



Western Michigan University
ScholarWorks at WMU

Dissertations

Graduate College

4-2016

A Metaevaluation of School Counseling Program Evaluations

Saeed Mohammed Almueed

Western Michigan University, sma5454@gmail.com

Follow this and additional works at: <https://scholarworks.wmich.edu/dissertations>



Part of the Educational Assessment, Evaluation, and Research Commons, and the Student Counseling and Personnel Services Commons

Recommended Citation

Almueed, Saeed Mohammed, "A Metaevaluation of School Counseling Program Evaluations" (2016).
Dissertations. 1430.

<https://scholarworks.wmich.edu/dissertations/1430>

This Dissertation-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact wmu-scholarworks@wmich.edu.



A META-EVALUATION OF SCHOOL COUNSELING PROGRAM
EVALUATIONS

by

Saeed Mohammed Almueed

A dissertation submitted to the Graduate College
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
Interdisciplinary Ph.D. in Evaluation
Western Michigan University
April 2016

Doctoral Committee:

Chris Coryn, Ph.D., Chair
Daniela Schroeter, Ph.D.
Beverly Vandiver, Ph.D.

A META-EVALUATION OF SCHOOL COUNSELING PROGRAM EVALUATIONS

Saeed Mohammed Almueed, Ph.D.

Western Michigan University, 2016

School counseling programs have been recognized as an educational component that ultimately benefits student achievement at most developmental stages.

Comprehensive school counseling programs are executed in alliance with a specific school counseling model, such as the Comprehensive Developmental Guidance model, or more recently, *The American School Counselor Association National Model* (ASCA, 2003, 2005, 2012). Although the school counseling literature reports positive outcomes associated with comprehensive school counseling programs, there are still challenges in their implementation and evaluation. Program evaluation as stressed in *The ASCA National Model* is to confirm the effectiveness and value of these programs and interventions. Yet, the school counseling literature has not clearly discussed the wide range of existing tools used in conducting sound program evaluations, nor has it thoroughly explored the quality of existing program evaluations. Consequently, the purpose of this sequential two-phase study was to conduct a metaevaluation of school counseling program evaluations.

A sample of school counseling evaluations was assessed regarding their conformity to the Joint Committee on Standards for Educational Evaluation (JCSEE) accuracy standards (Yarbrough et al., 2011), utilizing the Program Evaluation

Metaevaluation Checklist (Stufflebeam & Coryn, 2014). The authors of the evaluations engaged in semi-structured interviews designed to discover their perceptions of the capability of program evaluation to accurately determine the effectiveness and efficiency of school counseling programs and interventions. Findings of the study showed that the overall metaevaluation ratings conformed less to JCSEE standards. However, this low performance is further explained from the interview participants' perspectives. Interviews themes corresponded to strengths and weaknesses, methodology orientation, and key factors affecting the practice of evaluation in the field. The findings may assist the school counseling community and policymakers in gaining insight into the issues surrounding evaluation practice in the field and the use of metaevaluation as a tool to assess quality. To further these aims, it is also recommended that school counselors build more bridges and increase communication with the professional evaluation community.

© 2016 Saeed Mohammed Almueed

ACKNOWLEDGMENTS

First, I would like to begin my acknowledgments by thanking Allah (God) for his blessings and guidance, which were the lights that strengthened me throughout all of the years pursuing my graduate studies. I also would like to dedicate this work to my wonderful parents, brothers, and sisters for their ongoing help and encouragement.

Second, I would like to express my sincere gratitude to my advisor and committee chair, Dr. Chris Coryn, for his continuous support of my doctoral studies and related academic work. I want to thank Dr. Coryn especially for his professionalism, patience, motivation, and immense knowledge. I could not have imagined having a better advisor and mentor. Besides Dr. Coryn, I would like to thank the rest of my dissertation committee, Dr. Daniela Schroeter and Dr. Beverly Vandiver, for their insightful comments, kindness, and continuing encouragement, and also for their valuable feedback, which assisted me in widening my research from various perspectives.

Third, my sincere thanks also goes to Dr. Pedro Mateu, who helped me with the coding process. His knowledge and assistance were a valuable piece to my study. I would also like to extend my appreciation to the great person, Mary Ramlow, for her kindness and professional support.

Acknowledgments—Continued

Finally, huge thanks and gratitude go to my amazing wife and daughters, Rose and Rand, for their patience and love. They are the ones who inspire and empower me during the hard times and moments in this enjoyable journey.

Saeed Mohammed Almueed

TABLE OF CONTENTS

ACKNOWLEDGMENTS	ii
LIST OF TABLES	viii
CHAPTER	
I. INTRODUCTION	1
Background of the Study	2
Comprehensive School Counseling Programs.....	3
Program Evaluation in the School Counseling Context	6
Statement of the Problem.....	9
Purpose of the Study	9
Research Questions.....	10
Importance of the Study.....	11
Definition of Terms	12
II. REVIEW OF RELATED LITERATURE	15
Overview of Professional Evaluation	15
Evaluation Practice	17
Evaluation Standards	23
Metaevaluation: Concept and Practice	26
Perceptions of Program Evaluation in the Field of School Counseling	29
Accountability Practices of School Counseling Programs	34
Preparing School Counselors in Evaluation	38

Table of Contents—Continued

CHAPTER		
	Chapter II Summary.....	40
III.	METHODOLOGY	42
	Purpose of the Study and Research Questions.....	42
	Research Design	43
	Sample	44
	Metaevaluation	44
	Semi-structured Interviews.....	48
	Instrumentation	49
	Data Collection Procedures	51
	Data Analysis.....	53
	Chapter III Summary	56
IV.	RESULTS	57
	Results Related to Research Question 1	58
	Results of Program Evaluation Metaevaluation Checklist: Overall	58
	Results of Program Evaluation Metaevaluation Checklist: Sub-standards.....	60
	Findings Related to Research Question 2	72
	Identified Themes and Sub-themes	73
	Strengths	74

Table of Contents—Continued

CHAPTER		
	Increased Awareness of Program Evaluation	74
	Financial Assistance	79
	Weaknesses	80
	Competence	81
	Leadership Support	82
	Resource Allocation	83
	Methodology Orientation	84
	Common Methodologies	84
	Methodology Barriers	85
	Key Factors Influencing Program Evaluation Practice	86
	Pre-service Preparation	87
	In-service Preparation	88
	Discussion	89
V.	CONCLUSION	93
	Summary of Findings	94
	Accuracy Standard Ratings	94
	Strengths	97
	Weaknesses	98
	Methodology	98
	Factors Affecting Evaluation Practice	99

Table of Contents—Continued

CHAPTER	
Implications	100
Limitations	101
Contributions	104
Future Research	104
REFERENCES	106
APPENDICES	117
A. Example of Program Evaluation Metaevaluation Checklist- Accuracy Standard	117
B. Interview Protocol.....	121
C. Consent Letter	124
D. HSIRB Approval.....	128
E. Recruitment Email	130
F. Overall Metaevaluation Ratings For All Evaluations	132

LIST OF TABLES

1. Inclusionary Criteria for Evaluations	46
2. Evaluations Based on School Levels.....	47
3. Evaluations Based on Domains of School Counseling Programs.....	47
4. Example of Program Evaluation Metaevaluation Checklist-Accuracy Standard ...	50
5. Template of Data Analysis for Study’s Total Accuracy Strength.....	54
6. Template of Data Analysis for Study’s Total Accuracy Strength.....	59
7. Evaluation Ratings Based on the School Levels	59
8. Evaluation Ratings Based on the Domains of School Counseling Programs	60
9. Checkpoint Results for A1: Justified Conclusions and Decisions	61
10. Rating Scale Pertinent to A1: Justified Conclusions and Decisions	61
11. Checkpoint Results for A2: Valid Information	62
12. Rating Scale Pertinent to A2: Valid Information	63
13. Checkpoint Results for A3: Reliable Information.....	64
14. Rating Scale Pertinent to A3: Reliable Information.....	64
15. Checkpoint Results for A4: Explicit Program and Context Descriptions	65
16. Rating Scale Pertinent to A4: Explicit Program and Context Descriptions	66
17. Checkpoint Results for A5: Information Management	66
18. Rating Scale Pertinent to A5: Information Management	67
19. Checkpoint Results for A6: Sound Designs and Analyses.....	68
20. Rating Scale Pertinent to A6: Sound Designs and Analyses.....	69

List of Tables—Continued

21. Checkpoint Results for A7: Explicit Evaluation Reasoning	69
22. Rating Scale Pertinent to A7: Explicit Evaluation Reasoning	70
23. Checkpoint Results for A8: Communicating and Reporting.....	71
24. Rating Scale Pertinent to A8: Communicating and Reporting.....	72
25. Identified Themes and Sub-themes	73

CHAPTER I

INTRODUCTION

The United States has experienced several educational movements and related legislative changes over time. Ideally, these changes support educational professionals and systems in raising student achievement (House & Hayes, 2002). School counselors in particular have been an important part of instrumental changes designed to increase student achievement, not only academically, but also vocationally, personally, and socially (Sink, 2005). For school counselors, this responsibility is done primarily through the provision of comprehensive school counseling programs. Comprehensive school counseling programs (CSCPs) are designed and implemented based on a specific model, such as the Comprehensive Developmental Guidance model, or more recently, *The American School Counselor Association National Model* (ASCA, 2012; Palmer & Erford, 2012). Such programs are supportive to general educational goals, which place CSCPs among the other essential educational systems (Kozlowski & Huss, 2013). By providing these programs and interventions, school counselors align themselves with other state and national structures focused on student achievement across all grade levels (Martin, Carey, & DeCoster, 2009).

Given the importance of school counseling programs to comprehensive student services, professional school counselors over the past two decades have focused on developing models that utilize best practices as it relates to the design and implementation of these programs in schools (Martin & Carey, 2012; Sink, Akos, Turnbull, & Mvududu 2008). Research shows that best practice in the continuous development of school counseling models includes evaluating their effectiveness and

efficiency (ASCA, 2003). Yet, while there have been increasing demands for school counselors to demonstrate the effectiveness of their implemented programs, conducting program evaluation in school counseling remains limited (ASCA, 2012; Martin et al., 2009; Trevisan, 2002). The aim of this study was to assess the quality of current evaluation practice in the field of school counseling by conducting a metaevaluation of program evaluations performed thus far and then discovering the perceptions of evaluators regarding the viability of program evaluation performance in the field.

Background of the Study

As mentioned above, school counseling is guided by the implementation of comprehensive school counseling programs (ASCA, 2012). Many state departments of education and school counseling associations support the use of such programs as part of larger education reform movements (Martin & Carey, 2012; Martin et al., 2009). However, a national study found that the design, delivery, and sustainability of school counseling program implementation are widely inconsistent (Martin et al., 2009). Although the school counseling literature generally reports positive outcomes associated with comprehensive school counseling programs, there are still many challenges in their implementation and evaluation (Studer, Diambra, Breckner, & Heidel, 2011). In a study of professional school counselors in southwestern states, Studer and his colleagues (2011) found that a majority of respondents surveyed across school levels were low in terms of the regularity at which they collected and analyzed data. Similarly, in a national study surveyed by Martin et al. (2009), program evaluation in school counseling was infrequently conducted across the majority of states.

The sections below review the history of comprehensive school counseling programs and evaluation. Following this review, the purpose of the study and the research questions are outlined. The chapter concludes with the importance of the study, followed by definition of terms.

Comprehensive School Counseling Programs

The focus of school counseling early at the start of the profession was on vocational guidance, which was designed to prepare students for jobs associated with the economic and social developments of the time (Gysbers, 2001). Changes in educational trends and needs influenced a shift in the profession from a focus on vocation to a more practical vision of guidance that resulted in the manifestation of a student services model (Gysbers, 2001; McGamon, Carey, & Dimmitt, 2005). The student services model emerged to be more inclusive of other essential student needs beyond vocation (e.g., students with most needs of counseling services). With this model, the structure of guidance services and the role of counselors became more organized around a holistic set of aids for students (Gysbers, 2001); however, the role of the school counselors was not clearly understood by all stakeholders. It seemed to some that school counselors were basically performing tasks unrelated to their core responsibilities, such as clerical duties (Gysbers, 2001; Whiston & Sexton, 1998). Moreover, under this early model, school counseling programs targeted individuals who were in the most need rather than providing comprehensive services to all students in the school setting (McGamon et al., 2005). For some, this confusion caused school counseling programs to appear as irrelevant to the success of all students (Gysbers, 2001).

Because of the issues described above, there was a need in the profession to redefine the tasks of school counselors. This redefinition included structuring an organized model of school counseling. Furthermore, the school counseling profession sought to establish a structural guidance model that would serve students not only with service-based guidance, but also in consideration of their age and level of development. Accordingly, school counseling professionals became concerned with addressing students' academic, career, and personal needs (Gysbers, 2001; Zagelbaum, Kruczek, Alexander, & Crethar 2014), resulting in the Comprehensive Developmental Guidance (CDG) program model (Gysbers, 2001). Compared to the traditional student services model, the CDG model was more focused on student competence, making student competence the central element (McGamon et al., 2005). As the CDG model was subsequently developed, the counselor's tasks became more clearly identified, and facets of the counseling process were more delineated (Gysbers, 2001; McGamon et al., 2005).

The CDG model defined the direction and purposes of school counseling services, and positioned the model among other essential programs in the education system (McGamon et al., 2005). More specifically, CDG programs intentionally sought to develop student knowledge, skills, and awareness in academic, career, and personal/social domains in order to help fulfill the overall mission of the school and ultimately its educational goals (ASCA, 2012; McGamon et al., 2005; Sink et al., 2008). The CDG model was also organized around specific areas to enhance learning outcomes and promote positive experiences in schools. Attention was given to identify these learning outcomes and goals as core elements of the "guidance curriculum" (McGamon et al., 2005).

In 2003, *The ASCA National Model: A Framework for School Counseling Programs* (ASCA, 2003) was created as a result of continuous advances in education philosophy, the success of CDG model outcomes, and the rise in the importance of the accountability in the educational system (ASCA, 2003; Dahir et al., 2009; McGamon et al., 2005); however, the subsequent developments of CDG model in particular were the base for *The ASCA National Model's* emergence (ASCA, 2003, 2005; McGamon et al., 2005). While *The ASCA National Model* (ASCA, 2003) was established to support all goals and objectives identified in the CDG model and the academic mission of the school, attention was dedicated specifically to accountability and data-based practices (Sink et al., 2008). Accordingly, meeting accountability demands and making effective use of school and student information are essential facets of effective school counseling programs (McGamon et al., 2005).

Full implementation of comprehensive school counseling programs (CSCPs) based on *The ASCA National Model* (ASCA, 2003, 2005, 2012) has been shown to have a positive impact on student performance (Burkard, Gillen, Martines, & Skytte, 2012; Martin et al., 2009). Lapan (2012) assessed the findings of six studies about the implementation of CSCPs and revealed that students would more likely to benefit from schools that “fully implement a comprehensive school counseling program” (p. 84). However, there are limited studies that inspect how such programs are consistently executed (Burkard et al., 2012; Martin et al., 2009; Studer, Diambra, Breckner, & Heidel, 2011). Zagelbaum et al. (2014) indicate that there is variation in the implementation of CSCPs across states. In fact, Martin et al. (2009) conducted a national study examining the application of the school counseling models at the state level and found only 17 states

were progressing toward grounding their practices on the utilization of structured school counseling programs and models. While some states choose to benefit from existing models and counseling programs, other states establish and write their own (Zagelbaum et al., 2014).

Besides variation in the way CSCPs are implemented across states, there are also significant variations in how they are implemented across grades and by individual school counselors (Zagelbaum et al., 2014). Dahir, Burnham, and Stone (2009) found that school counselors implement counseling programs addressing the areas of social and personal development, but they were less attentive to the academic and career development of the students. Dahir et al. (2009) also found that CSCPs were more frequently observed at the middle school level than the other two school levels.

Recently, school counseling programs have been recognized as an educational component that, if applied effectively, ultimately benefits student achievement at most developmental stages. Yet, as mentioned, program audit or evaluation—one of the essential elements stressed in *The ASCA National Model* (ASCA, 2003, 2005, 2012)—must be consistently performed in order to confirm the effectiveness and value of these programs and interventions (Palmer & Erford, 2012). Evaluation is necessary not only for the academic success of students in schools, but also for the advancement and creditability of the school counseling profession in general (ASCA, 2012; Burkard et al., 2012).

Program Evaluation in the School Counseling Context

Program evaluation has been perceptible in the school counseling profession ever since school counseling and guidance models emerged. The benefits of program

evaluation, however, were not adequately acknowledged in the field until the publication of *The ASCA National Model* (ASCA, 2003). *The ASCA National Model* (ASCA, 2003, 2005, 2012) requires under its accountability section the use of program evaluation, which is considered an essential aspect of the model's implementation structure. In the accountability section of the model, program evaluation is to be used to document the effectiveness of school counseling programs on student achievement and learning outcomes. Beyond documentation, program evaluation entails reporting results to key stakeholders for funding purposes, program decisions, and further student development (Whiston & Aricak, 2008).

With the rapid increase of students' needs that are served by CSCPs, program evaluation has become an important endeavor that should gain considerable attention by the professionals and decision-makers in the field of school counseling for many reasons (Dimmit, 2009; Martin & Carey, 2012). First, education professionals and decision-makers are not likely to continue to recognize the value and support of these programs if their outcomes are not shown to produce justifiable results. Second, without quality formative evaluations, it is difficult to make informed decisions about the implementation and evaluation process while seeking to improve program interventions. Formative evaluation helps school counselors or evaluators assess the programs or interventions during their implementation phase in order to track problems that exist and offer immediate responses (Stufflebeam & Coryn, 2014). Third, without quality summative evaluations, it is difficult to make any decisions concerning the effectiveness and efficiency of the program. Summative evaluations provide school counselors or evaluators with the information needed to convince top stakeholders of the program's

merit and worth (Stufflebeam & Coryn, 2014). Similarly, financial recourses are the core elements that allow these counseling programs and interventions to perform well and produce expected outcomes. Although there are many funding channels, federal, state, and local profit and non-profit agencies (Vernon & Rainey, 2009), program evaluation is an essential process to validate the effectiveness and efficiency of such programs and interventions to the top stakeholders for spent recourses.

A significant amount of literature in the field of school counseling urges professionals to evaluate school counseling programs for their effects and outcomes (ASCA, 2012; Dimmitt, 2009; 2010; Martin et al., 2009; Trevisan, 2002). Despite *The ASCA National Model's* (ASCA, 2003, 2005, 2012) assertion that program evaluation should be an annual process conducted to demonstrate the impact of programs on student outcomes, Martin and Carey (2012) stressed that program evaluation is not sufficiently apparent among many state CSCPs. This state is true despite the fact that the school counseling field has been moving toward adopting new concepts, knowledge, and skills necessary to perform program evaluation (Trevisan, 2002).

A review of the literature concerning school counseling program evaluation revealed a limited number of studies that inspected program evaluation procedures, current methodologies, and quality mechanisms within the school counseling context. Scholars in the field of school counseling suggest there are a number of challenges that may influence the practice of program evaluation. These challenges include, but are not limited, a lack of training in evaluation (Trevisan, 2002), a lack of knowledge and skills of evaluation procedures (Dimmitt, 2010), mistrust of the evaluation process, and fear of evaluation outcomes (Astramovich, Coker, & Hoskins, 2005). Dimmitt (2010) stated that

without evaluation efforts, the impact of school counseling programs and interventions is not clearly justified. In reviewing the general mechanisms to evaluate school counseling programs, (Dimmitt, 2010) stressed, “one or more parts of what school counselors were doing did not seem to be having the intended impact” (p. 55).

Statement of the Problem

The school counseling profession has dedicated many efforts to sound school counseling programs that meet the desired goals of education and stated school missions. Program evaluation is among the best means for determining if such efforts have been fulfilled. Scholars in the field of school counseling reveal that program evaluation as a method to assess program effectiveness and quality is less commonly performed (Astramovich et al., 2005; Dimmitt, 2010; Studer et al., 2011; Trevisan, 2002). Furthermore, the school counseling literature has not clearly discussed the wide range of contemporary evaluation methods and existing tools used in conducting sound program evaluations. Consequently, examining the evaluations of school counseling program that have been conducted thus far is an essential endeavor in order to gain insight into: (a) how program evaluation is currently practiced in the field, and (b) if these evaluations adhere to professional evaluation standards.

Purpose of the Study

As stated, the school counseling profession has a dearth of literature exploring program evaluation practice as part of demonstrating the effectiveness and efficiency of CSCPs. Furthermore, the quality of existing program evaluations within the field of school counseling has not been made evident within the literature. Thus, the purpose of this sequential two-phase study was to conduct a metaevaluation of school counseling

program evaluations. The first phase of the study involved assessing a sample of school counseling evaluations in terms of their conformity to the Joint Committee on Standards for Educational Evaluation (JCSEE) accuracy standards (Yarbrough, Shulha, Hopson, & Caruthers, 2011). This process was done through the use of the Program Evaluation Metaevaluation Checklist (Stufflebeam & Coryn, 2014). The second phase engaged the authors of the evaluations in semi-structured interviews to discover their perceptions about the capability of program evaluation to accurately determine the effectiveness and efficiency of school counseling programs and interventions. The interview process attempted to uncover the evaluators' viewpoints in regard to program evaluation strengths and weaknesses, methodology selection, and other factors influencing program evaluation practice in the context of school counseling. Additionally, as Henry and Mark (2003) have urged the evaluation community to augment the research on metaevaluation literature, this study was also intended to add to the growing number of metaevaluations currently being conducted in the evaluation discipline itself.

Research Questions

This study attempted to answer the following research questions:

1. How well do school counseling program evaluations conform to the JCSEE accuracy standards (Yarbrough et al., 2011)?
2. What are the perceptions and experiences of the participating evaluators regarding the status and accuracy of school counseling program evaluation?
 - a. What do evaluators perceive as the strengths and weaknesses of school counseling program evaluation?

- b. How do evaluators align their evaluations to existing evaluation methods and practice?
- c. What are the major factors influencing program evaluation practice in school counseling context?

