from the planning and design phase of programs

YOUR ROLE: If you choose to participate, you will be asked questions about when and why you were excluded from the planning and design phase of programs. By participating in the survey, you are providing your informed consent.

YOUR TIME: This survey will take approximately 10 minutes to complete.

VOLUNTARY: Your participation in this survey is voluntary. You have the right not to answer any question and to quit the survey at any time. If you choose not to participate or to withdraw from the survey, there will be no penalty. Your decisions whether or not to participate in this study will not affect your status at your organization. You must be 18 or older to participate in the study. Once you complete the last question of the survey, it will automatically be submitted. No further action will be required from you.

WHY PARTICIPATE: No immediate benefits to you are anticipated through participating in this research. However, your responses may benefit other evaluators and program leaders in the future by helping us learn more about the extent to which evaluators are excluded from the planning and design phase of programs, and why. Your responses will be analyzed and results will be reported in my PhD dissertation. The results of this research may also be used in presentations or publications.

MINIMAL RISK: This study has been approved by Arizona State University's Institutional Review Board (IRB). There are no foreseeable risks or discomforts associated with your participation. The information you provide will only be accessed by me. Quotes may be used in my paper but your identity will remain confidential. Your personal and identifying information will be kept confidential and will not be disclosed to anyone outside of the research study. Your responses will be assigned an identification number and I am the only one who will have access to your assigned identification number. This identification list will be cleared from my files three years after completion of my dissertation.

QUESTIONS OR CONCERNS REGARDING THIS RESEARCH?: If you have any questions or concerns about the survey, if you have any questions about your rights as a subject/participant in this research, or if you feel you have been placed at risk, you can contact me at mgallag9@asu.edu. You may also contact the Principal Investigator, Dr. Jesse Lecy, at jlecy@asu.edu.

Thank you in advance for participating. I look forward to seeing your honest and comprehensive perspective.

Sincerely,
Matthew Gallagher
PhD Candidate, Community Resources and Development
Arizona State University
mgallag9@asu.edu

End of Block: Information Letter

Start of Block: Definitions

Q2 Definitions

Q3 The following definitions set the foundation for this research:

1. **Program:** A group of related projects managed in a coordinated way to obtain benefits not available from managing the projects individually, and may also include elements of on-going, operational work (Ward, 2011)
2. **Program Leader:** The person (or people) who oversees the day-to-day implementation of a program, provides strategic direction for its implementation, and/or

with whom all final decisions rest. May be referred to as a Program Manager or Program Director

3. **Evaluation:** The systematic collection and use of quantitative and qualitative data to aid program teams in developing knowledge about and managing program activities (Scheirer, 2012)

4. **Evaluator:** The person implementing the responsibilities of the evaluation profession

5. **Project Cycle:** The progressive phases of a program or project that begin with its planning and design, through implementation of activities to achieve planned outcomes, to evaluation and reporting of outputs, outcomes, and impact (Baum, 1970 and 1978)

6. **Program Planning and Design Phase:** All activities that occur prior to program launch and implementation. For new programs, this typically includes the initial identification of a program, the preparation of a logic model and evaluation plan, and development of a staffing plan, budget, and timeline. For ongoing programs that implement a next iteration, the planning and design phase is the period of time when these deliverables are being revised prior to re-launch

End of Block: Definitions

Start of Block: Exclusion Block

Q4 **Evaluator Exclusion**

To guide your responses for this section, please think back to your **three most recent** program experiences as an Evaluator. If you do not have at least three program experiences as an Evaluator, simply use your one or two most recent experiences to guide your responses.

In this section, the term '**excluded**' includes if you were actively excluded from a program's planning and design phase, if you were simply not included in the program planning and design phase, and/or if you were hired/contracted after the completion of the program planning and design phase.

Q5 During your three most recent experiences as an Evaluator, were you ever **excluded** from the planning and design phase of the projects/programs?

As a reminder, a program's planning and design phase covers all activities that occur **prior** to program launch and implementation.

- Yes
- No
- Unsure

Display This Question:

If During your three most recent experiences as an Evaluator, were you ever excluded from the planni... = No

Q6 You indicated that you do **not** have recent experience being excluded as an Evaluator from the planning and design phase of a project/program. If you would like to change this response, please use the 'Back' button at the bottom of the survey. If this response is correct, select 'Okay' and you will be taken to the end of this survey.

- Okay

Skip To: End of Survey If You indicated that you do not have recent experience being excluded as an Evaluator from the plan... = Okay

Q7 Select the phrase that best describes your exclusion experience:

- In my three most recent experiences, I have only experienced being excluded from the planning and design phase of federal/federally-funded projects or programs.
- In my three most recent experiences, I have only experienced being excluded from the planning and design phase of non-federal/non-federally-funded projects or programs (e.g. foundation-funded programs, non-profit programs, for-profit programs).
- In my three most recent experiences, I have experienced being excluded from the planning and design phase of both federal/federally-funded AND non-federal/federally-funded projects or programs.

End of Block: Exclusion Block

Start of Block: Exclusion Factors Block

Q8 Evaluator Exclusion Factors

In this section, you will be asked to review a comprehensive set of factors that may have influenced your exclusion from the planning and design phase of a project/program.