Importance of the Study

As previously stated, the purpose of this study was to explore and assess the quality and current practice of program evaluation within the field of school counseling. The researcher's interest to conduct such a study stemmed from the scarcity of evaluation efforts noted in the literature. Moreover, the field of school counseling as an endorsed component of the U.S. educational system appears to be lacking evaluation mechanisms needed to demonstrate the effectiveness, efficiency, and accountability demands of its programs and interventions. Hence, assessing the quality of school counseling program evaluations is an important attempt to recognize existing trends in the literature related to evaluation practice. More importantly, the study's findings may assist the school counseling community and policymakers in gaining insight into the issues surrounding evaluation practice in the field, and planning for undertaking sound evaluation work.

The metaevaluation technique employed in the study provides evaluators, interested practitioners, and professional school counselors with an evaluation tool that can be used to assess their evaluation endeavors for both improvement and accountability purposes. Based on the literature search, the use of metaevaluation for quality assessment has not been discussed in the school counseling literature. Finally, this study may motivate the evaluation community and school counseling profession to further examine evaluation practices along with its existing instruments in this specific field. The

knowledge and practice of evaluation will be beneficial for the CSCPs' legitimacy and funding purposes.

Definition of Terms

Comprehensive School Counseling Programs: In accordance to *The ASCA National Model* (ASCA, 2012), or other established state models, school counselors are responsible for planning, designing, implementing, and evaluating comprehensive school counseling programs that meet the stated goals of the school's mission, and support student growth in academic, career, personal/social development (ASCA, 2012; Palmer & Erford, 2012). According to ASCA (2012), these programs are comprehensive in scope, preventive in design, and developmental in nature. School counselors, therefore, should be responsive to implement programs, interventions, and activities that advance students' key skills, knowledge, and attitudes in three main areas: academic achievement, personal/social, and career education (Sink et al., 2008).

Program Evaluation: Prior to proceeding with defining the term *program evaluation* used in this study, it is important to shed light on the general definition of evaluation. One of the earlier definitions provided by Scriven (1991) was that evaluation is "the process of determining the merit, worth, and value of things, and evaluations are the products of that process" (p. 139). However, later on, there were developments to his definition by prominent evaluation scholars who perceive the importance of incorporating more vital and practical values to the definition. Therefore, Stufflebeam and Coryn (2014) define evaluation as "the systematic assessment of an object's merit, worth, probity, feasibility, safety, significance, and/or equity" (pp. 11-12).

Stufflebeam and Coryn (2014) further define *program evaluation* as “...assessments of any coordinated set of activities... that have been discernibly planned to meet assessed needs and defined goals” (p. 110). Similarly, Yarbrough et al. (2011) defined the program evaluation as follows:

The systematic investigation of the quality of programs, projects, subprograms, subprojects, and/or any of their components or elements, together or singly for the purpose of decision making, judgments, conclusions, findings, new knowledge, organizational development, and capacity building in response to the need of identified stakeholders leading to improvement and/or accountability in the users’ programs and systems ultimately contributing to organizational or social value. (p. xxv)

In the school counseling context, Dimmitt (2009) defined the program evaluation as “the purposeful and systematic collection and analysis of data or information used for the purpose of documenting the effectiveness, impact, and outcomes of programs, establishing accountability, and identifying areas needing change and improvement” (p. 395). This definition aligns with the definition provided by Rossi, Lipsey, and Freeman (2003). They defined program evaluation as “the use of social research methods to systematically investigate the effectiveness of social intervention programs in ways that are adapted to their political and organizational environments and are designed to inform social action to improve social conditions” (p. 16).

Metaevaluation: The simplest description of the term metaevaluation is derived from Scriven’s (1991) definition, “evaluation of evaluations” (p. 228). Operationally, Stufflebeam and Coryn (2014) elaborated on the definition further, describing it as “the

process of delineating, obtaining, and applying descriptive and judgmental information about an evaluation's utility, feasibility, propriety, accuracy, and accountability for the purposes of guiding the evaluation and reporting its strengths and weaknesses" (p. 635). This definition is more comprehensive to encompass the process undertaken to evaluate the evaluation reports or studies against the program evaluation standards put forth by the JCSEE.

CHAPTER II

REVIEW OF RELATED LITERATURE

This chapter presents literature relevant to the use of program evaluation in school counseling, beginning with a brief overview of evaluation as a profession. This overview includes a discussion on the evaluation development related to evaluation practice, evaluation standards, and metaevaluation. The next section sheds light on the perception of program evaluation within the context of school counseling. It covers the importance of evaluation in school counseling, accountability practices, and school counselors' preparation and training in evaluation methods and principles.

Overview of Professional Evaluation

Evaluation is profession in which the aim is to serve various societal, professional, and individual needs. Evaluation as a recognized profession today has gone through numerous developments over time. These developments in history, theory, methods, and practice have increased the profession's reliability in an array of different fields. The benefits of evaluation in public or private sectors are now evident, due primarily to its vital mission to ensure the quality, effectiveness, and efficiency of services, programs, and products (Stufflebeam & Coryn, 2014). In addressing the essential role of evaluation in public areas and its practical functions in legislation, Chouinard (2013) described evaluation as a valid tool for "...decision making, monitoring, standard-setting, improvement, reporting, and controlling of program activities and expenditures..." (p. 268). Moreover, the importance of evaluation is not only perceived as an important practice in the United States, but its procedures and endorsements are globally recognized.

Early in its history, the evaluation field experienced a number of transitions, including a great deal of debate and scholarly discussions to determine best practices in the profession. In fact, evaluation has been recognized as its own field since the beginning of 19th century, with many professionals from different fields participating in its theoretical and practical development (Stufflebeam & Coryn, 2014). The education movement in 1930s was the initial event that laid groundwork for the concept of evaluation to develop and expand. In particular, educational evaluation was introduced and strengthened with Ralph Tyler's substantial contributions in research (Mathison, 2005; Stufflebeam & Coryn, 2014). In the 1960s, there was a shift in perceptions about the use of evaluation, its role, and feasibility. This shift was reflected by a number of legislative and economic decisions (e.g., the War on Poverty and the Title I policy of the Elementary and Secondary Education Act) that were reliant upon evaluation in their implementation (Hogan, 2007; Shadish, Cook, & Leviton, 1991; Stufflebeam & Coryn, 2014).

LaVelle and Donaldson (2010) noted that evaluation's formal inception in the late 1960s had considerable attention from a wide range of educational professionals, further raising the recognition of evaluation as a profession. This professionalism was seen through the establishment of organizations (e.g., the American Evaluation Association), evaluation based programs (e.g., the Interdisciplinary Ph.D. program in Evaluation at Western Michigan University), and scholarly journals (*Evaluation Review* and *New Directions for Evaluation*) (LaVelle & Donaldson, 2010; Shadish et al., 1991; Stufflebeam & Coryn, 2014). Finally, the quality of evaluation efforts and tools were also given considerable attention in the field. Quality criteria, standards, and guiding

mechanisms were also addressed and established through credential organizations such as the Joint Committee on Standards for Educational Evaluation (Yarbrough, Shulha, Hopson, & Caruthers, 2011) and the American Evaluation Association's Guiding Principles (Newman, Scheirer, Shadish, & Wye, 1995).

Evaluation Practice

Integration of theory into practice. Evaluation has been theoretically, methodologically, and practically evolving over the last few decades. Following the creation of federal programs and funds for evaluation early in 1970s, accountability gained considerable attention (Donaldson & Lipsey, 2006). Building the evaluation knowledge base, as well as developing and employing the social science methods, were also notable objectives at that time (Mathison, 2005; Stufflebeam & Coryn, 2014). However, after the 1970s, the focus of the evaluation community and organizations shifted broadly to develop other essential components of evaluation practice (Donaldson & Lipsey, 2006). Specifically, the evaluation community increased the efforts to create and develop a number of evaluation approaches and tools to guide and assist evaluators in their practice (Stufflebeam & Coryn, 2014). For instance, theory-driven evaluation (Chen, 1990), utilization-focused evaluation (Patton, 2008), participatory evaluation (Cousin & Earl, 1992), empowerment evaluation (Fetterman, 2001), and program theory (Rogers, 2008) models, among other models, have captured the greatest attention among professional evaluators (Cullen, Coryn, & Rugh, 2011).

Numerous endeavors have been made by members of the evaluation community to clarify a realistic understanding of evaluation practice. Key evaluation scholars have gathered such efforts through long and productive discussions, writings, and debates.

Although perceptions of evaluation theories and theories application are important to support evaluation practice (Christie, 2003), there have been variations of accepting the feasibility and utilization of theories in real world practice (Coryn, Noakes, Westine, & Schroter, 2011). To this end, Christie (2003) conducted a study to examine the extent to which practicing evaluators of California's Healthy Start program carried out the theories of eight theorists. The finding was that evaluators in the sample were less likely to adhere to an entire theory. Particularly, the evaluators selectively utilized parts of the chosen theories rather than the whole theory.

In addition to Christie's (2003) examination of theory in practice, Coryn et al. (2011) reviewed scholars' endeavors relevant to theory-driven evaluations in the last two decades. They found "...very little empirical evidence exists to buttress the numerous theoretical postulations and prescriptions put forth for most evaluation approaches, including theory-driven forms of evaluation" (Coryn et al., p. 215). Recently, Rog (2015) investigated the concern of theory and practice. The author indicated "one idea for continuing to foster the integration between theory and practice is to identify ways in which more systematic reporting of how we designed and implemented our studies can be built into our reports and articles" (Rog, 2015, p. 234). In sum, the results of these studies indicate that it is crucial to clarify how and when the integration of evaluation theory and practice takes place in reality. More effort has to be evident so that practicing evaluators may have the necessary tools to provide sound judgments and conclusions.

Despite the contributions of studies investigating the feasibility and application of theories toward the advancement and support of good practices, other practical concerns have surfaced in the field's conversation regarding this topic. These concerns are the role

and influence of evaluation context and policy on practice (Torchim, 2009). That is, familiarity of the evaluation context and the knowledge of theoretical views are vital when designing and conducting evaluation activities, and should be considered in future evaluation practice (Rog, 2015). A discussion of the role of context and policy in evaluation is in the section below.

The role of context and policy in evaluation. With the growth of social programs and the evaluation profession, understanding the context wherein evaluation takes place is an essential skill of the practicing evaluator (Fitzpatrick, 2012). Evaluators, especially those new to the field, are often faced with challenges related to meeting the needs of programs, diverse stakeholders groups, and the community at large (Chouinard, 2013; Leviton, 2014). These diverse groups have different orientations and conditions that guide and shape how they function. Evaluators, nevertheless, are obligated to meet planned goals to strengthen the performance of the evaluand. Greene (2005) described the term *context* as, “the setting within which the evaluand (the program, policy, or product being evaluated) and thus the evaluation are situated. Context is the site, location, environment, or milieu for a given evaluand” (p. 83). Understanding the influence of the circumstances that surround the evaluand (e.g., the political, cultural, and social directives) is necessary in the life of any evaluation project (Fitzpatrick, 2012). Essentially, awareness of the evaluand’s contextual factors plays an important role to enhance and fulfill the success of the evaluation efforts.

One of the most significant contextual factors in any evaluation effort is the evaluation policy. Torchim (2009) defined the term *evaluation policy* as, “any rule or principle that a group or organization uses to guide its decisions and actions when doing

evaluation” (p. 16). In most cases, these rules and principles are influenced by the political and cultural contexts of the organization. Given the importance of political and cultural contexts in evaluation practice, the practicing evaluator is often under great pressure in regard to what evaluation methods are used (Chouinard, 2013; Christie, 2003). Accordingly, evaluation policies not only reflect current organizational context, but they also have an impact on future methods, preferences, and choices (Chouinard, 2013; Christie & Fleischer, 2010).

Evaluation methodology. As social and educational programs were developed, the desire for specific evaluation principles and techniques grew. Early in evaluation history, the involvement of social scientists enhanced reliance on the use of social science methods and perspectives to assess evaluation activities (Rossi, Lipsey, & Freeman, 2003). Although evaluation is performed through the lens of the social sciences, its activity is seen as a mixture of social science methods and the evaluand’s contextual dimensions, which are both used to eventually arrive at a value conclusion (Scriven, 1991; Shadish et al., 1991). The application of social science methods, however, has been particularly advantageous to the evaluation field. The essential task of program evaluation is to assess the performance of particular evaluand, and then compare it with established standards or criteria for final judgment. Thus, the systematic assessment of an evaluand’s performance entails the use of valid and reliable measures to enhance its creditability and document the evaluand’s success (Rossi et al., 2003).

Evaluation methodology has gained substantial attention within the evaluation profession. There are many conceptual and practical evaluation philosophies that are considered when discussing methodology. One of the most important aspects of the

discussion is the value of the evaluand. In spite of the central role of valuing in various phases of any evaluation, practicing evaluators, early in the field, did not pay attention to its importance (Shadish et al., 1991). Recently, Kallemeyn, Hall, Friche, and McReynolds, (2015) studied three theoretical aspects of evaluation practice (i.e., methods, value, and use) and found that value was the one of the domains less commonly addressed in the sample. This finding potentially has a negative affect on the overall outcome of an evaluation, as values are dimensions used in conjunction with social science methods to ultimately derive and present relevant evaluative conclusions (Davidson, 2005).

Michael Scriven, as a supporter of the “science of valuing” in evaluation (Shadish et al., 1991), has advanced the *logic of evaluation* concept as a distinct, practical, and guiding mechanism for evaluators to reach defensible conclusions and value judgments for their evaluative activities (Fournier, 1995). Specifically, Scriven (1995) described the logic of evaluation as “...the specific principles of reasoning that underlie the inference processes in all and only the fields of evaluation” (p. 49). The four general steps of the logic of evaluation are as follows: “(1) establishing criteria of merit, (2) constructing standards, (3) measuring performance and comparing with standards, and (4) synthesizing and integrating data into a judgment of merit or worth” (Fournier, 1995, p. 16). These four steps help to identify the extent to which evaluation methodology can be applied in any given evaluation. Davidson (2005) specified three similar aspects of evaluation methodology: (1) identifying and assessing the important of evaluation, (2) determining the merit or quality of the evaluand, and (3) collecting evidence to form an overall judgment of the evaluand.

Methodological preference has an essential role in the design of an evaluation. In essence, the background and skills of the evaluators and the level of stakeholders' involvement may determine the methodology of the evaluation. Azzam (2011) studied the influence of evaluator background and values on evaluation methodological preferences. The finding was that evaluators who tended to be objective were most likely to use rigorous methods that were experimental or quasi-experimental. However, evaluators who were more influenced by their subjective views tended to use qualitative designs. These findings indicate that the methodology choices, whether influenced by personal values or value of others (e.g., stakeholders), are crucial for the credibility and quality of the evaluation, and for evaluators as well.

Evaluation approaches. Dialogue concerning experimental and quasi-experimental methods dominated the field of evaluation for a period of time, especially in educational program evaluation (Donaldson, Christie, & Mark, 2008). In particular, growing interests of the evaluation community early on focused on further developing existing tools and creating others to evaluate program effectiveness in a quantitative manner. This awareness resulted in a number of standalone evaluation approaches and methods used to guide the evaluation performance and judgments. Stufflebeam and Coryn (2014) organized these developed evaluation approaches into groups based on specific their purposes. The authors identified more than 23 approaches most frequently used in evaluation today. Similarly, Alkin (2004) grouped the theorists along with their contributions on three distinctive dimensions, called “theory tree” (p. xi). This theory tree contains the field’s theorists’ perspectives pertinent to the evaluation methods, value, and use.

In summary, in spite of the good standing of the evaluation methodology, methods, and their uses in practice, the need for quality and guiding mechanisms to use them appropriately remains imperative. That is, evaluators need certain methods to ensure the accuracy of evaluative conclusions, and to increase credibility when doing evaluation. The quality of evaluation performance is significant to the advancement of the field. Quality aspects can be achieved through the knowledge and application of the evaluation standards in real world practice.

Evaluation Standards

Attempts to conduct well and high quality evaluations are one means of increasing the credibility of an evaluation and the profession at large (Ruhe & Boudreau, 2013). Prior to the period of professionalizing the evaluation field, guiding standards to check the quality of evaluation were not widely considered in evaluation practice. Moreover, it has been noted that low quality evaluations were perceived in the funded federal program evaluations (Stufflebeam & Coryn, 2014).

The need for quality mechanisms to guide and inform evaluation practice is crucial as the field continues to grow. As stated, evaluation efforts have a unique nature associated with its contexts, uses, needs, and values of the stakeholders (Ruhe & Boudreau, 2013). Conducting evaluations with the intent to use a specific set of standards is one way to guide the evaluation process and ensure the quality. Thus, two recognized sets of standards were established and developed in the evaluation field to assist with this process: The Joint Committee on Standards for Educational Evaluation (Yarbrough et al., 2011) and the American Evaluation Association's Guiding Principles for Evaluators (Newman et al., 1995).

The Joint Committee on Standards for Educational Evaluation. At the end of the 1960s, the evaluation community sought to fulfill a need to systematically investigate the quality of evaluations. This intention to ensure evaluation quality led evaluation professionals to form the Joint Committee on Standards for Educational Evaluation (JCSEE) in 1974. The inception of the JCSEE had an influential role on the establishment and development of professional evaluation standards related to programs, personnel, and students evaluations. The first standards related to program evaluation emerged in 1981, with two revisions issued in 1994 and 2011. During this time, the first standards related to personnel evaluation were issued in 1987, and second version emerged in 2009. Finally, the first standards for student evaluation emerged in 2004, and the second version is still under development (Stufflebeam & Coryn, 2014; Yarbrough et al., 2011).

The focus of the following paragraphs is on the Program Evaluation Standards (Yarbrough et al., 2011), as the accuracy portion of these standards was utilized for this study. In 1981, the JCSEE published the first set of Standards for Evaluations of Educational Programs, Projects, and Materials. In 1994, the JCSEE updated the first version, which led to the second revision, titled the Program Evaluation Standards. The second issue contained 30 standards that correspond to four main categories: utility, feasibility, propriety, and accuracy standards (Stufflebeam & Coryn, 2014).

The third version of the Program Evaluation Standards (Yarbrough et al., 2011) emerged in 2011, with an addition of evaluation accountability standards. Therefore, the current development of the JCSEE Program Evaluation Standards contains 30 standards stemming from five main categories that should be addressed by an evaluation. These five categories are evaluation utility, feasibility, propriety, accuracy, and accountability.

The fifth category added to the third version of Program Evaluation Standards addresses the role of metaevaluation, whether conducted internally or externally (Yarbrough et al., 2011).

Since the establishment and development of these standards took place in North America, their applications are limited to evaluations context and cultures in the U.S. and Canada (Stufflebeam & Coryn, 2014). Although the use of program evaluation standards is not mandatory, they are an optional process, which may assist guiding the evaluation at all phases (i.e., design, implementation, and reporting and disseminating and using the findings)(Ruhe & Boudreau, 2013). One of the productive methods to apply these standards is the use of checklists. The checklist is an important tool used to assist the practicing evaluators memorizing the standards, and ensures the comprehensive coverage of the standards intended for use (Yarbrough et al., 2011).

The American Evaluation Association's Guiding Principles for Evaluators.

In 1982, the Evaluation Research Society (ERS) published its second version of standards aimed for use in program evaluation endeavors. The ERS's standards contained 55 statements corresponding to six groups. These six groups include the following: (1) formulation and negotiation, (2) structure and design, (3) data collection and preparation, (4) data analysis and interpretation, (5) communication and disclosure, and (6) use of results (Stufflebeam & Coryn, 2014).

In 1986, the ERS and the Evaluation Network (ENet) were combined to form the American Evaluation Association (AEA, 2004). When the merging occurred, the AEA did not formally use the ERS's standards. Instead, AEA board sought to create its own guidelines. In 1994, a Task Force created a general draft of guiding principles for

evaluators (Newman et al., 1995). In 2004, after members' reviews and agreement on the first two revisions, the final draft of the guiding principles for evaluators was emerged and approved for use. The final version includes five guiding principles as follows: (1) systematic inquiry, (2) competence, (3) integrity and honesty, (4) respect for people, and (5) responsibility for general and public welfare (AEA, 2004). These guiding principles can be used in many disciplines to guide the conduct of sound evaluations (Stufflebeam & Coryn, 2014).

The JCSEE standards are central to education-based evaluations. Although the JCSEE standards could be possibly useful in other non-educational domains, the AEA Guiding Principles are more broadly utilized with most program evaluations (Stufflebeam, 2001). Evaluation standards are not an obligation required to conduct an evaluation; however, they are recommended to increase sound evaluations and credible results (Stufflebeam & Coryn, 2014; Yarbrough et al., 2011). Practicing evaluators, therefore, who have an affiliation with the AEA community are strongly encouraged to use the AEA Guiding Principles in their practice in order to yield credible and reliable findings (Stufflebeam, 2001).

Metaevaluation: Concept and Practice

Many evaluation professionals argue that recognition, understanding, and application of professional standards in practice should be part of the competence requirements for professional evaluators (Stevahn, 2005). In particular, the rapid advances of the evaluation field increased awareness of the need to check the extent to which evaluations are sound and justified. The methodology of applying the program evaluation standards can be done through different means. Metaevaluation is considered

one of the more productive methods used to practically apply program evaluation standards (Miller, 2008; Stufflebeam & Coryn, 2014; Yarbrough et al., 2011). More than four decades ago, metaevaluation was introduced as a valuable tool to assist in checking evaluation efforts (Stufflebeam, 2001b). Originated by Scriven in 1969, *metaevaluation* was first defined as “the evaluation of evaluations” (Scriven, 1991, p. 228). Davidson (2005) elaborated on Scriven’s definition by stating that metaevaluation “is a determination of the quality and/or value of evaluation” (p. 205).