Please note that if you were excluded from the planning and design phase in more than one of your three most recent experiences, please focus on just one of your experiences when responding to the factors presented below. Additionally, if you indicated that you have experienced this exclusion with a federal/federally-funded project or program, please focus on that experience when responding to the factors presented below.

Note: The factors presented below are randomized for each respondent.



Q9 How influential were the following factors when you, the evaluator, were excluded from the planning and design phase of the project/program?

	No influence	Very little influence	Some influence	Influential	Very influential	Not applicable / Unable to assess
Organization funding the program did not require evaluation activities	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Program leadership had no knowledge of or training on the benefits of evaluation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Program leadership had belief/understanding that evaluation only needs to occur at the end of a project/program	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Program leadership preferred to have program staff (who are not evaluators) manage evaluation activities	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Previous programs/projects had been successful without an evaluator, so there was no need to change	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Program's hiring process too difficult/lengthy to consider hiring an evaluator	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Program leadership followed a supervisor's (or someone else higher up) decision to not hire an evaluator	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Adding/Managing an evaluator would have added to an already heavy workload	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Evaluation findings would not be used in program decision-making, marketing, and/or funding decisions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Inadequate skillset of program leadership to interpret and/or apply evaluation results/findings	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Evaluator too expensive to include in budget	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Program leadership skeptical of the benefits of evaluation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Inadequate skillset of evaluators to provide benefits during the planning and design phase of programs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Program leadership had negative previous experience with evaluators	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Evaluator may discover negative program results	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Paperwork Reduction Act (PRA) processes mandated by Office of Management and Budget (OMB) too burdensome for designing/implementing evaluations	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Other (please specify):	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Other (please specify):	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

End of Block: Exclusion Factors Block

Start of Block: Demographics

Q10 **Demographics**

Q11 Are you:

- Male
 - Female
 - Non-binary / third gender
 - Prefer not to say
-

Q12 What is your highest level of education completed?

- Secondary School
 - Bachelor's Degree (4 years)
 - Associate's Degree (2 years)
 - Master's Degree
 - Doctorate Degree
 - Other (Please specify):

 - Prefer not to say
-



Q13 Indicate your total number of years of experience as an Evaluator:

[Please note that up to one decimal place will be accepted.]

End of Block: Demographics

APPENDIX B

QUESTIONNAIRE – PROGRAM LEADER VERSION

Q1 Hello!

My name is Matthew Gallagher, and I am a PhD Candidate at Arizona State University in the Community Resources and Development program. I have also been a member of the American Evaluation Association since 2009.

PURPOSE: My research explores the topic of evaluator exclusion during the planning and design phase of projects/programs. I am conducting this survey to gather data from experts in the fields of program evaluation and program leadership (i.e. Program Managers and Program Directors). The purpose of this survey research is four-fold:

1. To identify the extent to which evaluator exclusion from a program's planning and design phase occurs;
2. To identify a comprehensive set of factors that influence evaluator exclusion from the planning and design phase of programs;
3. To organize these factors into a taxonomy of internal and external locus of control; and
4. To explore the extent to which program leaders and evaluators demonstrate defensive attributions when rating these factors' influence on evaluator exclusion from the planning and design phase of programs

YOUR ROLE: If you choose to participate, you will be asked questions about when and why evaluators were excluded from the planning and design phase of programs. By participating in the survey, you are providing your informed consent.

YOUR TIME: This survey will take approximately 10 minutes to complete.

VOLUNTARY: Your participation in this survey is voluntary. You have the right not to answer any question and to quit the survey at any time. If you choose not to participate or to withdraw from the survey, there will be no penalty. Your decisions whether or not to participate in this study will not affect your status at your organization. You must be 18 or older to participate in the study. Once you complete the last question of the survey, it will automatically be submitted. No further action will be required from you.

WHY PARTICIPATE: No immediate benefits to you are anticipated through participating in this research. However, your responses may benefit other evaluators and program leaders in the future by helping us learn more about the extent to which evaluators are excluded from the planning and design phase of programs, and why. Your responses will be analyzed and results will be reported in my PhD dissertation. The results of this research may also be used in presentations or publications.

MINIMAL RISK: This study has been approved by Arizona State University's Institutional Review Board (IRB). There are no foreseeable risks or discomforts associated with your participation. The information you provide will only be accessed by me. Quotes may be used in my paper but your identity will remain confidential. Your personal and identifying information will be kept confidential and will not be disclosed to anyone outside of the research study. Your responses will be assigned an identification number and I am the only one who will have access to your assigned identification number. This identification list will be cleared from my files three years after completion of my dissertation.

QUESTIONS OR CONCERNS REGARDING THIS RESEARCH?: If you have any questions or concerns about the survey, if you have any questions about your rights as a subject/participant in this research, or if you feel you have been placed at risk, you can contact me at mgallag9@asu.edu. You may also contact the Principal Investigator, Dr. Jesse Lecy, at jlecy@asu.edu.

Thank you in advance for participating. I look forward to seeing your honest and comprehensive perspective.