Scriven’s definition and subsequent elaborations focused on the educational context; therefore, there was a need to expand this notion to include the many activities of evaluation (Stufflebeam & Coryn, 2014). Stufflebeam and Coryn (2014) extended the concept further to involve various conceptual and practical domains. They operationally defined metaevaluation as “the process of delineating, obtaining, and applying descriptive and judgmental information about an evaluation’s utility, feasibility, propriety, accuracy, and accountability for the purposes of guiding the evaluation and reporting its strengths and weaknesses” (p. 635). In this definition, there is a clear association between conducting metaevaluation and using the JCSEE program evaluation standards.

The uses of metaevaluation often differ in the life of an evaluation project. *Formative* or *proactive* metaevaluations are those activities conducted at the beginning stages of an evaluation to check for errors and plan for resolution strategies before or during the cycle of the project. *Summative* or *retrospective* metaevaluations on other hand are used upon the completion of the evaluation for accountability and stakeholders’ judgment purposes (Scriven, 1991, p. 229; Stufflebeam & Coryn, 2014, p. 634). Although these two purposes of metaevaluation are the most common in the evaluation

literature, *concurrent* metaevaluation was recently introduced as another form of metaevaluation. According to Hanssen, Lawrenz, and Dunet (2008), concurrent metaevaluation is “(a) is conducted simultaneously with the development and implementation of a new evaluation method; (b) has both formative and summative components; (c) is comprehensive in nature; and (d) includes multiple, original data collection methods” (p. 575).

Despite the fact that performing metaevaluation is a valuable and beneficial process for the credibility of evaluators’ work and even for the evaluation field, its practice, methods, and application need more development efforts (Cooksy & Caracelli, 2005, 2009). Cooksy and Caracelli (2005) indicated that metaevaluation could be used to assess one or many evaluations. In their metaevaluation of multiple evaluation reports, they found three important dimensions metaevaluators should be careful with when conducting such activities. These aspects contain (a) the way quality is defined, (b) the role of political and cultural values of the entity, and (c) the intention to use the metaevaluation findings (Cooksy & Caracelli, 2005). Cooksy and Caracelli (2009) analyzed 18 metaevaluations to show and validate how metaevaluation is practically implemented. The authors found that the practice and application of metaevaluation, in reality, varies. Specifically, setting criteria and employing methods varies across studies. The authors concluded the metaevaluations in their study had “a lack of clarity about what constitutes a metaevaluation” (Cooksy & Caracelli, 2009, p. 10).

Though there is confusion surrounding the practical implementation of metaevaluations, there have been many efforts to identify methods and techniques to simplify the process. Stufflebeam and Coryn (2014) suggested a set of tasks

metaevaluators could utilize at different stages of metaevaluation activity. Additionally, besides the availability of checklists as valuable tools to assist in performing a metaevaluation (such as the Key Evaluation Checklist (Scriven, 2007) and the Program Evaluation Metaevaluation Checklist (Stufflebeam & Coryn, 2014) among others), Davidson (2005) suggested other alternative methods to conduct metaevaluation. The alternatives she suggested include *second opinion* and *the hybrid approach*. The former is to compare the evaluation results conducted by different evaluators, while the latter is to inspect the results of the evaluation in regard to the points where there are disagreements (Davidson, 2005, p. 214).

Metaevaluation requires a certain set of skills in order to derive meaningful and justified outcomes and conclusions. Stufflebeam and Coryn (2014) listed six fundamental and necessary skills: (1) standards based knowledge, (2) methodology skills, (3) familiarity of evaluand context and needs, (4) propriety aspects, (5) skills needed for negotiation, and (6) dealing with those involved in the process. The proponents and opponents of metaevaluation generally agree on the vast advantages of such an approach. This study, however, utilized the Program Evaluation Metaevaluation Checklist (Stufflebeam & Coryn, 2014) as a method to assess the evaluations of school counseling programs.

Perceptions of Program Evaluation in the Field of School Counseling

In the 1970s, the school counseling profession recognized the need to restructure school counseling programs from its traditional orientation to a more comprehensive practice (Hatch, 2008; Studer et al., 2011). In particular, the traditional orientation was merely centered on vocational services (Gysbers, 2001; McGamon, Carey, & Dimmitt,

2005). However, notable changes in the nature of students' needs placed a greater burden on those in the field to develop more responsive school counseling models and frameworks (e.g., the Comprehensive Developmental Guidance (CDG) program model and *The ASCA National Model*) (ASCA 2003; Eschenauer & Chen-Hayes, 2005; Gysbers, 2001). These models are now the guiding mechanisms that drive the school counseling profession toward best practices.

Despite the documented successes of comprehensive school counseling programs, the implementation and evaluation of such programs are inconsistently performed across school districts (Gysbers & Lapan, 2001). Moreover, the value and effectiveness of these programs might not be perceived without a proper advocacy and use of evaluation and accountability procedures. Martin and Rallis (2014) noted that school counseling programs with an attention to evaluation and accountability activities are more successful than programs with little consideration of such activities. Therefore, there have been successive efforts to promote the use of program evaluation by the school counseling community. The intention to engage in evaluation practice and utilize its available tools is a professional responsibility to confirm the value and effectiveness of school counseling programs (Dimmitt, 2010).

Appeals for considering program evaluation as part of counseling programs have been continuously made in the field. *The ASCA National Model* (ASCA, 2003, 2005, 2012), the widely used framework for school counseling programs, addresses the central role of program evaluation. Additionally, school districts often establish standards and indicators to determine the successful implementation and functionality of comprehensive school counseling programs. Nevertheless, while program evaluation is a

strongly recommended practice to provide results of how these standards and indicators are achieved, much of the profession's emphasis has been devoted to the implementation of school counseling programs (Gysbers & Henderson, 2006; Martin, Carey, & DeCoster, 2009). Thus, evaluation practice still has not been widely observed within many states (Martin & Carey, 2012).

Just a few state of departments of education (e.g., Utah and Missouri Department of Education) support the use of program evaluation. These two state examples contributed to and enhanced the evaluation process in school counseling through the establishment of evaluation and monitoring systems to improve school counseling programs. Prior to this, financial struggles had caused a number of states to reduce or eliminate many counseling programs. But, the successful evaluation experiences in Utah and Missouri helped them to maintain funds for their counseling programs (Martin & Rallis, 2014).

The school counseling community has, for the most part, been largely reactive to the need for valid evaluative tools. Nevertheless, some scholars and professionals in the school counseling field have made considerable contributions to aid the implementation of evaluation procedures in schools. For instance, Carey and Dimmitt (2008) suggested a model of evidence-based practice for school counseling. The model functions through three connected processes: (1) the identification of the student's problem or needs, (2) the determination of the best practice interventions to address the problems or needs, and (3) the evaluation of the program effectiveness. It starts with a needs assessment to identify issues to be investigated. Then, the evaluator chooses among the best tools to address the issue or the needs. Finally, the model suggests evaluation activities to confirm the

effectiveness and efficiency of the program. The role of program evaluation here is expected to provide a judgment of the value of the programs being evaluated. The findings are then more likely to be valuable in reporting the program's success, areas for improvement, and/or accountability.

Program evaluation is not only an effective method to improve counseling programs, but it is also a defensible approach to legitimize these programs to policymakers for funding and accountability purposes (Carey & Dimmitt, 2008; Gysbers & Henderson, 2006; Hatch, 2008). For example, the Utah Department of Education uses program evaluation activities to report the extent to which counseling programs meet established achievement standards, which then justify the funds for these programs (Gysbers & Henderson, 2006). Similarly, Duarte and Hatch (2014) presented a case study of one of California's school districts that was awarded a federal grant. Evaluation results were shown to be the supportive mechanism to continue and expand counseling programs in four elementary and middle schools (Hatch, 2008). The aforementioned usage of evaluation implies that the evaluation process assists in maintaining needed financial assistance.

Many school counseling professionals have continued to find ways to improve and support the evaluation practice in school counseling context (Martin, Carey, & DeCoster, 2009; Martin & Rallis, 2014). However, there are challenges that may hinder the full implementation of school counseling program evaluation. One of these challenges is related to the resources needed to do the evaluation task. There are a number of necessary resources to implement an appropriate program evaluation. One of the resources is pertinent to financial aid. Funding sources for school counseling vary; funds

may come from federal government or non-federal government agencies (e.g., foundations or other agencies) (Lum, 2005). Often, school districts gain funding as a grant awarded from the federal government, such as the Elementary and Secondary School Counseling (ESSC) program. Primarily, the aim of this type of funding is to launch more school counseling programs and improve the quality of counseling services that are provided to students. Thus, program evaluation is one of the procedures required to report the results of the performances and improvements to maintain these counseling programs and their funding (Duarte & Hatch, 2014).

Other challenges may affect evaluation activities in schools as well. In a study conducted by Astramovich et al. (2005), a sample of school counselors reported time as one of the major difficulties in conducting evaluation. Similarly, a study conducted by Poynton (2013) found time and knowledge were the biggest obstacles preventing school counselors from collecting data and reporting the program effectiveness. Another challenge that school counselors may encounter is a lack of needed support from leadership and administration. In fact, Astramovich et al. (2005) also found that there is little assistance from the school staff and administrators when school counselors tend to evaluate their counseling programs.

In spite of evidence of a perceived lack of support by school administrators for evaluation activities in some studies (e.g., Astramovich et al., 2005), there are many efforts to encourage the evaluation practice within the field within the field at large. Such efforts are perceived through the establishment of incentive and resource programs. Incentive programs are provided according to evaluation endeavors and results (Martin & Carey, 2012). The evaluation initiatives in Utah and Missouri are two examples of states

that established an incentive program for those districts and school counselors performing evaluations of their counseling programs (Martin & Rallis, 2014). Also, the American School Counselors Association created an award program. This program is the Recognized ASCA Model Program (RAMP). RAMP recognizes the most successful schools that use evaluation to improve their school counseling programs. The award is also a method to encourage and support those schools that use of *The ASCA National Model* as the framework for their counseling programs (Young & Kaffenberger, 2011).

There are also two types of rewards given to those who advocate for program evaluation. These rewards are the Gysbers and Success awards (Martin & Carey, 2012, p. 139). Overall, it is apparent that there are continuous efforts to make program evaluation a recognized practice within the field. Furthermore, the recognition of program evaluation has increased with the establishment of accountability-based practices and demands in educational movements, especially in the school counseling context.

Accountability Practices of School Counseling Programs

With the inception of organized and comprehensive school counseling models, accountability became an endorsed concept to validate the viability and effectiveness of school counseling programs (Dahir & Stone, 2003; Palmer & Erford, 2012). Program audit or evaluation as stated in the ASCA requirement is a practical tool used to identify the strengths and weaknesses of counseling programs regarding four major components—accountability is the fourth component (ASCA, 2005; Palmer & Erford, 2012). The concept of accountability, therefore, has gained serious consideration in the school counseling profession, and a growing component of school counseling programs in general.

According to the literature, there are some misunderstandings in the school counseling profession as it relates to accountability and program evaluation (Astramovich & Coker, 2007). Astramovich and Coker (2007) addressed some of these misunderstandings. A primary issue of concern the authors noted is that two concepts are often used interchangeably. However, accountability is one of the evaluation purposes, which often used through the summative evaluations (Davidson, 2005). Program evaluation is a method used to document and improve the effectiveness and efficiency of counseling programs. Accountability may be accomplished when counselors evaluate their programs and report the results to the stakeholders to facilitate the decision-making process (Astramovich & Coker, 2007).

Dahir and Stone (2003) proposed an accountability model that supports the movement of evidence-based practices within school counseling programs. This model, M.E.A.S.U.R.E, contains seven elements: (1) Mission, (2) Elements, (3) Analyze, (4) Stakeholders, (5) Unite, (6) Reanalyze, and (7) Educate. The model emphasizes the use of data through these seven steps and strategies to exhibit the effectiveness of counseling programs and school counselors. It also stresses the need to align counseling programs with the overall improvement and accountability goals of school.

The first step of the M.E.A.S.U.R.E. model focuses on aligning the goals and implementation plans of counseling programs with the school mission. School counselors collaborate with school stakeholders to plan counseling programs to serve better the needs of all students. The second step attends to the collection of essential data about the students and counseling programs. The gathered data are then used to show the stakeholder group the overall developments and needs of students and programs. Here,

school counselors should incorporate different methods to obtain essential data. To this end, program evaluation is one of the methods that may be utilized to examine the effectiveness and impact of the programs and counselors' efforts. The third step analyzes the data elements to identify the effects or problems that students or programs may encounter (Dahir & Stone, 2003).

The fourth and fifth steps of the M.E.A.S.U.R.E. model take an account of the stakeholder groups that should be involved in the process. The model suggests the inclusion of a broad range of stakeholder groups from inside and outside the targeted schools. Once stakeholders are identified, they discuss strategies to develop an action plan. The action plan contains all elements and resources needed for the improvements of programs and the achievement of students. The sixth step tracks the results and changes occurred as a consequence of the preceding steps. This step illuminates what works and what does not work, and then allows for modifications accordingly. The final step demonstrates the effects and impacts of counseling programs. This step ensures the accountability of such programs pertinent to the student success (Dahir & Stone, 2003).

Bemak, Williams, and Chung (2014) proposed another accountability model to demonstrate the effectiveness of school counseling programs. The School Counselors' Domains of Accountability model contains four elements as indicators of program success. These indicators are students' grades, attendance, suspension, and disciplinary referrals. The authors affirm that such accessible data are beneficial to determine the effectiveness of the counseling programs and services. Furthermore, the successful use of this accountability tool could increase the credibility of counseling programs to key policymakers for decision-making process.

Another model, the Accountability Bridge Counseling Program Evaluation Model, was proposed and presented by Astramovich and Coker (2007) to help counselors review the processes and activities involved in program evaluation. The model consists of two primary cycles. Accountability here is a bridge connecting the outcomes of the two cycles. The first cycle is the counseling program evaluation. This cycle contains the evaluation components needed to conduct the program evaluation. This stage begins with planning the evaluation, implementing what was planned, monitoring the evaluation procedures and initial outcomes, and obtaining the evaluation results. Once the outcomes are analyzed and completed, the evaluation results are communicated through the accountability bridge to the stakeholders in the form of an evaluation report.

The second cycle is the counseling context evaluation. In this cycle, counselors present the evaluation results and get feedbacks from stakeholders. Next, counselors and stakeholders proceed to strategic planning to discuss the overall performance of the counseling program. At the strategic planning phase, stakeholders may recommend the allocation of necessary resources and revisions to further meet the program's goals and mission. Once strategic planning is addressed, a needs assessment is conducted to collect data from the stakeholders regarding the development and needs of counseling programs and services. The final stage of this cycle involves refining and setting program objectives. Once these objectives are determined, changes may be considered in the planning phase of the first cycle. This model promotes the continual process of evaluation between cycles. Hence, establishing clear guidelines for counselors to perform evaluation activities and report the results could enhance their abilities to engage confidently in evaluation practice (Martin & Rallis, 2014).

Finally, the success of program evaluation practice in the school counseling context would not be possible without the engagement of school counselors in the role of “change agent” (ASCA, 2005; Hays, 2010). Despite many attempts and tools to fully engage professional school counselors in the evaluation process, a number of factors may influence their involvement in the program evaluation process. One of the main obstacles school counselors face to conduct school counseling program evaluation is the unavailability of the necessary preparation and training in evaluation (Astramovich et al., 2005; Trevisan, 2000).

Preparing School Counselors in Evaluation

The field of school counseling has made serious attempts to highlight the importance of the role of school counselors in educational movements. According to Sink (2005), school counselors are a vital part within the educational system. Yet, in order to be seen as a vital part of the educational system, one of the school counselor’s expected duties is to show how their counseling programs are effective (Astramovich et al., 2005). Furthermore, with increasing federal and state mandates for evaluation and accountability, the expected involvement of school counselors in the evaluation process becomes crucial (Martin & Rallis, 2014).

The defined position of school counselors in *The ASCA National Model* (ASCA, 2005, 2012) and almost all state educational models requires the performance of evaluation work. Also, the school counseling literature often asserts the obligation of school counselors to evaluate counseling programs (Martin & Rallis, 2014). As Sink (2009) noted, “To serve as accountability leaders, school counselors have to take more seriously their evaluator role” (p. 72). However, it is difficult to expect evaluation

knowledge and activities from counselors while minimally supporting the development of evaluation preparation and training programs. This status is true despite the pre-service and in-service emphasis on the preparation of school counselors in evaluation skills throughout the field (Maras, Coleman, Gysbers, Herman, & Stanley, 2013)

Astramovich et al. (2005) studied the accountability and program evaluation practices of 28 school counselors in a school district. The findings were that interests and willingness to conduct program evaluation were high among school counselors. However, the biggest concern among school counselors was the need for evaluation training. More than half of the sample (75 percent) indicated the need for more training opportunity about evaluation methods.

Similarly, a statewide study was conducted to discover the training needs of 166 school counselors in Wisconsin (Burkard, Gillen, Martinez, & Skytte, 2012). This study revealed that evaluation was the least performed activity among counselors. The authors indicated that undeveloped evaluation skills were associated with the observed performance. Additionally, Trevisan (2002) reviewed the evaluation portion of certification requirements for school counselors in many states. He found that school counselors might not be learning the appropriate evaluation knowledge. The author attributed this lack of evaluation learning to the observed misunderstanding of the differences between research and evaluation.

Overall, the findings of the above-mentioned studies (i.e., Burkard et al., 2012; Trevisan, 2002) indicate school counselors might value and have interest in evaluation principles and concepts, but acquiring the skills for real application needs continued attention from the entire school counseling community. The concern of evaluation

preparation concerns was clearly evident in Trevisan's (2000) calls for more evaluation efforts in the school counseling field. He stressed that the scarcity of program evaluation activities may be primarily due to the absence of evaluation preparation and training (Trevisan, 2000).

Chapter II Summary

Reforms in departments of education and professional school counseling organizations clearly indicate a need for school counselors who are able to evaluate the effectiveness of their school counseling programs (e.g., ASCA, 2005, 2012). The educational environment usually aims to provide quality services for student success and development. To connect to this objective and show the capacity to serve the needs of students, school counselors may need to accelerate their efforts in acquiring the necessary evaluation knowledge and skills (Dimmitt, 2010). Counselors also need these skills for instances when resources for external evaluation might not be available to the unstable economy (Martin & Rallis, 2014).

There are many helpful resources available for school counselors to learn program evaluation. For instance, many textbooks may facilitate the learning process of evaluation principles and methods. Also, there are online resources available such as the American Evaluation Association along with its journals (i.e., *American Journal of Evaluation and New Directions for Evaluation*) and annual conferences (American Evaluation Association, 2016). Also, there are evaluation resources available in the school counseling field, such as The National Center for Education Evaluation and Regional Assistance, and The Center for School Counseling Outcome Research (Dimmitt, 2010). The connection with the evaluation community is a worthy effort to gain the necessary

assistance and recommendations needed to conduct program evaluation in the school counseling context. Specifically, the results obtained in this dissertation's metaevaluation of school counseling program evaluations may assist counselors in determining common areas for further training and preparation as it relates to program evaluation.

CHAPTER III

METHODOLOGY

Chapter III describes the methodology used to answer the research questions of this study. The section below restates the purpose of the study and its research questions as originally presented in Chapter I. Then, the remaining aspects of the methodology are discussed as follows: (a) design, (b) sample, (c) instrumentation, (d) data collection, and (e) data analysis.

Purpose of the Study and Research Questions

Program evaluation is an important aspect of comprehensive school counseling programs (ASCA, 2012). Yet, there is a dearth of literature exploring how program evaluation is used to demonstrate the effectiveness of CSCPs. Furthermore, there is also an absence of literature exploring the quality of existing program evaluations within the field of school counseling. Therefore, the purpose of this study was to assess the accuracy of program evaluations performed thus far in the school counseling profession by conducting a metaevaluation. The Program Evaluation Metaevaluation Checklist (Stufflebeam & Coryn, 2014) was used to inspect how evaluations conform to the JCSEE's accuracy standards (Yarbrough et al., 2011). Moreover, semi-structured interviews performed with the authors (i.e., evaluators) of evaluations to uncover their views about their program evaluations' strengths and weaknesses, evaluators' methodology choices, and other factors influencing program evaluation practice in the context of school counseling. Accordingly, this study attempted to answer the following research questions:

1. How well do school counseling program evaluations conform to the JCSEE accuracy standards (Yarbrough et al., 2011)?
2. What are the perceptions and experiences of the participating evaluators regarding the status and accuracy of school counseling program evaluation?
 - a. What do evaluators perceive as the strengths and weaknesses of school counseling program evaluation?
 - b. How do evaluators align their evaluations to existing evaluation methods and practice?
 - c. What are the major factors influencing program evaluation practice in school counseling context?

Research Design

This study utilized a sequential two-phase design to describe the nature and accuracy of program evaluation within the school counseling context. To this end, metaevaluation checklist and semi-structured interview strategies were employed. Metaevaluation was used to answer research question 1 specifically, as it was designed to assess compliance of existing CSCP evaluations with the accuracy standards contained within the JCSEE program evaluation standards (Yarbrough et al., 2011). Metaevaluation is a valuable process used to assess an evaluation once it is completed to ensure that the evaluation is sound, systematic, and producing accurate findings (Cooksy & Caraceli, 2005; Stufflebeam & Coryn, 2014). Simply put, metaevaluation allows one to assess the quality of an evaluation. The accuracy assessment in school counseling program evaluations is at the center of this study, making metaevaluation through the application of JCSEE program evaluation standards, the accuracy standards specifically, the most

appropriate methodology to answer what is currently unknown about the dependability of program evaluation in the school counseling profession.

Semi-structured telephone interviews were executed to address research question 2, with the aim to explore evaluators' perceptions of and experiences with program evaluation practice in the school counseling context. The interviews were primarily intended to discover strengths, weaknesses, methodologies, and key factors affecting the practice and performance of program evaluation in school counseling. The use of this method allowed the participants to share their experiences and perceptions about the phenomenon under study (Marshall & Rossman, 2006). In turn, combining what participants revealed about their experiences with the accuracy ratings from the metaevaluation checklist assisted with the interpretation of information obtained through this process, further facilitating the investigation of the value and nature of evaluation practice in the school counseling context.

Sample

The study had two different samples and sampling techniques. As described in the Research Design section, this study utilized both metaevaluation and interview procedures to answer research questions aimed at assessing evaluation practice in the school counseling field. The sample and sampling technique for each procedure is described below.

Metaevaluation

The sample of the metaevaluation consisted of published school counseling program evaluations. Such evaluations are mainly studies conducted in the field of school counseling examining the effectiveness of the counseling programs, services, and

activities implemented in elementary, middle, and high schools; that is, the evaluations are primarily conducted to assess the implementation and impact of school counseling programs on students' academic, personal/social, and vocational development areas at all grade levels. An extensive search of various databases was completed between December 1, 2014 to end of January 2015 to arrive at a comprehensive list of all available evaluations. The initial search strategy was performed through formal databases, such as PsycINFO, ERIC, ProQuest, Web of Science, Science Direct; school counseling associations, and departments of education at state and district levels. The key terms used to search these evaluations included *school counseling*, *school guidance*, *program evaluation*, *program effectiveness*, *counseling interventions*, and *counseling services*. The initial search strategies resulted in 128 evaluation reports.

The researcher developed inclusion criteria, depicted in Table 1, based on the literature review for evaluations to be eligible for this study. As previously stated, the practice of program evaluation in the school counseling field is growing (ASCA, 2012; Martin et al., 2009; Trevisan, 2002). As a result, the timeframe of reports ranged from 2005 to 2014 in order to locate a larger number of evaluations for systematic review and analysis. Additionally, the evaluations had to comprise programs, interventions, services, and activities provided to students in three areas of development as specified in *The ASCA National Model* (ASCA, 2012), which are academic, career, and personal/social. The methodology section of the evaluations had to involve a full description of the methods and approaches used in the evaluation process. This criterion allowed the researcher to perform an assessment of the evaluation's accuracy against the program evaluation standards (i.e., Yarbrough et al., 2011). Finally, all evaluations had to have

been conducted at all school levels (elementary, middle, and high schools), published in the United States, and written in English. Evaluations that only contained literature reviews, incomplete descriptions of the programs or interventions, or those that did not adequately describe the procedures used to conduct the evaluation were excluded from the sample.

Table 1

Inclusionary Criteria for Evaluations

Key Criteria	Description
Timeframe	Evaluation reports published from 2005 to 2014
CSCP characteristics	Included evaluations must contain counseling programs, interventions, services, and activities that were conducted in the following area (ASCA, 2012): <ul style="list-style-type: none"> • Academic development • Career development • Personal and social development
Result information	Detailed information must be present in the evaluation to allow for assessing the methods employed.
Setting and language	<ul style="list-style-type: none"> • School levels include: elementary, middle, and high schools. • Evaluation reports must be conducted in the United States. • The language used in the reports must be written in English.

Thus, once these criteria were applied to the 128 reports obtained previously, half of the evaluations (64) were eligible for inclusion. Of the 128 reports, 34 reports were excluded because they were merely literature reviews, 12 reports were conducted outside the U.S, and the remaining 18 reports did not contain the school counseling programs required for the study's purpose. Also, prior to the coding process, the two coders agreed to exclude 10 evaluations that had insufficient descriptions of the programs and

procedures needed to conduct the metaevaluation process. Therefore, 54 reports were eligible, meeting the study's criteria. These obtained evaluations were conducted at the three school levels as indicated in the study's criteria. Table 2 shows the evaluation reports pertinent to each school level. Also, Table 3 shows the school counseling program areas where the evaluations were conducted. The combined programs category indicates the integration of academic, career, and/or personal/social programs in one evaluation report.

Table 2

Evaluations Based on School Levels

School Levels	Sum of Evaluations	Percent of Evaluations
Elementary	17	31%
Middle	13	24%
High	24	44%
Total	54	100%

Table 3

Evaluations Based on Domains of School Counseling Programs

Program Domains	Sum of Evaluations	Percent of Evaluations
Academic	15	28%
Career	5	9%
Personal/social	27	50%
Combined programs	7	13%
Total	54	100%

Semi-structured Interviews

Purposeful sampling was utilized to select the sample to participate in the semi-structured interviews. Purposeful sampling is a nonrandom method used in accordance to prior knowledge of the population and the particular purpose of the research (Fraenkel & Wallen, 2006). Participation in the second phase depended primarily on the results of the first phase metaevaluation process. Subsequently, eligible participants for this study were authors whose evaluations were rated high and low on the Program Evaluation Metaevaluation Checklist (Stufflebeam & Coryn, 2014). These participants were expected to be school counselors, counselor educators, university researchers or professors, or researchers or evaluators from outside agencies. The initial sample size determined for this study was 10 participants; five participants from each rated category of the metaevaluation results; and more participants were expected to join as the interviews evolve. However, the response rate of participants was lower than expected. Invitations were sent to all 54 evaluators to obtain the determined 10 participants. Although three reminders were sent to the evaluators, only four participants were willing to participate in the interviews. Three other evaluators replied back with their rejection to participate. Consequently, three counselor educators and one school counselor comprised the interview sample. The gender of the participants was three male participants, and one female. Three participants were counselor educators who worked in academic based occupations; however, one counselor educator previously held a school counselor position at a school. Their experiences in evaluation ranged from 2 to 20 years (Median = 9.75).

Instrumentation

The study employed two instruments to collect the desired data. The metaevaluation data were collected using the Program Evaluation Metaevaluation Checklist (Stufflebeam & Coryn, 2014). The checklist was developed to judge how evaluations conform to the JCSEE program evaluation standards (Yarbrough et al., 2011). It consists of 30 sub-standards divided into five categories: (a) utility, (b) feasibility, (c) propriety, (d) accuracy, and (e) evaluation accountability. This study only used the accuracy standards to assess the methodological features of the selected evaluations (Appendix A). Also, the initial review of evaluation reports indicated that the reports had insufficient information to apply to the other four Program Evaluation Standards.

There are eight accuracy standards in the checklist. Each sub-standard of the accuracy standards includes six checkpoints and rating scales to assess the evaluation against the respective sub-standard's statement. The checkpoint was marked with a plus (+) if the necessary information was present in the checkpoint, a minus (-) if the information was not present, and a question mark (?) if there was not sufficient information to support the judgment. At the end of each standard assessment, the number of pluses were summed and linked to one of the ratings on a scale that ranged from *Excellent* if all checkpoints are met, to *Poor* when one or none of the checkpoints are met. Table 4 is an example of the first accuracy standard in the checklist.

Table 4

Example of Program Evaluation Metaevaluation Checklist-Accuracy Standard

The accuracy standards are intended to ensure that an evaluation employs sound theory, designs, methods, and reasoning in order to minimize inconsistencies, distortions, and misconceptions and produce and report truthful evaluation findings and conclusions.				
A1 Justified Conclusions and Decisions. <i>[Evaluation conclusions and decisions should be explicitly justified in the cultures and contexts where they have consequences.]</i>				
<input type="checkbox"/> Address each contracted evaluation question based on information that is sufficiently broad, deep, reliable, contextually relevant, culturally sensitive, and valid <input type="checkbox"/> Derive defensible conclusions that respond to the evaluation's stated purposes, e.g., to identify and assess the program's strengths and weaknesses, main effects and side effects, and worth and merit <input type="checkbox"/> Limit conclusions to the applicable time periods, contexts, purposes, and activities <input type="checkbox"/> Identify the persons who determined the evaluation's conclusions, e.g., the evaluator using the obtained information plus inputs from a broad range of stakeholders <input type="checkbox"/> Identify and report all important assumptions, the interpretive frameworks and values employed to derive the conclusions, and any appropriate caveats <input type="checkbox"/> Report plausible alternative explanations of the findings and explain why rival explanations were rejected				
<input type="checkbox"/> 6 Excellent	<input type="checkbox"/> 5 Very Good	<input type="checkbox"/> 4 Good	<input type="checkbox"/> 2-3 Fair	<input type="checkbox"/> 0-1 Poor

To conduct the semi-structured interviews, an interview protocol was developed to guide the process for collecting the data from the study's participants (Creswell, 2007). The semi-structured interview is an adequate type of interviewing process that helps uncover the topics of interest while allowing participants to add new and in-depth perspectives along the process (Galletta & Cross, 2013). As Jacob and Furgerson (2012) stated, "Interview protocols become not only a set of questions, but also a procedural guide for directing a new qualitative researcher through the interview process" (pp. 1-2). The script, contained in Appendix B, initially contain relevant background data, such as the time and date of interview, interviewee's job title and institution, the purpose of the

study, instructions for the interviewers to conduct interviews, and a predetermined set of open-ended questions. All interviews were administered over the telephone. The open-ended questions along with probes and prompts were formed to explore participant perceptions and experiences pertinent to program evaluation practice in school counseling. Additionally, during its development, feedback from the dissertation committee about the protocol structure was sought for further modifications.

Data Collection Procedures

An extensive search of relevant databases was conducted to identify all eligible evaluation studies. Then, the evaluations were thoroughly reviewed to assess them against the program evaluation accuracy standards (Yarbrough et al., 2011). Two raters assessed the reports during the metaevaluation phase. The raters, both males, included the researcher and a former doctoral student in the Interdisciplinary Ph.D. in Evaluation (IDPE) program at Western Michigan University (WMU). Raters began by assessing five studies independently. This independent assessment of five reports assisted in establishing the decision rules to assess the remainder of the reports. Each rater had a copy of the accuracy standards checklist prepared to assess each study against all eight standards.

Prior to conducting the assessment, instructions and rules for applying the checklist were described and clarified. These instructions stated that the following: (a) the checkpoint is marked with a (+) if it is addressed in the evaluation, (b) the checkpoint is marked with a (-) if it is not addressed in the evaluation, and (c) the checkpoint is marked (?) if it is addressed but not clearly stated in the evaluation. At the end of assessing the checkpoints of the standard, the rater summed the scores using the corresponding rating

scale of *Excellent*, *Very Good*, *Good*, *Fair*, and *Poor*. Inter-rater agreements were calculated to ensure consistency of the process for the checklist assessment. The calculated inter-rater reliability (Kappa) between the two coders was 81%.

Once all eligible evaluations were gathered and rated, contact information about the authors (evaluators) was collected for communication purposes. Telephone interview is often used when difficulty exists to directly reaching participants (Creswell, 2007). Next, HSIRB approval was obtained to conduct this second phase of the study (Appendix D). The evaluators were initially contacted via e-mail introducing the purpose of the study and inviting them to voluntarily participate in a semi-structured telephone interview (Appendix E). The invitation emails were sent to participants between September 1 and 25, 2015. Upon their agreement, participants were asked to electronically sign a consent form, which is contained in Appendix C.

An interview was scheduled for each participant from September 10 through 14, 2015. The duration of each interview was approximately 30 to 45 minutes, and audio-recorded with the participant's permission. During the interview, the interviewer started by ensuring the confidentiality of the participants and the process. Then, participants were provided with the option to answer the interview questions and elaborate openly on other perspectives through probing questions. At the end of the interviews, the audiotaped interviews were transcribed on the same day of the interviews to ensure no important information was neglected. Furthermore, all transcriptions and audiotapes were confidentially saved in different locations (Creswell, 2007) with a notecard containing primary information such as date, times, and other information for organizational reasons

(Marshall & Rossman, 2006). The recordings and transcripts were saved in a password remote server in the Evaluation Center at WMU for confidentiality purposes.

Creswell (2007) advocates using strategies to validate the accuracy of the findings. Therefore, peer-debriefing strategy was used to ensure the credibility of the findings in this study (Lincoln & Guba, 1985). *Peer review* or *debriefing* focuses on engaging an outsider to help “exploring aspects of the inquiry that might otherwise remain only implicit within the inquirer's mind” (Lincoln & Guba, 1985, p. 308). Thus, a colleague of the researcher, another doctoral student at Western Michigan University, engaged in this process. A discussion of the findings and interpretation allowed for detecting new perspectives and understanding of the data.

Data Analysis

The data analysis process in this study began by assessing the conformity of the selected CSCP evaluations to the accuracy criteria of the JCSEE program evaluation standards. Next, the interview data were thematically analyzed to report participants’ experiences and perspectives on the nature and accuracy of program evaluation in the context of school counseling.

The first phase involved the raters completing the last process in the checklist by gathering the overall strengths of each study. They obtained total scores as follows:

- Compile all sub-standards ratings of *Excellent* and multiply by 4;
- Compile all sub-standards ratings of *Very Good* and multiply by 3;
- Compile all sub-standards ratings of *Good* and multiply by 2; and
- Compile all sub-standards ratings of *Fair* and multiply by 1

Then, total scores were converted to total percentage that indicates the overall strength of the study. Table 5 shows the template that was used to obtain the total accuracy score of the study. A subsequent analysis was conducted using the same procedures to describe the methodological soundness of all of the studies. Hence, the analysis offered the evaluations' strengths and weaknesses, which would determine their overall quality.

Table 5

Template of Data Analysis for Study's Total Accuracy Strength

Scoring the Evaluation for ACCURACY:		Strength of the evaluation's provisions for ACCURACY:	
Add the following:		[] 29.44 (92%) to 32:	Excellent
Number of Excellent ratings (0-8) ____ x		[] 21.44 (67%) to 29.43:	Very Good
4 = ____		[] 13.44 (42%) to 21.43:	Good
Number of Very Good (0-8) ____ x		[] 5.44 (17%) to 13.43:	Fair
3 = ____		[] 0 (0%) to 5.43:	Poor
Number of Good (0-8) ____ x		(Total score) = %	
2 = ____			
Number of Fair (0-8) ____ x 1			
= ____			
Total score:			
= ____			

The second phase involved the interview data analysis. This process utilized thematic analysis procedures to analyze the data. Specifically, the analytical procedures outlined by Marshall and Rossman (2007) and Creswell (2007) were followed to guide this process. They organize the analysis process into several phases. Upon the completion of each interview, audiotaped recordings were transcribed and saved on a secured password remote server in the Evaluation Center at Western Michigan University. Qualitative data analysis software (ATLAS.ti 7) was utilized to manage and analyze transcribed data more efficiently. The program software (ATLAS.ti7) was used because

of its organized and straightforward features, such as simplicity, capability of reading, reviewing, writing, and coding the data, and analysis functions (Creswell, 2007). Then, the researcher reviewed the transcripts thoroughly and frequently to become familiarized with data. Throughout the review process, statements relevant to the problem being studied were obtained to form the initial themes and categories. Creswell (2007) suggested using a short list of themes that extends as the analysis unfolds. In this study, such themes were attentive to the three objectives of research question 2: (a) strengths and weaknesses, (b) methodology orientation, and (c) key factors affecting the practice and performance of program evaluation in school counseling. However, other sets of themes relevant to the study's purpose emerged from the data as well (Creswell, 2007). The identification of themes was then used to guide and develop the coding process.

Once themes and categories were initially identified, a coding system using different labels and colors was assigned to various segments of the transcripts throughout the documents. Moreover, a list of codes was also generated throughout the process to help uncover other relevant themes as new understanding arose (Marshall & Rossman, 2006). Thus, a codebook was created to incorporate all obtained codes along with their meanings to be used in the coding process (Carey, Morgan, & Oxtoby, 1996).

Additionally, the researcher took memos and notes to assist in apprehending concepts, patterns, and thoughts while reviewing and coding the data (Marshall & Rossman, 2006). Thus, early identification of the patterns in the data via memos was useful in writing the findings (Creswell, 2007). Only the researcher performed the thematic analysis. Although involving another coder is a valuable to ensure the reliability of the qualitative phase, the time constraints did not allow such involvement in the process.

Chapter III Summary

Chapter III described the methodology used to conduct this study. Overall, this study consisted of two phases, a metaevaluation phase and a semi-structured telephone interview phase. The metaevaluation and interview results were incorporated to describe the major findings of the study. The results are intended to describe the accuracy and nature of program evaluation in the school counseling context. The metaevaluation results were performed to reveal compliance to accuracy portion of the program evaluation standards (Yarbrough et al., 2011), while the interview responses uncovered the major perceptions of participants pertinent to program evaluation in the field. The total findings are presented in the following chapter, *Chapter IV, Results*.

CHAPTER IV

RESULTS

This chapter presents the findings of the study. This sequential two-phase study was designed to assess the evaluations of comprehensive school counseling programs. In Chapter I, two research questions were stated. The first research question focused on the extent to which a set of evaluations of school counseling programs conformed to the Joint Committee on Standards for Educational Evaluation (JCSEE) accuracy standards (Yarbrough et al., 2011). The first question was carried out using the Program Evaluation Metaevaluation Checklist (Stufflebeam & Coryn, 2014). The focus of the second research question was to understand the perceptions and experiences of participating evaluators regarding the status and accuracy of school counseling program evaluation. Specifically, the aim of the second research question was to discover the perceptions and experiences of evaluators in regard to (a) the strengths and weaknesses of school counseling program evaluation, (b) the alignment of evaluations to existing evaluation methods and practice, and (c) the major factors influencing program evaluation practice in the school counseling context. Semi-structured interviews were used to engage the evaluators to answer the second research question and related sub-questions. The organization of this chapter begins with the results obtained with the Program Evaluation Metaevaluation Checklist, then moves to the results from the semi-structured interviews, and ends with a discussion of the study's results.

Research Question 1: Metaevaluation of Program Evaluation

The first research question focused on whether school counseling program evaluations conformed to the JCSEE Program Evaluation Standards (Yarbrough et al., 2011), specifically the accuracy standards. There are eight accuracy standards, which focus on “the truthfulness of evaluation representation, propositions, and findings, especially those that support judgments about the quality of programs or program components” (Yarbrough et al., 2011, p. 158). Each standard includes six checkpoints, which were addressed in the evaluations examined in this study.

The following sections answer the first research question by presenting the overall results of the Metaevaluation Checklist (Stufflebeam & Coryn, 2014) for all studies, then presenting the results pertinent to each sub-standard of the accuracy standard.

Program Evaluation Metaevaluation Checklist: Overall

Table 6 shows the template used to obtain the total accuracy score for each evaluation. Appendix F provides the overall accuracy ratings for all 54 evaluations. The overall scores represented the strength of the evaluations for accuracy standard. The total accuracy ratings showed that 74% ($n = 40$) of the evaluations were rated Fair, while 26% ($n = 14$) of the evaluations were rated Poor. Although the evaluations only ranged from Fair to Poor ratings, the percentages displayed considerable variations within each rated evaluation. The overall ratings showed that the highest percentage was 41% while the lowest percentage was 6%. Two evaluation reports reached the highest rate whereas three of the evaluation reports fell below the 10th percentile of the total ratings.

Table 6

Template of Data Analysis for Study's Total Accuracy Strength

Scoring the Evaluation for ACCURACY:		Strength of the evaluation's provisions for ACCURACY:	
Add the following:		[] 29.44 (92%) to 32:	Excellent
Number of Excellent ratings (0-8)	__ x 4 =	[] 21.44 (67%) to 29.43:	Very Good
Number of Very Good (0-8)	__ x 3 =	[] 13.44 (42%) to 21.43:	Good
Number of Good (0-8)	__ x 2 =	[] 5.44 (17%) to 13.43:	Fair
Number of Fair (0-8)	__ x 1 =	[] 0 (0%) to 5.43:	Poor
Total score	= __	(Total score) = %	

Table 7 shows the evaluation ratings related to the school levels. The most rating scores of Fair were given to evaluations conducted at the high school level (83%).

Evaluations obtained at the elementary school level showed the highest percentage (35%) in the Poor category.

Table 7

Evaluation Ratings Based on the School Levels

School Level	Number of Evaluations	Percent of Evaluations	Percent of Fair Ratings	Percent of Poor Ratings	Percent Total
Elementary	17	31%	65%	35%	100%
Middle	13	24%	69%	31%	100%
High	24	44%	83%	17%	100%
Total	54	100%			

Table 8 shows the overall evaluation ratings pertinent to various domains of school counseling programs in which the evaluation took place. Half of the evaluations ($n = 27$) were obtained in the personal and social program area. The least evaluations ($n = 5$)

were in the career development area. Most of the Fair ratings were in the academic and personal/social programs while career had the most ratings in Poor category.

Table 8

Evaluation Ratings Based on the Domains of School Counseling Programs

Domain of Programs	Number of Evaluations	Percent of Evaluations	Percent of Fair Ratings	Percent of Poor Ratings	Percent Total
Academic	15	28%	80%	20%	100%
Career	5	9%	40%	60%	100%
Personal/social	27	50%	81%	19%	100%
Combined programs	7	13%	57%	43%	100%
Total	54	100%			

Program Evaluation Metaevaluation Checklist: Sub-standards

Table 9 shows the results related to A1: Justified Conclusions and Decisions. This sub-standard states, “Evaluation conclusions and decisions should be explicitly justified in the cultures and contexts where they have consequences” (Yarbrough et al., 2011, p. 165). As shown in Table 9, the results indicated that the majority of the studies, ranging from 50 to 53, appropriately addressed the statements items 1 through 3. Those reports addressing the first three items fairly stated the questions and purposes of the evaluation. For item 5, more than half of the studies ($n = 35$) addressed this checkpoint; however, only 14 evaluations confirmed the statements in items 4 and 6.

On the scale pertinent to A1: Justified Conclusions and Decisions shown in Table 10, the ratings of school counseling evaluations were high. Although none of the evaluations had a rating of Excellent, 69% of evaluations had ratings of Good and Very Good. Most rated reports were located in the Good and Very good categories because majority of the reports explicitly addressed the evaluation purposes, questions, and

activities. While 30% of the evaluations had Fair ratings, only one evaluation was rated Poor.