Sincerely,
Matthew Gallagher
PhD Candidate, Community Resources and Development
Arizona State University
mgallag9@asu.edu

End of Block: Information Letter

Start of Block: Definitions

Q2 Definitions

Q3 The following definitions set the foundation for this research:

1. **Program:** A group of related projects managed in a coordinated way to obtain benefits not available from managing the projects individually, and may also include elements of on-going, operational work (Ward, 2011)
2. **Program Leader:** The person (or people) who oversees the day-to-day implementation of a program/project, provides strategic direction for its implementation, and/or with whom all final decisions rest. May be referred to as a Program Manager or

Program Director

3. **Evaluation:** The systematic collection and use of quantitative and qualitative data to aid program teams in developing knowledge about and managing program activities (Scheirer, 2012)

4. **Evaluator:** The person implementing the responsibilities of the evaluation profession

5. **Project Cycle:** The progressive phases of a program or project that begin with its planning and design, through implementation of activities to achieve planned outcomes, to evaluation and reporting of outputs, outcomes, and impact (Baum, 1970 and 1978)

6. **Program Planning and Design Phase:** All activities that occur prior to program launch and implementation. For new programs, this typically includes the initial identification of a program, the preparation of a logic model and evaluation plan, and development of a staffing plan, budget, and timeline. For ongoing programs that implement a next iteration, the planning and design phase is the period of time when these deliverables are being revised prior to re-launch

End of Block: Definitions

Start of Block: Exclusion Block

Q4 **Evaluator Exclusion**

To guide your responses for this section, please think back to your **three most recent** program experiences as a Program Leader. If you do not have at least three experiences as a Program Leader, simply use your one or two most recent experiences to guide your responses.

In this section, the term '**excluded**' includes if an Evaluator was actively excluded from the program planning and design phase, if an Evaluator was simply not included in the program planning and design phase, and/or if an Evaluator was hired/contracted after the completion of the program planning and design phase.

Q5 During your three most recent experiences as a Program Leader, was an Evaluator ever **excluded** from the planning and design phase of the programs?

As a reminder, a program's planning and design phase covers all activities that occur **prior** to program launch and implementation.

- Yes
- No
- Unsure

Display This Question:

If During your three most recent experiences as a Program Leader, was an Evaluator ever excluded fro... = No

Q6 You indicated that you do **not** have recent experience excluding an Evaluator from the planning and design phase of a project/program that you have led. If you would like to change this response, please use the 'Back' button at the bottom of the survey. If this response is correct, select 'Okay' and you will be taken to the end of this survey.

- Okay

Skip To: End of Survey If You indicated that you do not have recent experience excluding an Evaluator from the planning and... = Okay

Q7 Select the phrase that best describes your recent experiences:

- In my three most recent experiences, I have only not included an Evaluator in the planning and design phase of federal/federally-funded projects or programs.
- In my three most recent experiences, I have only not included an Evaluator in the planning and design phase of non-federal/non-federally-funded projects or programs (e.g. foundation-funded programs, non-profit programs, for-profit programs).
- In my three most recent experiences, I have not included an Evaluator in the planning and design phase of both federal/federally-funded AND non-federal/federally-funded projects or programs.

End of Block: Exclusion Block

Start of Block: Exclusion Factors Block

Q8 **Evaluator Exclusion Factors**

In this section, you will be asked to review a comprehensive set of factors that may have influenced the decision to not include an Evaluator in the planning and design phase of a project/program.

Please note that if you did not include an Evaluator in the planning and design phase in more than one of your three most recent experiences, please focus on just one of your experiences when responding to the factors presented below. Additionally, if you indicated that you have done this in a federal/federally-funded project or program, please focus on that experience when responding to the factors presented below.

Note: The factors presented below are randomized for each respondent.



Q9 How influential were the following factors when you, the program leader, did not include an evaluator in the planning and design phase of the project/program?

	No influence	Very little influence	Some influence	Influential	Very influential	Not applicable / Unable to assess
Organization funding the program did not require evaluation activities	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Program leadership had no knowledge of or training on the benefits of evaluation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Program leadership had belief/understanding that evaluation only needs to occur at the end of a project/program	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Program leadership preferred to have program staff (who are not evaluators) manage evaluation activities	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Previous programs/projects had been successful without an evaluator, so there was no need to change	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Program's hiring process too difficult/lengthy to consider hiring an evaluator	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Program leadership followed a supervisor's (or someone else higher up) decision to not hire an evaluator	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Adding/Managing an evaluator would have added to an already heavy workload	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Evaluation findings would not be used in program decision-making, marketing, and/or funding decisions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Inadequate skillset of program leadership to interpret and/or apply evaluation results/findings	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Evaluator too expensive to include in budget	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Program leadership skeptical of the benefits of evaluation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Inadequate skillset of evaluators to provide benefits during the planning and design phase of programs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Program leadership had negative previous experience with evaluators	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Evaluator may discover negative program results	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Paperwork Reduction Act (PRA) processes mandated by Office of Management and Budget (OMB) too burdensome for designing/implementing evaluations	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Other (please specify):	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Other (please specify):	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q10 **Demographics**

Q11 Are you:

- Male
 - Female
 - Non-binary / third gender
 - Prefer not to say
-

Q12 What is your highest level of education completed?

- Secondary School
- Bachelor's Degree (4 years)
- Associate's Degree (2 years)
- Master's Degree
- Doctorate Degree
- Other (Please specify):

- Prefer not to say
-



Q13 Indicate your total number of years of experience as a Program Leader:

[Please note that up to one decimal place will be accepted.]