Table 9

Checkpoint Results for A1: Justified Conclusions and Decisions

A1	Justified Conclusions and Decisions	Percent Met Evaluations
1	Address each contracted evaluation question based on information that is sufficiently broad, deep, reliable, contextually relevant, culturally sensitive, and valid	98%
2	Derive defensible conclusions that respond to the evaluation's stated purposes, e.g., to identify and assess the program's strengths and weaknesses, main effects and side effects, and worth and merit	93%
3	Limit conclusions to the applicable time periods, contexts, purposes, and activities	94%
4	Identify the persons who determined the evaluation's conclusions, e.g., the evaluator using the obtained information plus inputs from a broad range of stakeholders	11%
5	Identify and report all important assumptions, the interpretive frameworks and values employed to derive the conclusions, and any appropriate caveats	65%
6	Report plausible alternative explanations of the findings and explain why rival explanations were rejected	15%

Table 10

Rating Scale Pertinent to A1: Justified Conclusions and Decisions

Rating Scale	Evaluation Reports
Excellent	0
Very Good	9
Good	28
Fair	16
Poor	1
Total	54

Table 11 shows the results related to A2: Valid Information. This accuracy sub-standard states, "Evaluation information should serve the intended purposes and support

valid interpretations” (Yarbrough et al., 2011, p. 171). Table 11 shows that the majority of evaluations ($n = 41$) addressed well the sixth checkpoint. The evaluation reports were able to document the information and procedures to answer the stated questions and purposes of the evaluation. Only 11 reports addressed the checkpoints in items 1 through 3. None of the evaluations addressed the statement in item 4. There was insufficient information to explicitly indicate how stakeholder groups assisted with the evaluation activities.

Table 11

Checkpoint Results for A2: Valid Information

A2	Valid Information	Percent Met Evaluations
1	Through communication with the full range of stakeholders develop a coherent, widely understood set of concepts and terms needed to assess and judge the program within its cultural context	6%
2	Ensure—through such means as systematic protocols, training, and calibration--that data collectors competently obtain the needed data	9%
3	Document the methodological steps taken to protect validity during data selection, collection, storage, and analysis	6%
4	Involve clients, sponsors, and other stakeholders sufficiently to ensure that the scope and depth of interpretations are aligned with their needs and widely understood	0%
5	Investigate and report threats to validity, e.g., by examining and reporting on the merits of alternative explanations	31%
6	Assess and report the comprehensiveness, quality, and clarity of the information provided by the procedures as a set in relation to the information needed to address the evaluation’s purposes and questions	76%

On the rating scale in Table 12, the overall ratings of the A2: Valid Information accuracy sub-standards were low. None of the evaluations had ratings of Excellent or Very Good. Only one report was rated Good, 33% were rated Fair, and the remaining reports (64%) had ratings of Poor.

Table 12

Rating Scale Pertinent to A2: Valid Information

Rating Scale	Evaluation Reports
Excellent	0
Very Good	0
Good	1
Fair	18
Poor	35
Total	54

Table 13 shows the results related to A3: Reliable Information. This sub-standard states, “Evaluation procedures should yield sufficiently dependable and consistent information for the intended uses” (Yarbrough et al., 2011, p. 179). Table 13 shows that more than half of the reports ($n = 31$) met the first checkpoint. Given that most evaluations used pre and post designs, these reports included different types of reliability. Due to insufficient information about those inspecting the reliability, the lowest number of reports ($n = 2$) addressed the criteria in item 3. Approximately, ten to 15 reports addressed items 2, 4, 5, and 6.

On the rating scale in Table 14, the overall ratings of this sub-standard of accuracy are low. None of the evaluations had ratings of Excellent or Very Good. Only (11%) of reports were rated Good. While there were 35% studies in the Fair category, the remaining 53% of studies had Poor ratings.

Table 13

Checkpoint Results for A3: Reliable Information

A3	Reliable Information	Percent Met Evaluations
1	Determine, justify, and report the needed types of reliability—e/g., test-retest, findings from parallel groups, or ratings by multiple observers—and the acceptable levels of reliability	57%
2	In the process of examining, strengthening, and reporting reliability, account for situations where assessments are or may be differentially reliable due to varying characteristics of persons and groups in the evaluation's context	20%
3	Ensure that the evaluation team includes or has access to expertise needed to investigate the applicable types of reliability	4%
4	Describe the procedures used to achieve consistency	22%
5	Provide appropriate reliability estimates for key information summaries, including descriptions of programs, program components, contexts, and outcomes	28%
6	Examine and discuss the consistency of scoring, categorization, and coding and between different sets of information, e.g., assessments by different observers	19%

Table 14

Rating Scale Pertinent to A3: Reliable Information

Rating Scale	Evaluation Reports
Excellent	0
Very Good	0
Good	6
Fair	19
Poor	29
Total	54

Table 15 displays the results related to A4: Explicit Program and Context Descriptions. This sub-standard states, “Evaluations should document programs and their contexts with appropriate detail and scope for the evaluation purposes” (Yarbrough et al., 2011, p. 185). In Table 15, more than half of the studies ($n = 51$) confirmed the statement on items 1. Similarly, about half of the studies ($n = 48$) confirmed the statement on item 3. There were 26 and 27 reports met the statement on items 2 and 5 respectively. However, only seven reports confirmed the statement on item 6. Regarding item 6, most evaluated programs focused on the context under study, and no indication on how such context could be applied to different contexts.

Table 15

Checkpoint Results for A4: Explicit Program and Context Descriptions

A4	Explicit Program and Context Descriptions	Percent Met Evaluations
1	Describe all important aspects of the program—e.g., goals, design, intended and actual recipients, components and subcomponents, staff and resources, procedures, and activities—and how these evolved over time	94%
2	Describe how people in the program’s general area experienced and perceived the program’s existence, importance, and quality	48%
3	Identify any model or theory that program staff invoked to structure and carry out the program	89%
4	Define, analyze, and characterize contextual influences that appeared to significantly influence the program and that might be of interest to potential adopters, including the context’s technical, social, political, organizational, and economic features	56%
5	Identify any other programs, projects, or factors in the context that may affect the evaluated program’s operations and accomplishments	50%
6	As appropriate, report how the program’s context is similar to or different from contexts where the program is expected to or reasonably might be adopted	13%

On the rating scale in Table 16, the overall ratings of the A4: Explicit Program and Context Descriptions sub-standard were high. Unlike the other accuracy standards,

this sub-standard had two studies that rated Excellent on the scale. Nearly, half of the reports (48%) were rated as Good and higher. Also, half of the reports (48%) were rated as Fair, with only two reports rated as Poor.

Table 16

Rating Scale Pertinent to A4: Explicit Program and Context Descriptions

Rating Scale	Evaluation Reports
Excellent	2
Very Good	8
Good	16
Fair	26
Poor	2
Total	54

Table 17 shows the results related to A5: Information Management. This accuracy sub-standard states, “Evaluations should employ systematic information collection, review, verification, and storage method” (Yarbrough et al., 2011, p. 193). Table 17 shows that first item was addressed in most reports ($n = 51$). Item 1 focuses on the methods used in the report to ensure the accuracy. Also, well documentation was found in more than half of the reports ($n = 31$). However, all reports did not meet the statements in items 3, 4, 5, and 6. Tasks included on those items were difficult to obtain because the reports reviewed were the only available documents for the study.

Table 17

Checkpoint Results for A5: Information Management

A5	Information Management	Percent Met Evaluations
1	Select information sources and procedures that are most likely to meet the evaluation’s needs for accuracy and be respected by the evaluation’s client group	94%

Table 17—Continued

A5 Information Management		Percent Met Evaluations
2	Ensure that the collection of information is systematic, replicable, adequately free of mistakes, and well documented	57%
3	Establish and implement protocols for quality control of the collection, validation, storage, and retrieval of evaluation information	0%
4	Document and maintain both the original and processed versions of obtained information	0%
5	Retain the original and analyzed forms of information as long as authorized users need it	0%
6	Store the evaluative information in ways that prevent direct and indirect alterations, distortions, destruction, or decay	0%

Table 18

Rating Scale Pertinent to A5: Information Management

Rating Scale	Evaluation Reports
Excellent	0
Very Good	0
Good	0
Fair	30
Poor	24
Total	54

Table 19 shows the results related to A6: Sound Designs and Analyses. This sub-standard states, “Evaluations should employ technically adequate designs and analyses that are appropriate for the evaluation purposes” (Yarbrough et al., 2011, p. 201). Table 19 displays the conformity of the majority of evaluations to items 1, 2, and 4. More than 40 evaluations addressed the statements in each of those three items that focused on the logical framework, data collection methods, and final analysis. Half of the evaluations ($n = 27$) met item 3 related to the procedural aspects of the report. The information needed

to assess item 5 regarding funding, scheduling, and metaevaluation were absent. Only 15 evaluations satisfied item 6 of this standard.

On the rating scale in Table 20, the overall ratings of the A6: Sound Designs and Analyses sub-standard were moderately high. Despite the absence of reports in the Excellent category, (46%) reports had ratings of Good and Very Good. The Fair category contained 44% of evaluations, while the remaining 9% of the reports had a Poor rating.

Table 19

Checkpoint Results for A6: Sound Designs and Analyses

A6	Sound Designs and Analyses	Percent Met Evaluations
1	Create or select a logical framework that provides a sound basis for studying the subject program, answering the evaluation's questions, and judging the program and its components	87%
2	Plan to access pertinent information sources and to collect a sufficient breadth and depth of relevant, high quality quantitative and qualitative information in order to answer the evaluation's questions and judge the program's value	80%
3	Delineate the many specific details required to collect, analyze, and report the needed information	50%
4	Develop specific plans for analyzing obtained information, including clarifying needed assumptions, checking and correcting data and information, aggregating data, and checking for statistical significance of observed changes or differences in program recipients' performance	78%
5	Buttress the conceptual framework and technical evaluation design with concrete plans for staffing, funding, scheduling, documenting, and metaevaluating the evaluation work	0%
6	Plan specific procedures to avert and check for threats to reaching defensible conclusions, including analysis of factors of contextual complexity, examination of the sufficiency and validity of obtained information, checking on the plausibility of assumptions underlying the evaluation design, and assessment of the plausibility of alternative interpretations and conclusions	28%

Table 20

Rating Scale Pertinent to A6: Sound Designs and Analyses

Rating Scale	Evaluation Reports
Excellent	0
Very Good	8
Good	17
Fair	24
Poor	5
Total	54

Table 21 shows the results related to A7: Explicit Evaluation Reasoning. This sub-standard states, “Evaluation reasoning leading from information and analyses to findings, interpretations, conclusions, and judgments should be clearly and completely documented” (Yarbrough et al., 2011, p. 209). Most reports ($n = 50$) adequately corresponded to item 2. Item 6 was concerned with assessment of factors influenced the reports’ conclusions and was confirmed by 40 reports. However, a small number of reports were located in items 4 and 5. The fourth checkpoint was the least item addressed because the reports did not sufficiently document the stakeholder group included in the studies.

Table 21

Checkpoint Results for A7: Explicit Evaluation Reasoning

A7	Explicit Evaluation Reasoning	Percent Met Evaluations
1	Clearly describe all the assumptions, criteria, and evidence that provided the basis for judgments and conclusions	63%
2	In making reasoning explicit, begin with the most important questions, then, as feasible, address all other key questions, e.g., those related to description, improvement, causal attributions, accountability, and costs related to effectiveness or benefits	93%

Table 21—Continued

A7 Explicit Evaluation Reasoning	Percent Met Evaluations
3 Document the evaluation’s chain of reasoning, including the values invoked so that stakeholders who might embrace different values can assess the evaluation’s judgments and conclusions	65%
4 Examine and report how the evaluation’s judgments and conclusions are or are not consistent with the possibly varying value orientations and positions of different stakeholders	2%
5 Identify, evaluate, and report the relative defensibility of alternative conclusions that might have been reached based on the obtained evidence	11%
6 Assess and acknowledge limitations of the reasoning that led to the evaluation’s judgments and conclusions	74%

On the rating scale in Table 22, the overall ratings of sub-standard A7: Explicit Evaluation Reasoning was moderately high. Despite the absence of evaluations on the Excellent and Very Good categories, (35%) of the reports had ratings of Good. The Fair category contained (61%) of the evaluations, whereas only (4%) was included in the Poor category.

Table 22

Rating Scale Pertinent to A7: Explicit Evaluation Reasoning

Rating Scale	Evaluation Reports
Excellent	0
Very Good	0
Good	19
Fair	33
Poor	2
Total	54

Table 23 shows the results related to A8: Communicating and Reporting. This sub-standard states, “Evaluation communications should have adequate scope and guard against misconceptions, biases, distortions, and errors” (Yarbrough et al., 2011, p. 217). It was difficult to obtain the information needed to meet this sub-standard in the reports. All evaluations were in the form of printed reports. For this reason, all evaluations were considered to meet the statement in item 4. Because of the absence of reporting methods, all evaluations had a rating of Poor, as shown in Table 24.

Table 23

Checkpoint Results for A8: Communicating and Reporting

A8	Communicating and Reporting	Percent Met Evaluations
1	Reach a formal agreement that the evaluator will retain editorial authority over reports	0%
2	Reach a formal agreement defining right-to-know audiences and guaranteeing appropriate levels of openness and transparency in releasing and disseminating evaluation findings	0%
3	Schedule formal and informal reporting in consideration of user needs, including follow-up assistance for applying findings	0%
4	Employ multiple reporting mechanisms, e.g., slides, dramatizations, photographs, PowerPoint©, focus groups, printed reports, oral presentations, telephone conversations, and memos	100%
5	Provide safeguards, such as stakeholder reviews of draft reports and translations into language of users, to assure that formal evaluation reports are correct, relevant, and understood by representatives of all segments of the evaluation’s audience	0%
6	Consistently check and correct draft reports to assure they are impartial, objective, free from bias, responsive to contracted evaluation questions, accurate, free of ambiguity, understood by key stakeholders, and edited for clarity.	0%

Table 24

Rating Scale Pertinent to A8: Communicating and Reporting

Rating Scale	Evaluation Reports
Excellent	0
Very Good	0
Good	0
Fair	0
Poor	54
Total	54

Research Question 2: Interviews with the Evaluators

This section addresses the second research question. The aim of the research question was to understand the perceptions and experiences related to the status and accuracy of program evaluation in the school counseling context. Specifically, three sub-questions were formed to explain and clarify the phenomenon of the study. These sub-questions focused on the strengths and weaknesses of program evaluation, methodology selection and alignment, and other key factors that may influence program evaluation practice.

Methodology in analyzing the interviews is reported in the method section (see pp. 54-55). The Qualitative Data Analysis software-ATLAS.ti 7 was used to assist in managing and analyzing the data more easily and efficiently. As a result of the process used, related codes were grouped to form eight major themes, along with five sub-themes. These themes were aligned to research question two as shown in Table 25.

Table 25

Identified Themes and Sub-themes

Research Question Domains	Themes	Sub-themes
Strengths	<ul style="list-style-type: none"> • Increased Awareness of Program Evaluation 	<ul style="list-style-type: none"> • Perceptions of Program Evaluation Roles • Evaluation Experience • Evaluation Activities and Supports
Weaknesses	<ul style="list-style-type: none"> • Financial Assistance • Competency • Leadership Supports • Resource Allocation 	
Methodology Orientation	<ul style="list-style-type: none"> • Common Methodologies 	<ul style="list-style-type: none"> • Methodological Barriers
Key Influential Factors	<ul style="list-style-type: none"> • Pre-Service Preparations • In-Service Preparations 	<ul style="list-style-type: none"> • Counselor Education Programs

The following sections provide a description of the participants and the findings related to the identified themes and sub-themes

Identified Themes and Sub-themes

Eight themes emerged as a result of the coding analysis process. Those themes were then organized to answer research question 2: *What are the perceptions and experiences of the participating evaluators regarding the status and accuracy of school counseling program evaluation?* Some relevant sub-themes were also derived from the major themes. The second research question has three sub-questions. First, sub-question one focused on the strengths and weaknesses of program evaluation in school counseling field. Two major themes speak to the strengths of program evaluation, with three sub-themes, and three relevant themes speak to the weaknesses of program evaluation, with

one sub-theme. Sub-question two centered on the participants' perceptions about methodology orientation when evaluating school counseling programs. There was one major theme along with one sub-theme related to the methodology question.

Finally, the intention of sub-question three was to understand factors that may have an impact on program evaluation practice in the field. Two themes emerged that shed light on the influential factors in this sub-question. The themes and a sub-theme for each sub-question in research question 2 are discussed in detail in the paragraphs below.

Strengths of Program Evaluations

The following two themes answer the second research question, sub-question 1 pertinent only to **strengths**: *What do evaluators perceive as the strengths and weaknesses of school counseling program evaluation?*

Increased Awareness of Program Evaluation

This theme focused on uncovering the participants' understanding of school counseling program evaluation. Participants identified some areas that displayed their awareness of school counseling program evaluation. All participants were responsive and attentive to show how program evaluation is an essential part of the school counseling. They pointed out the internal and external forces that play roles in strengthening the evaluation process within the field. Therefore, three sub-themes were developed during the analytical process and were relevant to the theme. These sub-themes encompassed (a) perceptions of program evaluation roles, (b) participants' evaluation experiences in the field, and (c) instances of evaluation activities and supports.

Perceptions of program evaluation roles. Three participants agreed with the various benefits and purposes of program evaluation, which were aimed to strengthen the

school counseling programs. Two participants also discussed the ultimate goal of program evaluation in school counseling. One participant talked about the objective of program evaluation stating, “It was important to conduct comprehensive evaluation of the program in order to determine if it was functioning effectively in the interests of children.” Another participant described the purpose of conducting program evaluation in school noting, “The benefits of evaluation are to see if what you are doing is having a positive effect, and what you need to change it or come up with a different strategy.” She added, “You could be doing something, and you think it is great. But, if you are not seeing any results from your data and criteria that you are using, then you are wasting your time.” These participants emphasized the significance of evaluation to meet the student needs. Another participant specified areas where program evaluation may take place. He said, “Intervention evaluation, and in particular, classroom-based interventions, are the area [*sic*] that I’ve seen most of the program evaluation happened.”

Two participants described the contributions of the ASCA National Model in encouraging the engagement of evaluation activities in school counseling programs. One counselor educator stated, “With the implementation of the ASCA National Model over 10 years ago or 12 years ago now, there was a lot of emphasis on using data, which often times involves engaging program evaluations.” Another participant also indicated how she benefited from the training and support of the ASCA National Model. Thus, as evident from the participant responses, the ASCA National Model has been useful in supporting program evaluation.

Three participants pointed out the roles and responsibilities of program evaluation to enhance progress in the field. One participant stated, “Program evaluation is not

something that's extraordinarily popular, but it has significant implications for the future of our profession." This response spoke to what another participant said in this regard. He specified, "There are a number of bigger competent people who are interested in this topic." Moreover, these responses aligned well with the explanation provided by the third participant. He described school counseling programs as one component of the entire school program. He further noted, "The environment is very supportive of program delivery."

There were three participants who were concerned about the improvements needed to enhance the practice of program evaluation. Two participants addressed their concerns about effective ways of practicing evaluation. One of these participants hoped to see more improvement of the practice stating "I think it should be handled better than it is." He continued, "I haven't given up on it. I think it is still possible to improve." Similarly, the other participant responded to the same concern saying, "I think we are starting to see more improvement."

Some of the participants described how they view evaluation in school counseling programs somewhat differently. The school counselor noted, "When it comes to counselors evaluating what they are doing, we are going more into that as a profession." She further explained, "So you are always trying to strive to do better. The evaluation piece is a way to continue to address issues and improve our intervention" One counselor educator agreed with the view of the school counselor saying, "Properly trained school counselors are in a good position to engage in some effective program evaluation if they work collaboratively as a team." Conversely, another counselor educator explained the motivations that drive the school counselors-evaluators to conduct program evaluation in

schools. He stressed that counselors are motivated to conduct evaluations in schools due to the feeling of job insecurity. Thus, this feeling of job insecurity leads school counselors to perform evaluation activities in order to prove their position, not to improve programs.

Two participants mentioned the benefits of collaboration during the evaluation process. One interview participant spoke to the benefits school counselors may have when working together with the school stakeholders. He said, “the students, counselors, and sometimes administrators collaboratively working together to figure out what they want to look at.” The school counselor had the same thought of collaboration with other staffs and teachers when evaluating her program. She revealed, “So the teachers together, we collaborate on procedures, and then we present it to the entire school.”

Evaluation experience. The interviews revealed the evaluation experience of the participants. Most of the participants shared common evaluation experience, except one participant who had a different experience working as a school counselor as opposed to a counselor educator. The evaluation experiences were classified into two categories. The first category consisted of evaluation experience that was evident by the projects they talked about in the interviews. The second category consisted of the evaluation experience participants gained in teaching program evaluation within counseling programs at universities and other related institutions.

In regard to participants’ evaluation experience, one participant was involved in evaluation projects that ranged from big-scale to small-scale projects in small and large districts. He also worked on federal and state grant evaluations. Another participant also stressed his evaluation working experience, stating, “I have done several program evaluations, most of the time with the students that [*sic*] are interns.” Yet another

participant, a counselor educator, revealed his evaluation experiences when he was a school counselor, indicating he was able to conduct school counseling program evaluations. Later in his career as a counselor educator, he became an evaluator focusing on school counseling research and evaluation, especially evaluation capacity building.

The participants also had a teaching experiences related to program evaluation. Three participants engaged in teaching about evaluation to graduate students. They specified that program evaluation was introduced to students during their research method courses. One counselor educator identified his teaching role stating the following:

I have also been involved in educating youth counselors or perspective counselors in evaluation methods in order to help them evaluate their work when they are in schools, and it involved efforts consulting with counselors in schools to help them evaluate specific interventions.

Evaluation activities and supports. Finally, in regard to strengths, participants also talked about some of the best examples of where program evaluation took place. Two participants identified Utah and Missouri as examples of states that devoted considerable attention to evaluation. One counselor educator stressed the role of Utah state's initiatives noting, "There are some places that are really bright like in Utah. There is a whole state which is using evaluation very, very effectively" Participants also articulated the role of these states in allocating resources and other outstanding efforts for school counseling and evaluation of counseling programs. Moreover, they added that evaluation was one of the means to financially support the counseling programs in Utah. One participant responded, "Those evaluations are actually used to invest millions of dollars in school counseling." He continued stating, "Increasingly school counseling is

funded over other things because of the robust evaluation data that is present within that state.” This response spoke to the capacity of evaluation to determine a counseling program’s effectiveness.

Financial Assistance

The second theme under Strengths was Financial Assistance. The participants’ responses regarding funding sources of program evaluation varied. During the coding process, the sources of financing were grouped into external and local funds. Federal and state funds were classified as external funding. District and school funds were placed in the local category. Two participants discussed external funds from federal and state grants. One participant indicated that federal funds were given to schools where counseling programs were implemented. The participant provided an example of federal funds such as the Elementary and Secondary Education Act. Regarding state funds, one participant noted, “The State Department of Education provides support for schools to do program evaluation for school counseling.” These responses indicated the roles federal and state funds have in enhancing program evaluation activities in schools.

At the local funding level, districts seemed to have major financial influence over the support of program evaluations. One participant spoke of district control over federal funds noting, “The contract is going to be held by the school district again even if those funds were federally generated.” Additionally, two participants agreed that large districts tended to provide funds for particularly large program evaluation initiatives. A counselor educator noted, “If the district is going to do a full-blown program evaluation, typically a district comes up with some funding for that.” The school counselor said, “I work for a big school district, so we have a lot of support. They provide a lot of resources for us.”

On the other hand, the local school funding level did not appear to be powerful in providing financial assistance for internal evaluation activities. Two participants elaborated on the role of school funding. One participant noted, “For smaller program evaluation initiatives...the evaluations are really something that is the cost of the evaluation both in terms of time and in terms of material”

Overall, this theme was considered a strength for program evaluations because financial allocations are needed to perform evaluations. Participants described the role of federal, state, and district funds allocated for school counseling programs that have program evaluation as a component of their structure. One participant noted that districts had to provide an evidence of program implementation in order to secure the state funds. He continued, “But with those funds you are required to evaluate the program and to share your evaluation results.” Similarly, another participant identified evaluation responsibilities when conducting evaluations for federal and district funds. He spoke about his roles in grant-funded evaluation noting, “I was the principal investigator, which would be the person in charge of designing, conducting, and supervising the evaluation activities.” He further specified his role with district evaluation funds noting, “I am essentially the main evaluator. I was responsible for instrument development, district selection, evaluation design, the analysis, and write-ups.”

Weaknesses of Program Evaluations

The following three themes answer the second research question, sub-question 1 pertinent only to **weaknesses**: *What do evaluators perceive as the strengths and weaknesses of school counseling program evaluation?*

Competence

One of the weaknesses participants identified about program evaluations in school counseling was the evaluator's competence. Interviews revealed a consensus among participants that evaluation competency in the field overall was not at a satisfactory level. Two of the four participants indicated the importance of acquiring skills specific to the context of the school counseling. One of these participants noted, "We have seen that there have been evaluations that have been conducted externally by evaluators that do not have a background in school counseling." The other participant pointed out the difficulty of evaluating school counseling programs without prior "working knowledge." He further explained that counseling programs include certain structural elements. Such elements as, "the history of the field, purposes, and the interventions," would assist in understanding the structure of the programs when evaluating them.

Three participants also expressed their concerns about the knowledge acquired by evaluators in the field. Although the field has been moving to promoting program evaluation, one participant indicated evaluators' lack of knowledge and skills are factors affecting the practice. Similarly, another participant stated, "We really do need to maintain some level of proficiency on program evaluation." In defining this proficiency, another participant explicated the skills required for conducting an effective evaluation. Specifically, he spoke to the personal and social skills of the evaluators. He stressed the role of "interpersonal skills and interpersonal insight because the process is here only a social process." He also stressed the technical skills needed to enhance the use of the evaluations saying, "You need to be a good writer and a good oral communicator because if the results are not literally expressed, they are not going to be used."

Finally, participants had several ideas about factors that may be influencing program evaluation competence in school counseling. A factor one participant mentioned is the fear of evaluation. He noted, “People are actually fearful of evaluation. They do not feel like they have been trained well enough. They feel like if they actually evaluate what is really happening, then there could be negative ramifications of that.” One participant expressed his feeling that these competency issues are barriers to the field’s advancement regarding program evaluation implementation. He attributed this issue to the isolation of the school counseling field from the evaluation profession stating, “People who are involved in evaluation school counseling do not track the advances in theory or practice and the general feel of the evaluation, education evaluation.” A counselor educator also mentioned this gap noting, “There is no connection with say the evaluation profession like American Evaluation Association, almost none.”

Leadership Support

The participants described the connections between leadership support and the success of program evaluation as another weakness in school counseling. Two participants called for more leadership engagement to aid program evaluation activities. On the school level, one participant noted, “The leadership has to be involved in this because if you do not have the leadership of the principal, then you are not going to get the support you need.” She explained the role of counselor for programs’ improvement stating, “If they want their school to improve effectively, the counselor can provide a lot of things that the principal can not do or is not aware of.” At the district level, a counselor educator talked about the shortage of school counseling leaders noting, “We have very few formal school counseling leaders within the system.” He then talked about the

change of counselor' role stating, "We've also seen district level positions where they used to be in charge of school counseling programs, and now they are wearing many hats."

Regarding leadership aids, three participants agreed that program evaluation is an optional component of school counseling programs. However, one participant stressed the necessity of evaluation practice to both the school counselor as an evaluator and the leadership of the school or the district. Another participant addressed the position and support of the stakeholder in the cycle of the evaluation. The counselor educator suggested the need to increase their involvement more noting, "I think [stakeholders' engagement] is good practice, but it is not common practice within school counseling based evaluation." He explained how ASCA expects stakeholder involvement stating, "*The ASCA National Model* suggests that school counseling programs have advisory part of the stakeholders, and stakeholder have an input into everything including program evaluation and that they receive results of any evaluation activities and help guide the program." He added, "Stakeholder group is a thing that almost never gets implemented in schools and so the mechanisms from that model uses to ensure that stakeholders have input and evaluation really does not." Thus, this participant brought to light the importance of support inside the school noting, "Program evaluation works best when it is inspired from within and supported from within."

Resource Allocation

This theme is related to the necessary resources allocated to conduct program evaluation. The participants talked about the most needed resources, which included time and money. Almost all participants agreed that the biggest challenge facing the

implementation of program evaluation in schools was time. One participant stated, “We do not really create a lot of time to do effective program evaluation at the school, the districts, or the state level.” Likewise, another participant noted, “Time and resources are not allocated in schools to do way better evaluation.” Related to time was money. Despite the funding sources reported by the participants as a strength of program evaluations, the necessary funds distributed to do program evaluation are short. Three participants agreed that insufficient funds were poorly distributed for program evaluation activities. One participant noted, “There is not enough material resources to do a good job, and rarely is there other resources for them to consult with a professional evaluator needed to design implementation analysis of evaluation.” These responses shed light into the funding channels that neglect program evaluation and its role in program success.

Methodology Orientation

The following theme answers the second research question, sub-question 2: *How do evaluators align their evaluations to existing evaluation methods and practice?*

Common Methodologies

Participants were in consensus about the popular methodologies used to evaluate school counseling programs. Three participants identified pre-post designs as one of the most common designs employed by the evaluators in the field. Also, the four participants agreed on the use of quantitative and qualitative methods in those evaluations. Only one participant indicated the use of mixed method evaluations noting, “I think you are looking at kind of some mixed methods evaluation activities.”

All asserted that survey was the most preferred method by the evaluators. Two participants revealed the common use of qualitative methods, such as interviews and

focus group. On the other hand, the school counselor identified rubrics and feedback as evaluation tools used in the counseling program. She noted, “Asking for feedback all the time is a form of our evaluation.” She also mentioned, “We have rubrics in place so it is kind of an evaluation tool.”

Furthermore, one counselor educator addressed the role of the field for developing evaluation tools for evaluators. He identified two models used to guide and assist evaluators in evaluating their programs. One of them is the Carol Dahir’s (2003) model. He explained the model saying, “She has a model of doing evaluation work that is basically a blend of databases.” The other assistant model to conduct evaluations is the RAMP model (Young & Kaffenberger, 2011), as indicated by the participant. Another counselor educator articulated the importance of using the evaluation approaches and models developed in the evaluation field. Specifically, he stressed the need for more participatory approaches due to the nature of most counseling programs. He noted, “I see investing in a participatory evaluation process is probably the better way to approach this work within at least all levels.” He was also an advocate for evaluation capacity building in school counseling field. He spoke to this aspect of evaluation saying, “You have to actually invest in developing capacity at all those different levels if you can even expect to get high quality evaluation products.” He further stated, “So I think that without doing any of that, you are going to run into some pretty significant issues.”

Methodological Barriers

Two participants revealed their concerns about the methodologies currently employed in the field by the evaluators. These two participants called for using more rigorous methodologies to evaluate counseling activities. One participant stated, “Your

methodologies really need to reflect that you are putting an outsiders' view of the program and then you are presenting that to all of the stakeholders involved." Thus, he pointed out that methodology "does not need to be that sophisticated," but strongly insisted on including "some professional methodologies" when conducting program evaluations. For instance, he described the methods used in the Utah evaluations noting, "Yet, when you look at the methodologies, they are mostly survey data and just to multiple stakeholders groups and they synthesize that and present it."

Similarly, the other participant added to the methodology concerns stating, "they are very simple and no power methods." He attributed these issues of methodology to the pre-service preparation, stating, "They might understand research methods and statistics, but they leave not being able to apply that knowledge to evaluation." To this end, the counselor educator said regarding methodology, "very few that actually approached more systematically and actually collected real data from practitioners."

Key Factors Influencing Program Evaluation Practice

The following two themes answer the second research question, sub-question 3: *What are the major factors influencing program evaluation practice in school counseling context?* All participants talked about factors that have an essential role in evaluation practice. These factors were grouped into two categories. The first category was associated with pre-service preparation. The second category concerned participants' perceptions of in-service preparations. Some of the participants emphasized the school counselor's evaluation role as an internal evaluator. Other participants stated that external evaluators should conduct the evaluation work. The following presents the participants views regarding the preparation of school counseling on program evaluation.

Pre-service Preparation

Three participants discussed elements of evaluation preparation for school counselors before working in schools. Specifically, they talked about certification requirements needed to ensure the acquisition of necessary knowledge and skills. Three participants also stressed the state's role in requesting the necessary evaluation skills and knowledge for evaluation work. One participant explained this issue, noting, "Program evaluation could be facilitated by states if their licensing standards gave more attention to program evaluation knowledge and skills." Another participant stated, "Our crediting bodies do not really have a good handle on the nature of learning that needs to happen in order for a counselor to be competent to do an evaluation." The other participant noted, "There is no license or standard that school counseling graduate students need to know about program evaluation."

Three counselor educators discussed teaching program evaluation to prospective school counselors. One participant noted, "My students that [sic] have gone through my courses and have some skills in conducting program evaluation." Another participant also involved in teaching program evaluation to his students expanded on that notion, stating, "It has been about nine or 10 years working with students on conducting program evaluation in schools." Additionally, program evaluation was suggested to be a prerequisite for the graduate programs, especially at the doctorate level. One participant talked about this involvement, saying, "It is necessary for particularly the counselor educator at the doctorate level to become better educated in evaluation theory and practice."

Two participants spoke about issues involved in counselor education related to teaching evaluation. One of the participants noted, “If we do not treat our master’s level work force in an appropriate and actual more dynamic evaluation practice, then we are going to even have trouble in implementing the national model.” The other counselor educator spoke to the nature of such issues stating, “There are challenges around what is appropriate curriculum and competencies that need to be learned.”

Two participants, however, suggested where those evaluation courses should take place. Both agreed that evaluation learning should be part of the research courses. One participant noted, “The best place we have to engender program evaluation skills in most graduate programs is in the context of the research courses.” The other participant similarly added, “We do a real disservice if we do not explicitly include school counseling program evaluations within our research methodology course.” Thus, training school counselors in evaluation skills was the most emphasized concept in the interviews. Participants revealed that training does not have considerable attention in the field, which is necessary to successfully evaluate counseling programs.

In-service Preparation

The participants recognized the necessity for providing on-going professional development for schools counselors in evaluation. Three participants spoke about the role of professional development in extending the evaluation knowledge gained in counselor education. Two participants identified sources of professional development provided to counselors in program evaluation. This professional development is provided in the form of conferences, district-based professional development training, and webinars. One participant stated, “ I see professional development that's provided in program evaluation

most often comes in the context of conferences.”

Although professional development has been provided in the field, there were issues addressed by three participants. One participant noted, “The trainings that I have actually seen have been superficial surface level and really have not provided the training that is necessary to do program evaluation.” Similarly, in discussing the conferences offered in the field for the counselors, two participants discussed the problems with this form of professional development, stating, “the conferences...only pick up a small percentage of counselors that need the skills.” Another participant related to this issue, stating, “takes away their opportunities to gain more knowledge through professional development in the area of program evaluation.” Furthermore, this participant attributed this issue to the absence of the leadership assistance stating, “School counselors are really left to their own devices to determine what they need to do for professional development.”

Thus, participants were attentive to the importance of training counselors in program evaluation. They ascribed the success of school counseling program evaluation to organized plans and efforts to prepare those counselors. They pointed out the means for such training through carefully designed counselor education programs and professional development. Because of the possible of limited funds allocated for evaluation activities, the participants believed counselors are the best professionals to obtain the necessary knowledge and skills to fill this gap in education.

Discussion

The aim of this metaevaluation study was to examine the conformity of school counseling program evaluations to the JCSEE accuracy standards. Also, school

counseling professionals were engaged in semi-structured interviews to understand the position and accuracy of program evaluation in the field. The methodology used was designed to understand the nature of and the connection between program evaluation and school counseling programs from different perspectives. School counseling programs are essential structures of the general education system (Kozlowski & Huss, 2013). Such programs and interventions are useful to assist the development and success of students in schools nationwide. Thus, attention to the effectiveness, value, and accountability of these programs is crucial to the success and performance of school counseling programs. Accordingly, meeting accountability demands and making effective use of school and student information are essential facets of effective school counseling programs (McGamon et al., 2005).

The overall metaevaluation ratings showed little conformity to JCSEE standards. However, this low performance is understood from different perspectives. Generally, evaluation is different than research. In evaluation, considerable attention is paid to many conceptual and theoretical notions such as the value and quality of the programs or the evaluand (Stufflebeam & Coryn, 2014). However, school counseling evaluations were likely designed from a research focus rather than an evaluation one. The school counseling professionals relate these concerns to the competency of those counselors expected to perform the evaluation activities. Such concerns were also mentioned in the school counseling literature. Dimmitt (2010) raised the lack of evaluation knowledge and skills as one of the issues involved in the implementation of program evaluation in the field.

Despite the reported lack of evaluation competence, knowledge, and skills, the

findings speak to the good performances revealed in the evaluation reports. These performances can be found through the high ratings in four accuracy standards. These standards were (a) A1: Justified Conclusions and Decisions, (b) A4: Explicit Program and Context Descriptions (c) A6: Sound Designs and Analyses, and (d) A7: Explicit Evaluation Reasoning. Evaluation reports that met these four standards were judged to have good evaluation practice. The reports explicitly addressed these four standards although validity and reliability standards were low. This contradiction is due to the type of information presented in the reports. As stated above, the reports were executed from a research perspective that affects the representation of information needed to meet the standards' evaluative criteria. For example, inclusion of stakeholders is one of the important aspects of the validity standards. The information about stakeholders was not enough presented in the reports which affect the rating scores of validity. Another issue is the documentation of the validity procedures in the reports (i.e., validity threats). This methodological issue was also reported by the participants who expressed their concerns about the knowledge and skills competences needed for those who evaluate the school counseling programs.

On the other hand, the financial and leadership assistance were reported in the interviews to play major roles in school counseling evaluation practice. Such assistance should increase the opportunity to train school counselors whether in pre-service or in-service training or both. Also, the leadership and top policymakers should allocate appropriate funds for program evaluation activities. Investing in training school counselors to internally assess their programs is a valuable practice. This investment may ensure the implementation of program evaluation even when there are insufficient

resources budgeted for evaluations.

The findings at the school counseling program level indicated that more evaluation reports were found in the social and personal domain. Similarly, in their study, Dahir et al. (2009) found that school counseling programs implemented by school counselors focused on social and personal development and that school counseling programs were observed more at the middle school level. In contrast, the current study documented more evaluations at the high school level than at the other two school levels. Thus, the current findings indicate that the implementation of school counseling programs along with expected evaluation activities is developing at the high school level.

CHAPTER V

CONCLUSION

The purpose of this dissertation was to examine comprehensive school counseling program evaluations. Specifically, the aim of the study was to gain an insight into (a) how program evaluation is currently practiced in the field, and (b) if these evaluations adhere to professional evaluation standards. Scholars and practitioners in the field of school counseling have questioned the capacity and practice of such program evaluations. These scholars stress that program evaluation as a method to assess program effectiveness is less commonly performed (Astramovich et al., 2005; Dimmitt, 2010; Studer et al., 2011; Trevisan, 2002). Furthermore, the school counseling literature has not clearly discussed the wide range of contemporary evaluation methods and existing tools used to conduct sound program evaluations. Therefore, the two research questions guide this study:

- 1 How well do school counseling program evaluations conform to the JCSEE accuracy standards (Yarbrough et al., 2011)?
- 2 What are the perceptions and experiences of the participating evaluators regarding the status and accuracy of school counseling program evaluation?
 - a. What do evaluators perceive as the strengths and weaknesses of school counseling program evaluation?
 - b. How do evaluators align their evaluations to existing evaluation methods and practice?
 - c. What are the major factors influencing program evaluation practice in school counseling context?

These two research questions were answered through the use of a metaevaluation approach and semi-structured interviews. Thus, the structure of this chapter contains a summary of the study's major findings, implications, limitations, contributions, and recommendations for future research.

Summary of Findings

The study was performed in two phases to reach its proposed goals. First, the Program Evaluation Metaevaluation Checklist (Stufflebeam & Coryn, 2014) was used to assess the conformity of 54 school counseling program evaluations to the Joint Committee on Standards for Educational Evaluation (JCSEE) accuracy standards (Yarbrough et al., 2011). Because of the nature of the information these program evaluations included, the accuracy standard was the only appropriate criteria for the assessment. Second, according to the results of the initial phase, four authors of program evaluations were purposefully selected and engaged in semi-structured interviews to further discover their perspectives and experiences about program evaluation practice in the field.

Accuracy Standard Ratings

The major findings from research question 1 revealed that, overall, the metaevaluation ratings derived in this study were low. The low rating is clarified through the rating analysis for each accuracy standards. The next two paragraphs show the low and high ratings for each sub-standard.

Rating decisions were determined according to where the evaluations fell within the scale of *Excellent* through *Poor*, as a final step of the analysis. A reasonable level of adherence to four of the accuracy sub-standards was manifested. These four sub-

standards were as follows: (a) A1: Justified Conclusions and Decisions, (b) A4: Explicit Program and Context Descriptions, (c) A6: Sound Designs and Analyses, and (d) A7: Explicit Evaluation Reasoning. Overall, these four criteria show that the evaluation of comprehensive school counseling programs and interventions attend to the evaluation questions, document methods, and steps that logically led to the reported conclusions. Also, the evaluation reports in this category clearly described the important features of the programs and interventions.

The results of this study showed low conformity to the criteria in the other four accuracy sub-standards. These standards were as follows: (a) A2: Valid Information, (b) A3: Reliable Information, (c) A5: Information Management, and (d) A8: Communicating and Reporting. The low amount of information included in the reports is one of reasons these accuracy standards were rated low. Specifically, such evaluations did not provide the amount of information needed to address the statements in the checklist. The standards A2 and A3 seem contradicting with the first standard A1. However, the raters treated the standard A1 with regard to the other criteria than validity and reliability of information. Although valid and reliable information are essential, other criteria such as broad, contextually relevant, and culturally sensitive information were considered which increased the ratings of this standard. Additionally, the type of information regarding management and communication were not accessible for these evaluations. Finally, the evaluation reports examined in this study failed to meet statement requiring reporting full range of reliability procedures because the pre-post designs were prevalent among evaluations. Moreover, the involvement of stakeholders in the evaluation process was not clearly stated to know their role in evaluations.

The evaluations gathered in this study may have adequate methods and procedures to collect and analyze the data. However, the types of information reported in these evaluations influenced the judgments against the program evaluation standards. Particularly, the JCSEE advises evaluators to pay attention to certain types of information that are expected to be present in an evaluation report (Yarbrough et al., 2011). Therefore, communicating and appropriately reporting the evaluation findings are very crucial steps in the evaluation course. These steps ensure that the program stakeholders understand the evaluation findings and processes, which then may enhance the likelihood to use the evaluation and inform decisions (Stufflebeam and Coryn, 2014). The attention to the evaluation stakeholders is emphasized by the JCSEE due primarily to the significance of their inputs.

The interaction between the evaluator and stakeholders assists in determining the information needed for the evaluation. Such information is likely directed by the agreed upon evaluation purposes and questions (Yarbrough et al., 2011). Depending on the report format and the nature of stakeholder needs, the types of information and report components vary. In some cases, evaluation reports are expected to provide an executive summary, descriptions of the program and context, descriptions of the methodology steps and procedures, complete presentation of findings and conclusions; and appendices containing procedures, instruments, and other relevant data and activities (Stufflebeam and Coryn, 2014). In other cases, reports simply provide information about the progress and achievement of specific evaluand. Such evaluations tend to provide information pertinent to the evaluand's "...background, structure, implementation, costs, main effects, and side effects (Stufflebeam and Coryn, 2014, p. 606). Thus, the JCSEE emphasizes that

without a sufficient details and documentation of the evaluations, judging the evaluation quality can be difficult (Yarbrough et al., 2011).

Strengths of Program Evaluations

Semi-structured interviews using a thematic analysis technique were performed to answer the second research question. Eight themes emerged, which shed light on the concerns and experiences of the participants in this study regarding different aspects of program evaluation. Specifically, the interview responses described the strengths and weaknesses of program evaluations. The strengths of program evaluation reported were the increased awareness of program evaluation and financial support present in the field. The participants highlighted the benefits and values of program evaluation in the field. Moreover, there was also an emphasis on the value of program evaluation as a means to assess the effectiveness of the implemented counseling programs (ASCA, 2012; Martin et al., 2009; Trevisan, 2002). Finally, the interviews shed light on contributions of the ASCA National Model to promote program evaluation. Particularly, the ASCA National Model (ASCA, 2003; 2012) stresses the role of program evaluation under the accountability section.