Q14 Are you a certified Project Management Professional through the Project Management Institute?

- Yes
 - No
 - Unsure
-

Q15 Are you a member of a local chapter of the Project Management Institute?

- Yes
 - No
 - Unsure
-

Display This Question:

If Are you a member of a local chapter of the Project Management Institute? = Yes

Q16 Please name the local PMI chapter of which you are a member:

End of Block: Demographics

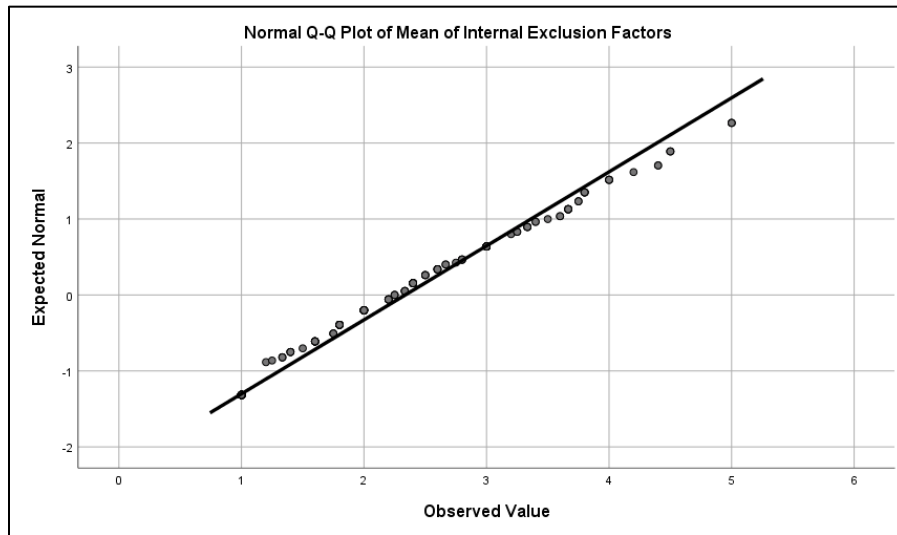
APPENDIX C
CHI-SQUARE TEST RESULTS

Respondent Role and Exclusion Status - Crosstab					
			Exclusion Status		Total
			Yes	No	
Role	Program Evaluator	Count	187	62	249
		Expected Count	185.7	63.3	249.0
	Program Leader	Count	42	16	58
		Expected Count	43.3	14.7	58.0
Total		Count	229	78	307
		Expected Count	229.0	78.0	307.0

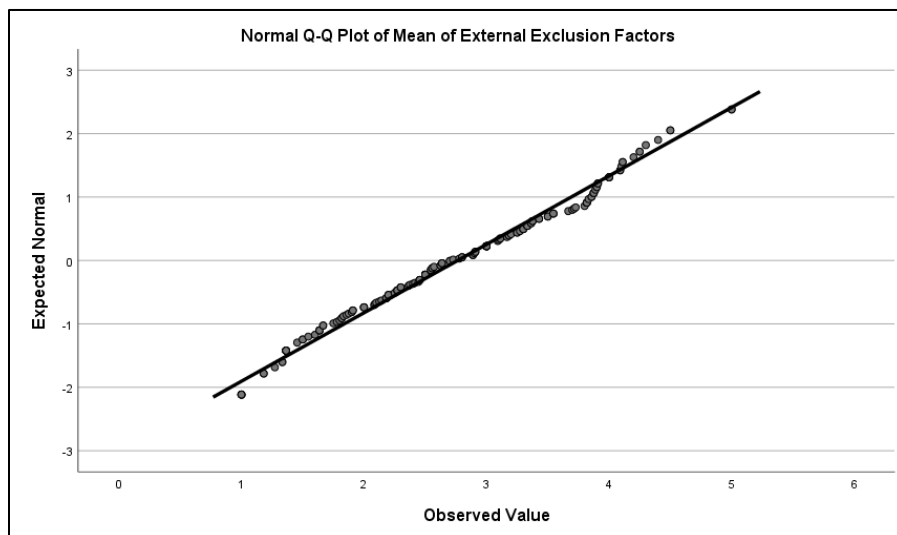
Respondent Role and Exclusion Status - Chi-Square Tests					
	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.179 ^a	1	0.672		
Continuity Correction ^b	0.065	1	0.798		
Likelihood Ratio	0.177	1	0.674		
Fisher's Exact Test				0.738	0.393
Linear-by-Linear Association	0.179	1	0.673		
N of Valid Cases	307				

APPENDIX D
TESTS OF NORMALITY

	Kolmogorov-Smirnova			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Mean of Internal Exclusion Factors	0.097	169	0.001	0.945	169	0.000



	Kolmogorov-Smirnova			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Mean of External Exclusion Factors	0.065	173	0.075	0.982	173	0.022



APPENDIX E
PAIRED T-TEST RESULTS

Paired Samples Statistics - Program Evaluators					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mean of Internal Exclusion Factors	2.2473	131	1.00233	0.08757
	Mean of External Exclusion Factors	2.6058	131	0.89596	0.07828