The interviews also showed other strengths of program evaluation through participants' involvement in the evaluation efforts. They illustrated their experiences with examples of their evaluative work. Participants were also able to provide examples of State's department of education that endorsed the practice of program evaluation. Particularly, they explained the role of Utah and Missouri states, which is similar to Gysbers and Henderson (2006) who also addressed the position of Utah regarding program evaluation. Participants in this study indicated that the Utah State uses program

evaluation activities to report the extent to which counseling programs meet established achievement standards, which then justify the funds for these programs. The strength of program evaluation was also evident as the participants noted the manifestation of funds allocated for program evaluation activities. They highlighted growing funding sources as the program evaluation culture is being enhanced in the field. The school counseling field endorses the evaluation as a proper method to report findings of the programs' effectiveness to stakeholders for funding and accountability purposes (Carey & Dimmitt, 2008; Gysbers & Henderson, 2006; Hatch, 2008).

Weaknesses of Program Evaluations

Three themes emerged that are associated to difficulties related to the progress and accuracy of program evaluation in school counseling. These themes were competency, leadership support, and resource allocations. The participants were attentive to the necessary competency needed to conduct program evaluation. Evaluation knowledge and skills were among the greatest concerns regarding the competency expected of school counselors as they evaluate counseling programs. Similarly, Dimmitt (2010) addressed the insufficiency of knowledge and skills of evaluation procedures as one of the evaluation challenges in the field. Additionally, the participants recognized the essential role of leadership assistance on the success of program evaluation. Such leadership involvement plays an important role to allocate resources for evaluation activities in schools or districts. Therefore, time and financial resources were the most reported challenges present in schools and districts.

Methodology

One theme was in understanding methodological orientations when conducting

program evaluations. The participants reported that social science methodologies are the mechanisms normally used to evaluate programs. Examples of designs and methods participants provided include pre-post designs, quantitative, qualitative, and mixed methods. Moreover, it was indicated that there are models that have been specifically established for evaluation purposes, such as Dahir's (2003) model for evaluation. Nevertheless, participants stressed the need for more rigorous methodology and skills to evaluate school counseling programs. In this regard, participants saw the necessity to connect with the evaluation profession to gain appropriate knowledge and skills concerning methodologies and approaches specifically developed for program evaluation. The need for evaluation skills is affirmed by Astramovich et al. (2005), who reported that 75 percent of school counselors indicated a need for more training opportunities in the area of evaluation methods.

Factors Affecting Evaluation Practice

Two themes corresponded to factors affecting evaluation practice. These themes were pre-service and in-service preparation. Participants specified the significance of preparing and training school counselors in evaluation concepts and tools. Participants spoke about the advantages of teaching prospective counselors about evaluation concepts and methodologies in their graduate education. Teaching evaluation to graduate students was evident through the teaching experiences of three participants. They recommended deep inclusion and development of evaluation courses within counselor education programs. They also stressed the need to develop certification requirements regarding evaluation knowledge and skills of school counselors. Participants perceived these requirements as criteria to promote evaluation practice more firmly in the field.

In-service training on evaluation was another factor affecting the practice of evaluation. Participants identified concerns about the nature and frequency of these professional developments when they are offered. Specifically, participants reported that professional development trainings have an insufficient focus on program evaluation. Such training concerns confirm the results of a statewide evaluation conducted to understand the evaluation training needs of 166 school counselors (Burkard et al., 2012). The authors concluded that program evaluation was the least activity among counselors. This low performance was due to the shortage of evaluation skills by those school counselors

Implications

The major implication of the study is having a better understanding of the position of program evaluation within the school counseling context. The program evaluation standards were used as away to identify the quality of existing program evaluation. Although using only one standard does not necessarily offer a complete assessment of the evaluation quality, it highlights the areas of concern for future evaluation endeavors.

Overall, the goal of a school counseling program evaluation is to assist improvement in the program and facilitate decision-making processes. Individuals in the school counseling profession also perceive program evaluation as an accountability measure. Thus, the purpose of this study was to promote metaevaluation as an evaluation-mechanism help to meet this accountability goal. Furthermore, beyond the viability of metaevaluation for accountability aims, it is also a beneficial tool to guide and develop the direction of ongoing evaluation activities.

The findings of this study also indicated that school counseling evaluation

practice should build more bridges and communicate effectively with the professional evaluation community. Evaluation knowledge, skills, methods, and standards are concepts that need further consideration in the field. Enhancing these needed competences can be achieved through an emphasis on building evaluation capacity within schools and districts where counseling programs are implemented.

In conclusion, the program evaluation standards contain the components of successful evaluation. The absence of such components decreases the credibility of the evaluation efforts, which then may lead to question the effectiveness of school counseling programs. Thus, the results of this study showed that greater attention should be given to the application of program evaluation standards in school counseling in order to increase the trustworthiness of these efforts. More than ten years ago, Trevisan (2000) recommended the field to rethink about using the program evaluation standards when evaluating school counseling programs.

Limitations

The study has limitations pertinent to the sample selection, information in the reports, and instrument. First, there were difficulties obtaining program evaluation reports from locations, such as state department of educations and evaluation agencies. Although several attempts and contacts with these locations were initiated, many of these requests were denied. These program evaluations whether funded internally or externally should be transparently available and present for public use to review the process and outcomes or to inspect any other areas of concerns. Therefore, the evaluation reports are expected be present in simple and readable structures to allow different stakeholders and researchers to understand the impact of the evaluation data on school counseling

programs and ultimately the student achievements. There might be another factor affecting the accessibility of those reports as Donaldson, Gooler, and Scriven (2002) called it “Excessive Evaluation Anxiety.” The authors identified the difficulty of obtaining the access to the evaluation information as one of the consequences of anxiety related to the evaluation work. Some evaluators or stakeholders worry about the criticism or evaluative views of others about the evaluation process or outcomes (Donaldson, Gooler, and Scriven, 2002). However, this difficulty of obtaining the evaluation reports influenced the decision to assess accessed reports on only the accuracy standard. Moreover, the number of interview participants was less than expected, although four reminders to participate were sent. This difficulty to contact authors may be due to the older dates of the reports selected for the study.

Second, the evaluation reports lacked sufficient information to address the standard statements. Since the study relied on the evaluation reports only, insufficiency of information challenged the raters to judge firmly some of the checkpoints with many criteria. Wingate (2009) asserted that there are some difficulties rating metaevaluation that relies only on evaluation reports. Specifically, some of the accuracy standards showed low ratings as a result of a lack of enough information in the reports. Contacting the authors of these reports is a recommended practice for this kind of metaevaluation study; however, the publication date of most reports was older, which prevented communication with those authors. Also, interviewing a small portion of the authors was helpful to gain insight into the reason relevant to the information issue in the reports. They highlighted the competency of evaluators related to the evaluation knowledge and skills as one of the challenges in the field. Therefore, this challenge has an impact on the

way they report the relevant evaluation information and findings.

Third, the Program Evaluation Metaevaluation Checklist (Stufflebeam & Coryn, 2014) is a well-developed instrument to facilitate the application of program evaluation standards; however, there are several issues inherent in the checklist. The checklist tends to be comprehensive and inclusive of the program evaluation standards. Therefore, the intention to include most aspects and concepts of the standards led to involve statements with many concepts. Some of the checkpoints include multiple criteria to judge information in the evaluation reports. The inclusion of multiple criteria in each checkpoint caused the raters some level of confusion during the rating process. For example, a checkpoint may require three to five different criteria to address that checkpoint. This issue led the two raters to perform many rounds and modify decision rules to check for thorough understanding of these statements. Thus, Wingate (2009) suggested deleting or developing the confusing checkpoints to increase the reliability of the checklist.

Related to the aforementioned issue is the scoring choice. As described in Chapter III, the checklist has three scoring option: addressed, not addressed, or partially addressed. The response options of not addressed and partially addressed also made the rating task more difficult. Therefore, the scoring option and the inclusion of multiple criteria in some checkpoints are practical difficulties faced during the rating process. Moreover, the checklist appears to be designed to assess an evaluation as a whole. Stufflebeam (1999) suggested the evaluation fails if certain standards are not met. Therefore, it is a challenge to decide the failure of an evaluation if only one standard is used. Finally, the use of this checklist with only evaluation reports entails collecting

additional information, which may not be feasible due to some constraints such as the date of the reports as the case of this study. It would be a practical tool when an updated version of the checklist is designed for completed reports only.

Contributions

The findings of this study shed light onto the issues and motivations surrounding program evaluation in the school counseling field. The use of a metaevaluation approach should motivate policymakers and other groups investing in program evaluation to promote this methodology to assess the quality of program evaluations. The school counseling field has been promoting evaluation as an essential method to assess the value of programs. Therefore, it is beneficial to endorse the “program evaluation standards” and other related quality standards and criteria. This endorsement can assist to guide the production of sound and credible evaluations. Also, the presence of such quality mechanisms would support school counseling’s quest to legitimize their position among other education profession and systems. Finally, scholars of the evaluation profession urge the community to increase metaevaluation efforts to clearly understand the evaluation practice (Henry & Mark, 2003). Thus, this study contributes to the body of metaevaluation literature. It helps to recognize evaluation practice in differing contexts, as the case of the current study’s context.

Future Research

Evaluation systems have clearly been established in the school counseling field. Further examination of these systems would provide a clear view of evaluation practice from different perspectives (i.e., Utah and Missouri evaluation systems related to training opportunities provided to school counselors in evaluation). Also, the current study

revolves around the insights of evaluators only. Therefore, further research should gain the views of different groups of individuals who have stakes in evaluation. Such groups may include the directors of school counseling departments, teachers, principals, and parents. Also, the abovementioned limitations related to the checklist used in the study entails further examination of the instrument validity. Finally, a revised and updated checklist to be used with only evaluation reports is recommended.

REFERENCES

- Alkin, M. (2004). *Evaluation roots: Tracing theorists' views and influences*. Thousand Oaks, CA: Sage Publications.
- American School Counselor Association (2003). *The ASCA national model: A framework for school counseling programs*. Washington, DC: Author.
- American School Counselor Association (2005). *The ASCA national model: A framework for school counseling programs* (2nd ed.). Washington, DC: Author.
- American School Counselor Association. (2012). *The ASCA national model: A framework for school counseling programs* (3rd ed.). Alexandria, VA: Author.
- Azzam, T. (2011). Evaluator characteristics and methodological choice. *American Journal of Evaluation*, 32, 376–391. doi: 10.1177/1098214011399416
- Astramovich, R. L., & Coker, J. K. (2007). Program evaluation: The accountability bridge model for counselors. *Journal of Counseling & Development*, 85, 162-172. doi: 10.1002/j.1556-6678.2007.tb00459.x
- Astramovich, R. L., Coker, J. K., & Hoskins, W. J. (2005). Training school counselors in program evaluation. *Professional School Counseling*, 9, 49-54. doi: 10.5330/prsc.9.1.80442r1q5n29x546
- Bemak, F., Williams, J., & Chung, R. (2014). Four critical domains of accountability for school counselors. *Professional School Counseling*, 18, 100–110. doi: 10.5330/prsc.18.1.q40379257k35n1kx

- Burkard, A., Gillen, M., Martinez, M., & Skytte, S. (2012). Implementation challenges and training needs for comprehensive school counseling programs in Wisconsin high schools. *Professional School Counseling, 16*, 136-145. doi: 10.5330/PSC.n.2012-16.136
- Carey, J., & Dimmitt, C. (2008). A model for evidence-based elementary school counseling: Using school data, research, and evaluation to enhance practice. *The Elementary School Journal, 108*, 422–430. doi: 10.1086/589471
- Carey, J., Morgan, M., & Oxtoby, M. (1996). Intercoder agreement in analysis of responses to open-ended interview questions: Examples from tuberculosis research. *Cultural Anthropology Methods Journal, 8*, 1-5. doi: 10.1177/1525822X960080030101
- Chen, H. (1990). *Theory-driven evaluations*. Newbury Park, CA: Sage Publications.
- Chouinard, J. (2013). The practice of evaluation in public sector contexts: A response. *American Journal of Evaluation, 34*, 266–269. doi: 10.1177/1098214013478146
- Christie, C., & Fleischer, D. (2010). Insight into evaluation practice: A content analysis of designs and methods used in evaluation studies published in North American evaluation-focused journals. *American Journal of Evaluation, 31*, 326–346. doi: 10.1177/1098214010369170
- Christie, C. (2003). What guides evaluation? A study of how evaluation practice maps onto evaluation theory. *New Directions for Evaluation, 2003*, 7–36. doi: 10.1002/ev.72

- Cooksy, L., & Caracelli, V. (2009). Metaevaluation in practice: Selection and application of criteria. *Journal of Multidisciplinary Evaluation*, 6(11), 1-15.
- Cooksy, L., & Caracelli, V. (2005). Quality, context, and use: Issues in achieving the goals of metaevaluation. *American Journal of Evaluation*, 26, 31-42. doi: 10.1177/1098214004273252
- Coryn, C., Noakes, L., Westine, C., & Schroter, D. (2011). A systematic review of theory-driven evaluation Practice from 1990 to 2009. *American Journal of Evaluation*, 32, 199–226. doi: 10.1177/1098214010389321
- Creswell, J. (2007). *Qualitative inquiry & research design: Choosing among five approaches* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Cousins, B., & Earl, L. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis*, 14, 397–418. doi: 10.3102/01623737014004397
- Cullen, A., Coryn, C., & Rugh, J. (2011). The politics and consequences of including stakeholders in international development evaluation. *American Journal of Evaluation*, 32, 345–361. doi: 10.1177/1098214010396076
- Dahir, C., Burnham, J., & Stone, C. (2009). Listen to the voices: School counselors and comprehensive school counseling programs. *Professional School Counseling*, 12, 182-192. doi:10.5330/PSC.n.2010-12.182
- Dahir, C., & Stone, C. (2003). Accountability: A M.E.A.S.U.R.E of the impact school counselors have on student achievement. *Professional School Counseling*, 6(3), 214–221.
- Davidson, J. (2005) *Evaluation methodology basics: The nuts and bolts of sound evaluation*. Thousand Oaks, CA: Sage Publications.

- Dimmitt, C. (2010). Evaluation in school counseling: Current practices and future possibilities. *Counseling Outcome Research and Evaluation, 1*, 44–56. doi: 10.1177/2150137810361306
- Dimmitt, C. (2009). Why evaluation matters: Determining effective school counseling practices. *Professional School Counseling, 12*, 395-399. doi: 10.5330/PSC.n.2010-12.395
- Donaldson, S. I., Christie, C. A, & Mark, M. M. (2008). *What counts as credible evidence in applied research and evaluation practice ?* Thousand Oaks, CA: Sage Publications.
- Donaldson, S. I., Gooler, L. E., & Scriven, M. (2002). Strategies for managing evaluation anxiety: Toward a psychology of program evaluation. *American Journal of Evaluation, 23*, 261-273. Doi: 10.1177/109821400202300303
- Duarte, D., & Hatch, T. (2014). Successful implementation of federally funded violence prevention elementary school counseling program: Results bring sustainability. *Professional School Counseling, 18*(1), 71–82.
- Eschenauer, R., & Chen-Hayes, S. F. (2005). The transformative individual school counseling model: An accountability model for urban school counselors. *Professional School Counseling, 8*(3), 244–248.
- Fetterman, D. (2001). *Foundations of empowerment evaluation*. Thousand Oaks, CA: Sage.
- Fitzpatrick, J. (2012). An introduction to context and its role in evaluation practice. *New Directions for Evaluation, 2012*, 7–24. doi: 10.1002/ev

- Fournier, D. M. (1995). Establishing evaluative conclusions: A distinction between general and working logic. *New Directions for Evaluation*, 68, 15–32. doi: 10.1002/ev.1017
- Fraenkel, J., & Wallen, N. (2006). *How to design and evaluate research in education* (6th ed.). Boston, MA: McGraw-Hill.
- Galletta, A., & Cross, W. (2013). *Mastering the semi-structured interview and beyond: From research design to analysis and publication*. New York, NY, USA: New York University Press (NYU Press). Retrieved from <http://www.ebrary.com>
- Greene, J. C. (2005). Context. In S. Mathison (Ed.), *Encyclopedia of evaluation* (pp. 82–84). Thousand Oaks, CA: Sage.
- Gysbers, N. (2001). School guidance and counseling in the 21st century: Remember the past into the future. *Professional School Counseling*, 5(2), 96-105.
- Gysbers, N., & Lapan, R. (2001). The implementation and evaluation of comprehensive school guidance programs in the United States : Progress and prospects. *International Journal for Educational and Vocational Guidance*, 1, 197–208. doi: 10.1023/A:1012297314669
- Gysbers, N., & Henderson, P. (2006). *Developing & managing your school guidance and counseling program* (4th ed.). Alexandria, VA: American Counseling Association.
- Hanssen, C., Lawrenz, F., & Dunet, D. (2008). Concurrent meta-evaluation: A critique. *American Journal of Evaluation*, 29, 572–582. doi: 10.1177/1098214008320462
- Hatch, T. (2008). Professional challenges in school counseling: Organizational, institutional and political. *Journal of School Counseling*, 6(22), 1–31.

- Hays, D. (2010). Introduction to counseling outcome research and evaluation. *Counseling Outcome Research and Evaluation, 1*, 1–5. doi: 10.15241/im.4.5.455
- Henry, G., & Mark, M. (2003). Toward an agenda for research on evaluation. *New Directions in Evaluation, 97*, 69–80. doi: 10.1002/ev.77
- Hogan, R. (2007). The historical development of program evaluation: Exploring the past and present. *Online Journal of Workforce Education & Development, 2*(4), 1-14. Retrieved from <http://opensiuc.lib.siu.edu/ojwed/vol2/iss4/5/>
- House, R. M., & Hayes, R. L. (2002). School counselors: Becoming key players in school reform. *Professional School Counseling, 5*(4), 249-256.
- Jacob, S. A., & Ferguson, S. P. (2012). Writing interview protocols and conducting interviews: Tips for students new to the field of qualitative research. *The Qualitative Report, 17*, 1-10.
- Kallemeyn, L., Hall, J., Friche, N., & McReynolds, C. (2015). Cross-continental reflections on evaluation practice: Methods, use, and valuing. *American Journal of Evaluation, 36*, 339–357. doi: 10.1177/1098214015576400
- Kozlowski, K., & Huss, S. (2013). *Training of School Counselors*. Paper presented at the 2013 American Counseling Association Conference, March 22-24, Cincinnati, OH.
- Lapan, R. T. (2012). Comprehensive school counseling programs: In some schools for some students but not in all schools for all students. *Professional School Counseling, 16*, 84-88. doi:10.5330/PSC.n.2012-16.84

- LaVelle, J., & Donaldson, S. (2010). University-based evaluation training programs in the United States 1980--2008: An empirical examination. *American Journal of Evaluation*, 31(1), 9–23. <http://doi.org/10.1177/1098214009356022>
- Leviton, L. (2014). Generative insights from the Eleanor Chelimsky forum on evaluation theory and practice. *American Journal of Evaluation*, 35, 244–249. doi: 10.1177/1098214013503701
- Lincoln, Y. S., & Guba, E. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.
- Maras, M., Coleman, S., Gysbers, N., Herman, K., & Stanley, B. (2013). Measuring evaluation competency among school counselors. *Counseling Outcome Research and Evaluation*, 4, 99–111. doi: 10.1177/2150137813494765
- Marshall, C., & Rossman, G. (2006). *Designing qualitative research* (4th ed.). Thousand Oaks, CA: Sage Publications.
- Martin, I., & Carey, J. C. (2012). Evaluation capacity within state-level school counseling programs: A cross-case analysis. *Professional School Counseling*, 15, 132–143. doi:10.5330/PSC.n.2012-15.132
- Martin, I., Carey, J., & DeCoster, K. (2009). A national study of the current status of state school counseling models. *Professional School Counseling*, 12, 378–386. doi: 10.5330/PSC.n.2010-12.378
- Martin, I., & Rallis, S. (2014). Building on strengths and addressing challenges: Enhancing external school counseling program evaluation. *Journal of School Counseling*, 12(10), 29.
- Mathison, S. (2005). *Encyclopedia of evaluation*. Thousand Oaks, CA: SAGE Publications, Inc.

- McGannon, W., Carey, J., & Dimmit, C. (2005). *The current status of school counseling outcome research*. (Research Monograph No. 2) Amherst, MA: University of Massachusetts at Amherst, School of Education.
- Miller, R. (2008). The oral history of evaluation: The professional development of Daniel L. Stufflebeam. *American Journal of Evaluation*, 29(4), 555–571.
- Newman, D., Scheirer, M., Shadish, W., & Wye, C. (1995). Guiding principles for evaluators. *New Directions for Program Evaluation*, 1995, 19–26. doi: 10.1002/ev.1706
- Palmer, L., & Erford, B. (2012). Predicting student outcome measures using the ASCA national model program audit. *The Professional Counselor*, 2(2), 152-159.
Retrieved from http://tpcjournal.nbcc.org/wp-content/uploads/2012/09/predicting-student-outcome_Palmer-Manuscript_p152-159.
- Patton, M. (2008). *Utilization-focused evaluation* (4th ed). Thousand Oaks: Sage Publications.
- Rogers, P. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*, 14, 29–48. doi: 10.1177/1356389007084674
- Rog, D. (2015). Infusing theory into practice, practice into theory: Small wins and big gains for evaluation. *American Journal of Evaluation*, 36, 223–238. doi: 10.1177/1098214015573068
- Rossi, H., Lipsey, M., & Freeman, H. (2003). *Evaluation: a systematic approach* (7th ed.). Thousand Oaks, CA: Sage.