Paired Samples Test - Program Evaluators									
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Mean of Internal Exclusion Factors - Mean of External Exclusion Factors	-0.35843	0.73332	0.06407	-0.48519	-0.23168	-5.594	130	0.000

Paired Samples Statistics - Program Leaders					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 2	Mean of Internal Exclusion Factors	2.4828	31	1.02508	0.18411
	Mean of External Exclusion Factors	3.1655	31	0.73661	0.13230

Paired Samples Test - Program Leaders									
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 2	Mean of Internal Exclusion Factors - Mean of External Exclusion Factors	-0.68267	0.90491	0.16253	-1.01459	-0.35075	-4.200	30	0.000

APPENDIX F

INDEPENDENT SAMPLE T-TEST RESULTS – ALL EXCLUSION FACTORS

Group Statistics – Individual Factors					
Factors Internal to the Evaluation Field					
Factor	Role	N	Mean	Std. Deviation	Std. Error Mean
Evaluator too expensive to include in budget	Program Evaluator	112	2.38	1.472	0.139
	Program Leader	28	2.89	1.343	0.254
Program leadership skeptical of the benefits of evaluation	Program Evaluator	119	2.62	1.426	0.131
	Program Leader	26	2.77	1.366	0.268
Inadequate skillset of evaluators to provide benefits during the program planning and design phase	Program Evaluator	113	1.71	1.200	0.113
	Program Leader	21	2.62	1.203	0.263
Program leadership had negative previous experience with evaluators	Program Evaluator	92	1.87	1.215	0.127
	Program Leader	24	1.96	1.268	0.259
Evaluator may discover negative program results	Program Evaluator	115	2.37	1.429	0.133
	Program Leader	25	1.96	1.457	0.291
Factors External to the Evaluation Field					
Factor	Role	N	Mean	Std. Deviation	Std. Error Mean
Organization funding the program did not require evaluation activities	Program Evaluator	112	2.35	1.444	0.136
	Program Leader	29	3.90	1.205	0.224
Program leadership had no knowledge of or training on the benefits of evaluation in the planning and design phase	Program Evaluator	120	3.33	1.480	0.135
	Program Leader	28	3.07	1.562	0.295
Program leadership had belief/understanding that evaluation only needs to occur at the end of a program	Program Evaluator	123	3.52	1.462	0.132
	Program Leader	29	2.79	1.449	0.269
Program leadership had program staff (who are not evaluators) plan and/or manage evaluation activities	Program Evaluator	120	2.68	1.512	0.138
	Program Leader	28	3.75	1.531	0.289
Previous programs had been successful without an evaluator, so there was no need to change	Program Evaluator	109	2.49	1.372	0.131
	Program Leader	30	3.67	1.295	0.237
Program's hiring process too difficult/lengthy to consider hiring an evaluator	Program Evaluator	103	2.08	1.363	0.134

	Program Leader	25	2.52	1.558	0.312
Program leadership followed a supervisor's (or someone else higher up) decision to not hire an evaluator	Program Evaluator	97	2.11	1.413	0.143
	Program Leader	25	3.36	1.524	0.305
Adding/Managing an evaluator would have added to an already heavy workload	Program Evaluator	120	2.48	1.384	0.126
	Program Leader	29	2.76	1.405	0.261
Evaluation findings would not be used in program decision-making, marketing, and/or funding decisions	Program Evaluator	116	2.31	1.398	0.130
	Program Leader	26	3.04	1.399	0.274
Inadequate skillset of program leadership to interpret and/or apply evaluation results/findings	Program Evaluator	115	2.82	1.565	0.146
	Program Leader	29	3.17	1.391	0.258
PRA processes mandated by OMB too burdensome for designing/implementing evaluations	Program Evaluator	81	1.46	0.988	0.110
	Program Leader	19	2.00	1.247	0.286

Independent Samples Test – Individual Factors										
Factors Internal to the Evaluation Field										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Evaluator too expensive to include in budget	Equal variances assumed	1.607	0.207	-1.664	138	0.098	-0.509	0.306	-1.114	0.096
Program leadership skeptical of the benefits of evaluation	Equal variances assumed	0.090	0.764	-0.481	143	0.631	-0.147	0.306	-0.753	0.458
Inadequate skillset of evaluators to provide benefits during the program planning and design phase	Equal variances assumed	0.011	0.915	-3.193	132	0.002	-0.911	0.285	-1.476	-0.347
Program leadership had negative previous experience with evaluators	Equal variances assumed	0.054	0.817	-0.316	114	0.753	-0.089	0.281	-0.646	0.468

Evaluator may discover negative program results	Equal variances assumed	0.440	0.508	1.281	138	0.202	0.405	0.316	-0.220	1.031
Factors External to the Evaluation Field										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Organization funding the program did not require evaluation activities	Equal variances not assumed	4.633	0.033	-5.906	50.895	0.000	-1.548	0.262	-2.075	-1.022
Program leadership had no knowledge of or training on the benefits of evaluation in the planning and design phase	Equal variances assumed	0.000	0.995	0.834	146	0.405	0.262	0.314	-0.358	0.882
Program leadership had belief/understanding that evaluation only needs to occur at the end of a program	Equal variances assumed	0.167	0.684	2.414	150	0.017	0.727	0.301	0.132	1.322
Program leadership had program staff (who are not evaluators) plan and/or manage evaluation activities	Equal variances assumed	0.418	0.519	-3.354	146	0.001	-1.067	0.318	-1.695	-0.438
Previous programs had been successful without an evaluator, so there was no need to change	Equal variances assumed	1.252	0.265	-4.222	137	0.000	-1.180	0.280	-1.733	-0.628
Program's hiring process too difficult/lengthy to consider hiring an evaluator	Equal variances assumed	2.140	0.146	-1.415	126	0.159	-0.442	0.313	-1.061	0.176
Program leadership followed a supervisor's (or someone else higher up)	Equal variances assumed	0.372	0.543	-3.870	120	0.000	-1.247	0.322	-1.884	-0.609

decision to not hire an evaluator										
Adding/Managing an evaluator would have added to an already heavy workload	Equal variances assumed	0.006	0.940	-0.987	147	0.325	-0.284	0.287	-0.851	0.284
Evaluation findings would not be used in program decision-making, marketing, and/or funding decisions	Equal variances assumed	0.682	0.410	-2.400	140	0.018	-0.728	0.303	-1.328	-0.128
Inadequate skillset of program leadership to interpret and/or apply evaluation results/findings	Equal variances assumed	2.839	0.094	-1.115	142	0.267	-0.355	0.318	-0.984	0.274
PRA processes mandated by OMB too burdensome for designing/implementing evaluations	Equal variances assumed	3.852	0.053	-2.048	98	0.043	-0.543	0.265	-1.070	-0.017

APPENDIX G

INDEPENDENT SAMPLE T-TEST RESULTS – PROGRAM EVALUATOR TYPE

Group Statistics – Program Evaluator by Program Type					
Factor	Program Type	N	Mean	Std. Deviation	Std. Error Mean
Factors Internal to the Evaluation Field					
Evaluator too expensive to include in budget	Federal	79	2.38	1.470	0.165
	Non-federal	31	2.35	1.450	0.260
Program leadership skeptical of the benefits of evaluation	Federal	82	2.62	1.411	0.156
	Non-federal	35	2.63	1.477	0.250
Inadequate skillset of evaluators to provide benefits during the program planning and design phase	Federal	81	1.64	1.186	0.132
	Non-federal	31	1.90	1.248	0.224
Program leadership had negative previous experience with evaluators	Federal	62	1.89	1.203	0.153
	Non-federal	28	1.86	1.297	0.245
Evaluator may discover negative program results	Federal	77	2.25	1.416	0.161
	Non-federal	36	2.64	1.457	0.243
Factors External to the Evaluation Field					
Organization funding the program did not require evaluation activities	Federal	80	2.30	1.488	0.166
	Non-federal	31	2.45	1.362	0.245
Program leadership had no knowledge of or training on the benefits of evaluation in the planning and design phase	Federal	84	3.33	1.459	0.159
	Non-federal	35	3.29	1.545	0.261
Program leadership had belief/understanding that evaluation only needs to occur at the end of a program	Federal	86	3.45	1.508	0.163
	Non-federal	35	3.71	1.363	0.230
Program leadership had program staff (who are not evaluators) plan and/or manage evaluation activities	Federal	82	2.78	1.532	0.169
	Non-federal	35	2.51	1.463	0.247
Previous programs had been successful without an evaluator, so there was no need to change	Federal	74	2.49	1.377	0.160
	Non-federal	33	2.52	1.417	0.247
Program's hiring process too difficult/lengthy to consider hiring an evaluator	Federal	72	1.97	1.278	0.151
	Non-federal	30	2.37	1.542	0.282
Program leadership followed a supervisor's (or someone else higher up) decision to not hire an evaluator	Federal	66	2.20	1.438	0.177
	Non-federal	30	1.97	1.377	0.251
Adding/Managing an evaluator would have added to an already heavy workload	Federal	82	2.46	1.390	0.153
	Non-federal	35	2.57	1.399	0.237
Evaluation findings would not be used in program decision-making, marketing, and/or funding decisions	Federal	79	2.29	1.425	0.160
	Non-federal	34	2.35	1.368	0.235

Inadequate skillset of program leadership to interpret and/or apply evaluation results/findings	Federal	82	2.80	1.519	0.168
	Non-federal	32	2.78	1.680	0.297
PRA processes mandated by OMB too burdensome for designing/implementing evaluations	Federal	58	1.55	1.062	0.140
	Non-federal	21	1.14	0.655	0.143

Independent Samples Test – Program Evaluator by Program Type										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Diff.	Std. Error Diff.	95% Confidence Interval of the Difference	
									Lower	Upper
Factors Internal to the Evaluation Field										
Evaluator too expensive to include in budget	Equal variances assumed	0.242	0.624	0.080	108	0.936	0.025	0.310	-0.590	0.640
Program leadership skeptical of the benefits of evaluation	Equal variances assumed	0.201	0.655	-0.023	115	0.982	-0.007	0.289	-0.579	0.566
Inadequate skillset of evaluators to provide benefits during the program planning and design phase	Equal variances assumed	0.207	0.650	-1.028	110	0.306	-0.261	0.254	-0.765	0.242
Program leadership had negative previous experience with evaluators	Equal variances assumed	0.064	0.801	0.107	88	0.915	0.030	0.281	-0.528	0.588
Evaluator may discover negative program results	Equal variances assumed	0.204	0.653	-1.359	111	0.177	-0.392	0.288	-0.964	0.180
Factors External to the Evaluation Field										
Organization funding the program did not require evaluation activities	Equal variances assumed	0.871	0.353	-0.493	109	0.623	-0.152	0.308	-0.761	0.458
Program leadership had no knowledge of or training on the benefits of evaluation	Equal variances assumed	0.242	0.624	0.159	117	0.874	0.048	0.299	-0.544	0.639

in the planning and design phase										
Program leadership had belief/understanding that evaluation only needs to occur at the end of a program	Equal variances assumed	1.048	0.308	-0.886	119	0.377	-0.261	0.294	-0.844	0.322
Program leadership had program staff (who are not evaluators) plan and/or manage evaluation activities	Equal variances assumed	0.209	0.648	0.872	115	0.385	0.266	0.305	-0.338	0.871
Previous programs had been successful without an evaluator, so there was no need to change	Equal variances assumed	0.423	0.517	-0.099	105	0.922	-0.029	0.291	-0.605	0.548
Program's hiring process too difficult/lengthy to consider hiring an evaluator	Equal variances assumed	3.669	0.058	-1.335	100	0.185	-0.394	0.295	-0.981	0.192
Program leadership followed a supervisor's (or someone else higher up) decision to not hire an evaluator	Equal variances assumed	1.023	0.314	0.737	94	0.463	0.230	0.313	-0.390	0.851
Adding/Managing an evaluator would have added to an already heavy workload	Equal variances assumed	0.007	0.932	-0.384	115	0.702	-0.108	0.281	-0.665	0.449
Evaluation findings would not be used in program decision-making, marketing, and/or funding decisions	Equal variances assumed	0.094	0.760	-0.214	111	0.831	-0.062	0.289	-0.634	0.510

Inadequate skillset of program leadership to interpret and/or apply evaluation results/findings	Equal variances assumed	1.801	0.182	0.072	112	0.942	0.024	0.326	-0.623	0.670
PRA processes mandated by OMB too burdensome for designing/impl ementing evaluations	Equal variances not assumed	10.873	0.001	2.048	57.870	0.045	0.409	0.200	0.009	0.809

APPENDIX H

INDEPENDENT SAMPLE T-TEST RESULTS – PROGRAM LEADER TYPE

Group Statistics - Program Leader by Program Type

Factors Internal to the Evaluation Field					
Factor	Program Type	N	Mean	Std. Deviation	Std. Error Mean
Evaluator too expensive to include in budget	Federal	19	2.63	1.383	0.317
	Non-federal	6	3.17	1.169	0.477
Program leadership skeptical of the benefits of evaluation	Federal	18	2.44	1.381	0.326
	Non-federal	6	3.17	0.983	0.401
Inadequate skillset of evaluators to provide benefits during the program planning and design phase	Federal	13	2.85	1.405	0.390
	Non-federal	7	2.14	0.690	0.261
Program leadership had negative previous experience with evaluators	Federal	16	1.88	1.088	0.272
	Non-federal	6	2.17	1.835	0.749
Evaluator may discover negative program results	Federal	19	1.84	1.385	0.318
	Non-federal	4	1.25	0.500	0.250
Factors External to the Evaluation Field					
Organization funding the program did not require evaluation activities	Federal	19	3.84	1.344	0.308
	Non-federal	7	4.14	0.690	0.261
Program leadership had no knowledge of or training on the benefits of evaluation in the planning and design phase	Federal	21	2.95	1.596	0.348
	Non-federal	5	3.20	1.643	0.735
Program leadership had belief/understanding that evaluation only needs to occur at the end of a program	Federal	20	2.60	1.353	0.303
	Non-federal	6	3.17	1.835	0.749
Program leadership had program staff (who are not evaluators) plan and/or manage evaluation activities	Federal	19	3.74	1.522	0.349
	Non-federal	6	4.17	1.602	0.654
Previous programs had been successful without an evaluator, so there was no need to change	Federal	20	3.50	1.318	0.295
	Non-federal	7	4.29	0.756	0.286
Program's hiring process too difficult/lengthy to consider hiring an evaluator	Federal	18	2.33	1.572	0.370
	Non-federal	5	3.00	1.225	0.548
Program leadership followed a supervisor's (or someone else higher up) decision to not hire an evaluator	Federal	17	3.53	1.586	0.385
	Non-federal	5	2.40	1.342	0.600
Adding/Managing an evaluator would have added to an already heavy workload	Federal	19	2.63	1.342	0.308
	Non-federal	7	3.14	1.676	0.634

Evaluation findings would not be used in program decision-making, marketing, and/or funding decisions	Federal	19	2.74	1.408	0.323
	Non-federal	5	3.80	1.095	0.490
Inadequate skillset of program leadership to interpret and/or apply evaluation results/findings	Federal	19	3.05	1.580	0.363
	Non-federal	7	3.14	0.900	0.340
PRA processes mandated by OMB too burdensome for designing/implementing evaluations	Federal	12	2.00	1.279	0.369
	Non-federal	6	1.83	1.329	0.543

Independent Samples Test - Program Leader by Program Type										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Diff.	Std. Error Diff.	95% Confidence Interval of the Difference	
									Lower	Upper
Factors Internal to the Evaluation Field										
Evaluator too expensive to include in budget	Equal variances assumed	1.976	0.173	-0.853	23	0.402	-0.535	0.627	-1.832	0.762
Program leadership skeptical of the benefits of evaluation	Equal variances assumed	2.406	0.135	-1.177	22	0.252	-0.722	0.614	-1.995	0.550
Inadequate skillset of evaluators to provide benefits during the program planning and design phase	Equal variances assumed	3.615	0.073	1.235	18	0.233	0.703	0.569	-0.493	1.899
Program leadership had negative previous experience with evaluators	Equal variances not assumed	5.172	0.034	-0.366	6	0.726	-0.292	0.797	-2.215	1.631
Evaluator may discover negative program results	Equal variances assumed	2.491	0.129	0.830	21	0.416	0.592	0.713	-0.891	2.075
Factors External to the Evaluation Field										
Organization funding the program did not require	Equal variances not assumed	6.081	0.021	-0.745	21	0.465	-0.301	0.404	-1.141	0.539

evaluation activities										
Program leadership had no knowledge of or training on the benefits of evaluation in the planning and design phase	Equal variances assumed	0.022	0.882	-0.310	24	0.759	-0.248	0.798	-1.895	1.400
Program leadership had belief/understanding that evaluation only needs to occur at the end of a program	Equal variances assumed	1.182	0.288	-0.830	24	0.415	-0.567	0.683	-1.976	0.842
Program leadership had program staff (who are not evaluators) plan and/or manage evaluation activities	Equal variances assumed	0.187	0.670	-0.596	23	0.557	-0.430	0.721	-1.921	1.062
Previous programs had been successful without an evaluator, so there was no need to change	Equal variances assumed	3.211	0.085	-1.482	25	0.151	-0.786	0.530	-1.878	0.306
Program's hiring process too difficult/lengthy to consider hiring an evaluator	Equal variances assumed	1.826	0.191	-0.872	21	0.393	-0.667	0.764	-2.256	0.923
Program leadership followed a supervisor's (or someone else higher up) decision to not hire an evaluator	Equal variances assumed	0.573	0.458	1.441	20	0.165	1.129	0.783	-0.505	2.764
Adding/Managing an evaluator would have added to an	Equal variances assumed	0.215	0.647	-0.807	24	0.428	-0.511	0.634	-1.819	0.796

already heavy workload										
Evaluation findings would not be used in program decision-making, marketing, and/or funding decisions	Equal variances assumed	1.229	0.279	-1.559	22	0.133	-1.063	0.682	-2.477	0.351
Inadequate skillset of program leadership to interpret and/or apply evaluation results/findings	Equal variances assumed	3.354	0.079	-0.142	24	0.889	-0.090	0.637	-1.405	1.224
Paperwork Reduction Act (PRA) processes mandated by Office of Management and Budget (OMB) too burdensome for designing/implementing evaluations	Equal variances assumed	0.106	0.749	0.257	16.000	0.800	0.167	0.648	-1.206	1.539

APPENDIX I
IRB APPROVAL LETTER

EXEMPTION GRANTED

[Jesse Lecy](#)
[WATTS: Interdisciplinary Programs](#)

-
jdlscy@asu.edu

Dear [Jesse Lecy](#):

On 9/1/2021 the ASU IRB reviewed the following protocol:

Type of Review:	Initial Study
Title:	Evaluator Exclusion in the Planning and Design Phase of Federally-funded Program
Investigator:	Jesse Lecy
IRB ID:	STUDY00014354
Funding:	None
Grant Title:	None
Grant ID:	None
Documents Reviewed:	<ul style="list-style-type: none"> • Evaluator Version - Dissertation Research - Evaluator Exclusion.pdf, Category: Measures (Survey questions/Interview questions /interview guides/focus group questions); • IRB Application - Dissertation Research.docx, Category: IRB Protocol; • IRB Information Letter - Dissertation Research.pdf, Category: Consent Form; • IRB Survey Emails - Dissertation Research.pdf, Category: Recruitment Materials; • Program Leader Version - Dissertation Research - Evaluator Exclusion.pdf, Category: Measures (Survey questions/Interview questions /interview guides/focus group questions);

The IRB determined that the protocol is considered exempt pursuant to Federal Regulations 45CFR46 (2) Tests, surveys, interviews, or observation on 9/1/2021.

In conducting this protocol you are required to follow the requirements listed in the INVESTIGATOR MANUAL (HRP-103).

If any changes are made to the study, the IRB must be notified at research.integrity@asu.edu to determine if additional reviews/approvals are required. Changes may include but not limited to revisions to data collection, survey and/or interview questions, and vulnerable populations, etc.

REMINDER - All in-person interactions with human subjects require the completion of the ASU Daily Health Check by the ASU members prior to the interaction and the use of face coverings by researchers, research teams and research participants during the interaction. These requirements will minimize risk, protect health and support a safe research environment. These requirements apply both on- and off-campus.

The above change is effective as of July 29th 2021 until further notice and replaces all previously published guidance. Thank you for your continued commitment to ensuring a healthy and productive ASU community.

Sincerely,

IRB Administrator

cc: Matthew Gallagher