- Ruhe, V., & Boudreau, J. D. (2013). The 2011 program evaluation standards: A framework for quality in medical education programme evaluations. *Journal of Evaluation in Clinical Practice*, 19, 925–932. doi: 10.1111/j.1365-2753.2012.01879.x
- Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Newbury Park, CA: Sage Publication.
- Scriven, M. (2007.) Key Evaluation Checklist. Retrieved from <http://www.wmich.edu/evaluation/checklists>
- Scriven, M. (1995). The logic of evaluation and evaluation practice. *New Directions for Evaluation*, 1995, 49–70. doi: 10.1002/ev.1019
- Shadish, W., Cook, T., & Leviton, L. (1991). *Foundations of program evaluation: Theories of practice*. Newbury Park, CA: Sage.
- Sink, C. A., Akos, P., Turnbull, R. J., & Mvududu, N. (2008). An investigation of comprehensive school counseling programs and academic achievement in Washington state middle schools. *Professional School Counseling*, 12, 43-53. doi: 10.5330/PSC.n.2010-12.43
- Sink, C. A. (2005). Comprehensive school counseling programs and academic achievement: A rejoinder to Brown and Trusty. *Professional School Counseling* 9, 9-12. doi: 10.5330/prsc.9.1.m2g9176x846618r1
- Stevahn, L. (2005). Establishing essential competencies for program evaluators. *American Journal of Evaluation*, 26, 43–59. doi: 10.1177/1098214004273180
- Studer, J. R., Diambra, J. F., & Breckner, J. A., & Heidel, R. E. (2011). Obstacles and success in implementing the ASCA national model in schools. *Professional School Counseling*. 9(2), 2-26.

- Stufflebeam, D. (1999). *Program evaluations metaevaluation checklist*. Kalamazoo, MI: Western Michigan University.
- Stufflebeam, D. (2001b). The metaevaluation imperative. *American Journal of Evaluation*, 22, 183–209. doi: 10.1016/S1098-2140(01)00127-8
- Stufflebeam, D. (2001a). Evaluation checklists: Practical tools for guiding and judging evaluations. *American Journal of Evaluation*, 22, 71–79. doi: 10.1016/S1098-2140(01)00119-9
- Stufflebeam, D., & Coryn, C. (2014). *Evaluation theory, models, and applications* (2nd ed.). San Francisco, CA: Jossey-Bass.
- Torchim, W. (2009). Evaluation policy and evaluation practice. *New Directions for Evaluation*, 123, 13–32. doi: 10.1002/ev
- Trevisan, M. (2002). Evaluation capacity in K-12 school counseling programs. *American Journal of Evaluation*, 23, 291-305. doi: 10.1177/109821400202300305
- Trevisan, M. (2000). The status of program evaluation expectations in state school counselor certification requirements. *American Journal of Evaluation*, 21, 81–94. doi: 10.1177/109821400002100107
- Vernon, D., & Rainey, J. S. (2009). The school counselor as grant writer. *Journal of School Counseling*, 7(19), 21.
- Whiston, S. C., & Aricak, T. O. (2008). Development and initial investigation of the school counseling program evaluation scale. *Professional School Counseling*, 11, 253-261. doi: 10.5330/PSC.n.2010-11.253

- Whiston, S. C., & Sexton, T. L. (1998). A review of school counseling outcome research: Implications for practice. *Journal of Counseling and Development*, 76, 412-426.
doi: 10.1002/j.1556-6676.1998.tb02700.x
- Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2011). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). Thousand Oaks, CA: Sage.
- Young, A., & Kaffenberger, C. (2011). The beliefs and practices of school counselors who use data to implement comprehensive school counseling programs. *Professional School Counseling*, 15(2), 67-76.
- Zagelbaum, A., T, Kruczek., C, Alexander., & H, Crethar. (2014). Where are we now? An updated review of the school counseling literature for trends and themes. *Journal of School Counseling* 12(4), 25.

Appendix A

Example of Program Evaluation Metaevaluation Checklist-Accuracy Standard

Example of Program Evaluation Metaevaluation Checklist-Accuracy Standard

<p><i>THE ACCURACY STANDARDS ARE INTENDED TO ENSURE THAT AN EVALUATION EMPLOYS SOUND THEORY, DESIGNS, METHODS, AND REASONING IN ORDER TO MINIMIZE INCONSISTENCIES, DISTORTIONS, AND MISCONCEPTIONS AND PRODUCE AND REPORT TRUTHFUL EVALUATION FINDINGS AND CONCLUSIONS.</i></p>
<p>A1 Justified Conclusions and Decisions. <i>[Evaluation conclusions and decisions should be explicitly justified in the cultures and contexts where they have consequences.]</i></p>
<p><input type="checkbox"/> Address each contracted evaluation question based on information that is sufficiently broad, deep, reliable, contextually relevant, culturally sensitive, and valid</p> <p><input type="checkbox"/> Derive defensible conclusions that respond to the evaluation's stated purposes, e.g., to identify and assess the program's strengths and weaknesses, main effects and side effects, and worth and merit</p> <p><input type="checkbox"/> Limit conclusions to the applicable time periods, contexts, purposes, and activities</p> <p><input type="checkbox"/> Identify the persons who determined the evaluation's conclusions, e.g., the evaluator using the obtained information plus inputs from a broad range of stakeholders</p> <p><input type="checkbox"/> Identify and report all important assumptions, the interpretive frameworks and values employed to derive the conclusions, and any appropriate caveats</p> <p><input type="checkbox"/> Report plausible alternative explanations of the findings and explain why rival explanations were rejected</p>
<p><input type="checkbox"/> 6 Excellent <input type="checkbox"/> 5 Very Good <input type="checkbox"/> 4 Good <input type="checkbox"/> 2-3 Fair <input type="checkbox"/> 0-1 Poor</p>
<p>A2 Valid Information. <i>[Evaluation information should serve the intended purposes and support valid interpretations.]</i></p>
<p><input type="checkbox"/> Through communication with the full range of stakeholders develop a coherent, widely understood set of concepts and terms needed to assess and judge the program within its cultural context</p> <p><input type="checkbox"/> Assure—through such means as systematic protocols, training, and calibration—that data collectors competently obtain the needed data</p> <p><input type="checkbox"/> Document the methodological steps taken to protect validity during data selection, collection, storage, and analysis</p> <p><input type="checkbox"/> Involve clients, sponsors, and other stakeholders sufficiently to ensure that the scope and depth of interpretations are aligned with their needs and widely understood</p> <p><input type="checkbox"/> Investigate and report threats to validity, e.g., by examining and reporting on the merits of alternative explanations</p> <p><input type="checkbox"/> Assess and report the comprehensiveness, quality, and clarity of the information provided by the procedures as a set in relation to the information needed to address the evaluation's purposes and questions</p>
<p><input type="checkbox"/> 6 Excellent <input type="checkbox"/> 5 Very Good <input type="checkbox"/> 4 Good <input type="checkbox"/> 2-3 Fair <input type="checkbox"/> 0-1 Poor</p>
<p>A6 Reliable Information. <i>[Evaluation procedures should yield sufficiently dependable and consistent information for the intended uses.]</i></p>
<p><input type="checkbox"/> Determine, justify, and report the needed types of reliability—e.g., test-retest, findings from parallel groups, or ratings by multiple observers—and the acceptable levels of reliability</p> <p><input type="checkbox"/> In the process of examining, strengthening, and reporting reliability, account for situations where assessments are or may be differentially reliable due to varying characteristics of persons and groups in the evaluation's context</p> <p><input type="checkbox"/> Assure that the evaluation team includes or has access to expertise needed to investigate the applicable types of reliability</p> <p><input type="checkbox"/> Describe the procedures used to achieve consistency</p> <p><input type="checkbox"/> Provide appropriate reliability estimates for key information summaries, including descriptions of programs, program components, contexts, and outcomes</p> <p><input type="checkbox"/> Examine and discuss the consistency of scoring, categorization, and coding and between different sets of information, e.g., assessments by different observers</p>
<p><input type="checkbox"/> 6 Excellent <input type="checkbox"/> 5 Very Good <input type="checkbox"/> 4 Good <input type="checkbox"/> 2-3 Fair <input type="checkbox"/> 0-1 Poor</p>
<p>A4 Explicit Program and Context Descriptions. <i>[Evaluations should document programs and their contexts with appropriate detail and scope for the evaluation purposes.]</i></p>
<p><input type="checkbox"/> Describe all important aspects of the program—e.g., goals, design, intended and actual recipients, components and subcomponents, staff and resources, procedures, and activities—and how these evolved</p>

<p>over time</p> <p><input type="checkbox"/> Describe how people in the program's general area experienced and perceived the program's existence, importance, and quality</p> <p><input type="checkbox"/> Identify any model or theory that program staff invoked to structure and carry out the program</p> <p><input type="checkbox"/> Define, analyze, and characterize contextual influences that appeared to significantly influence the program and that might be of interest to potential adopters, including the context's technical, social, political, organizational, and economic features</p> <p><input type="checkbox"/> Identify any other programs, projects, or factors in the context that may affect the evaluated program's operations and accomplishments</p> <p><input type="checkbox"/> As appropriate, report how the program's context is similar to or different from contexts where the program is expected to or reasonably might be adopted</p>
<p><input type="checkbox"/> 6 Excellent <input type="checkbox"/> 5 Very Good <input type="checkbox"/> 4 Good <input type="checkbox"/> 2-3 Fair <input type="checkbox"/> 0-1 Poor</p>
<p>A5 Information Management. <i>[Evaluations should employ systematic information collection, review, verification, and storage methods.]</i></p>
<p><input type="checkbox"/> Select information sources and procedures that are most likely to meet the evaluation's needs for accuracy and be respected by the evaluation's client group</p> <p><input type="checkbox"/> Ensure that the collection of information is systematic, replicable, adequately free of mistakes, and well documented</p> <p><input type="checkbox"/> Establish and implement protocols for quality control of the collection, validation, storage, and retrieval of evaluation information</p> <p><input type="checkbox"/> Document and maintain both the original and processed versions of obtained information</p> <p><input type="checkbox"/> Retain the original and analyzed forms of information as long as authorized users need it</p> <p><input type="checkbox"/> Store the evaluative information in ways that prevent direct and indirect alterations, distortions, destruction, or decay</p>
<p><input type="checkbox"/> 6 Excellent <input type="checkbox"/> 5 Very Good <input type="checkbox"/> 4 Good <input type="checkbox"/> 2-3 Fair <input type="checkbox"/> 0-1 Poor</p>
<p>A6 Sound Designs and Analyses. <i>[Evaluations should employ technically adequate designs and analyses that are appropriate for the evaluation purposes.]</i></p>
<p><input type="checkbox"/> Create or select a logical framework that provides a sound basis for studying the subject program, answering the evaluation's questions, and judging the program and its components</p> <p><input type="checkbox"/> Plan to access pertinent information sources and to collect a sufficient breadth and depth of relevant, high quality quantitative and qualitative information in order to answer the evaluation's questions and judge the program's value</p> <p><input type="checkbox"/> Delineate the many specific details required to collect, analyze, and report the needed information</p> <p><input type="checkbox"/> Develop specific plans for analyzing obtained information, including clarifying needed assumptions, checking and correcting data and information, aggregating data, and checking for statistical significance of observed changes or differences in program recipients' performance</p> <p><input type="checkbox"/> Buttress the conceptual framework and technical evaluation design with concrete plans for staffing, funding, scheduling, documenting, and metaevaluating the evaluation work</p> <p><input type="checkbox"/> Plan specific procedures to avert and check for threats to reaching defensible conclusions, including analysis of factors of contextual complexity, examination of the sufficiency and validity of obtained information, checking on the plausibility of assumptions underlying the evaluation design, and assessment of the plausibility of alternative interpretations and conclusions</p>
<p><input type="checkbox"/> 6 Excellent <input type="checkbox"/> 5 Very Good <input type="checkbox"/> 4 Good <input type="checkbox"/> 2-3 Fair <input type="checkbox"/> 0-1 Poor</p>
<p>A7 Explicit Evaluation Reasoning. <i>[Evaluation reasoning leading from information and analyses to findings, interpretations, conclusions, and judgments should be clearly and completely documented.]</i></p>
<p><input type="checkbox"/> Clearly describe all the assumptions, criteria, and evidence that provided the basis for judgments and conclusions</p> <p><input type="checkbox"/> In making reasoning explicit, begin with the most important questions, then, as feasible, address all other key questions, e.g., those related to description, improvement, causal attributions, accountability, and costs related to effectiveness or benefits</p> <p><input type="checkbox"/> Document the evaluation's chain of reasoning, including the values invoked so that stakeholders who might embrace different values can assess the evaluation's judgments and conclusions</p> <p><input type="checkbox"/> Examine and report how the evaluation's judgments and conclusions are or are not consistent with the possibly varying value orientations and positions of different stakeholders</p>

<input type="checkbox"/> Identify, evaluate, and report the relative defensibility of alternative conclusions that might have been reached based on the obtained evidence <input type="checkbox"/> Assess and acknowledge limitations of the reasoning that led to the evaluation's judgments and conclusions	
<input type="checkbox"/> 6 Excellent <input type="checkbox"/> 5 Very Good <input type="checkbox"/> 4 Good <input type="checkbox"/> 2-3 Fair <input type="checkbox"/> 0-1 Poor	
A8 Communicating and Reporting. <i>[Evaluation communications should have adequate scope and guard against misconceptions, biases, distortions, and errors.]</i>	
<input type="checkbox"/> Reach a formal agreement that the evaluator will retain editorial authority over reports <input type="checkbox"/> Reach a formal agreement defining right-to-know audiences and guaranteeing appropriate levels of openness and transparency in releasing and disseminating evaluation findings <input type="checkbox"/> Schedule formal and informal reporting in consideration of user needs, including follow-up assistance for applying findings <input type="checkbox"/> Employ multiple reporting mechanisms, e.g., slides, dramatizations, photographs, powerpoint®, focus groups, printed reports, oral presentations, telephone conversations, and memos <input type="checkbox"/> Provide safeguards, such as stakeholder reviews of draft reports and translations into language of users, to assure that formal evaluation reports are correct, relevant, and understood by representatives of all segments of the evaluation's audience <input type="checkbox"/> Consistently check and correct draft reports to assure they are impartial, objective, free from bias, responsive to contracted evaluation questions, accurate, free of ambiguity, understood by key stakeholders, and edited for clarity	
<input type="checkbox"/> 6 Excellent <input type="checkbox"/> 5 Very Good <input type="checkbox"/> 4 Good <input type="checkbox"/> 2-3 Fair <input type="checkbox"/> 0-1 Poor	
Scoring the Evaluation for ACCURACY Add the following: Number of Excellent ratings (0-8) _____ x 4 = _____ Number of Very Good (0-8) _____ x 3 = _____ Number of Good (0-8) _____ x 2 = _____ Number of Fair (0-8) _____ x 1 = _____ Total score: _____ = _____	Strength of the evaluation's provisions for ACCURACY: <input type="checkbox"/> 29 (92%) to 32: Excellent <input type="checkbox"/> 27 (84%) to 31: Very Good <input type="checkbox"/> 22 (69%) to 30: Good <input type="checkbox"/> 10 (32%) to 21: Fair <input type="checkbox"/> 0 (0%) to 9: Poor _____ (Total score) ÷ 32 = _____ x 100 = _____ %

Appendix B
Interview Protocol

Interview Protocol

Date of Interview:
Time of Interview:
Job title
Institution
Years of experience in evaluation

I. Introduction:

Thank you for agreeing to participate in this interview. Your responses will help the current study to understand the nature of program evaluation in the field of school counseling. The interview will focus on your experiences evaluating the school counseling programs and interventions. Please remember to answer these questions from your own perspective.

The primary purpose of this study is to explore and assess aspects of program evaluation practice in the field of school counseling. This interview intends to discover the experiences and perspectives of those who are explicitly involved in the evaluation process and activities to assess the effectiveness and values of school counseling programs, interventions, and services.

II. Interview Questions:

1. Please let's start with your own experience of conducting program evaluation in school counseling settings?

Probes:

- a. How many years have you been involved in these evaluation activities?
- b. What is your primary role in these evaluation projects?
- c. Is evaluating school counseling programs mandated or optional? If it's mandated, by whom?

2. What are the evaluator's qualifications to evaluate such programs?

Probes:

- a. Who sets these qualification requirements?
- b. Role of school counseling education to prepare school counselor to conduct sound program evaluation.

3. Could you speak about the professional development offered for school counselors as evaluators?

Probes:

- a. Pre-service/on-service training opportunity on evaluation practice and methods?
- b. Other institutional or organizational assistance- such as conferences, workshops...etc.

4. Who is responsible of funding these program evaluation activities in schools?

Probes

- a. Has evaluation results helped to fund counseling programs and services?

5. Could you tell me what are the major strengths of program evaluation in school counseling context?

Probes:

- a. Role of state and/or district leaders to support evaluation activities?
- b. School administrative supports and participation?
- c. The strengths of involving the stakeholders in the evaluation process?

6. Could you talk about the obstacles that face the implementation of evaluation activities?

Probes:

- a. Challenges at the school, district, or state levels.
- b. Strategies to overcome these challenges?

7. Evaluation as a discipline has a wide range of evaluation methods and approaches. From your experience in evaluation, what methods and approaches are mostly employed to evaluate the school counseling programs and interventions?

8. Can you talk about how program evaluation quality is assured, so that evaluation outcomes would be more credible to decision makers or profession at large?

Probes:

- a. How do you assure your own evaluation's quality?
- b. If used, what evaluation standards are used?

9. Currently, what are the major factors influencing the practice of program evaluation in the field?

10. Is there anything else do you want to add that I should know about?

A. Interview Closing:

- Thank you for your time and participation.
- Reassure confidentiality once again.
- Ask permission to send back the interview transcripts to check and verify his/her responses.

Appendix C
Consent Letter

Western Michigan University
Interdisciplinary Ph.D. in Evaluation

Principal Investigator: Dr. Chris Coryn
Student Investigator: Saeed Almueed
Title of Study: Metaevaluation of school counseling program evaluation

You have been invited to participate in a research project titled "*Metaevaluation of school counseling program evaluation*." This project will serve as Saeed Almueed's dissertation for the requirements of the degree of Doctor of Philosophy. This consent document explains the purpose of this research project and goes over all of the time commitments, the procedures used in the study, and the risks and benefits of participating in this research project. Please read this consent form carefully and completely and please ask any questions if you need more clarification.

What are we trying to find out in this study?

The study is conducted to discover the perceptions and perspective of the participants about the capability of program evaluation to determine the quality, effectiveness, and efficiency of school counseling programs and interventions. This study aims to gain insight into (a) how program evaluation is currently practiced in the field, (b) if these evaluations adhere to professional evaluation standards, and (c) the program evaluations' strengths and weaknesses.

Who can participate in this study?

Eligible participants for this study are authors of the evaluation studies that were obtained for a metaevaluation process as the first phase of this study. These individuals may be school counselors, counselor educators, university researchers and professors, or researchers and evaluators from outside agencies.

Where will this study take place?

This study will take place in the Evaluation Center, at Western Michigan University. The study will utilize a semi-structured telephone interview technique, where participants answer questions via telephone from a location of their convenience.

What is the time commitment for participating in this study?

The duration of each interview is approximately 30 to 45 minutes.

What will you be asked to do if you choose to participate in this study?

You will be asked to participate in an interview about your own experience and perceptions of program evaluation in school counseling context. You have the option to answer the interview questions and elaborate openly on their perspectives.

The interviews will be audio-recorded for transcription and analysis purposes. To ensure your confidentiality during the interview, no questions will be asked that could reveal your personal identity. In case you choose not to be recorded, note taking will be the only method used to obtain the interview responses.

What information is being measured during the study?

As a result of participating in the study, the doctoral candidate will be able to collect information pertinent to (a) program evaluation's strengths and weaknesses, (b) evaluation methodology selection, and (c) key factors affecting the practice and performance of program evaluation in school counseling.

What are the risks of participating in this study and how will these risks be minimized?

Other than the time consumed during the interview, there are no known or anticipated risks with participating in this study.

What are the benefits of participating in this study?

The results of this study hope to benefit both the school counseling and evaluation communities. The participants of the study are given the opportunity to provide their experiences, understandings, and voices over the topic understudy.

Are there any costs associated with participating in this study?

There are no costs associated with participating in this study.

Is there any compensation for participating in this study?

Participation in this study will not include any compensation.

Who will have access to the information collected during this study?

Only the principal investigator and student investigator will have access to the collected information. The information will be kept confidential in password-protected files on the remote server in the Evaluation Center at Western Michigan University. Your name will not be disclosed in the dissertation or report as a result of this study. However, anonymous quotations may be used with your permission.

What if you want to stop participating in this study?

You can choose to stop participating in the study at anytime for any reason. You will not suffer any prejudice or penalty by your decision to stop your participation. You will experience NO consequences either academically or personally if you choose to withdraw from this study.

The investigator can also decide to stop your participation in the study without your consent.

Should you have any questions prior to or during the study, you can contact the primary investigator, Saeed Almueed at 269-779-9326 or saeed.m.almueed@wmich.edu. You may also contact the Chair, Human Subjects Institutional Review Board at 269-387-8293 or the Vice President for Research at 269-387-8298 if questions arise during the course of the study.

This study was approved by Western Michigan University Human Subjects Institutional review Board (HSIRB) on (date). Please do not participate in this study after (one year after approval).

I have read this informed consent document. The risks and benefits have been explained to me. I agree to take part in this study.

I grant the study investigators permission to audio-record the interview:

☐ I agree

☐ I disagree

Clicking the button below indicates my consent to participate in the study.

☐ I agree to participate in the interview.

☐ I do not agree to participate in the interview.

Contact information to schedule interview:

Participant Name: _____

Participant phone number: _____

SUBMIT

Appendix D
HSIRB Approval

WESTERN MICHIGAN UNIVERSITY



Human Subjects Institutional Review Board

Date: August 28, 2015

To: Chris Coryn, Principal Investigator
Saeed Almueed, Student Investigator for dissertation
Daniela Schroeter, Co-Principal Investigator
Beverly Vandiver, Co-Principal Investigator

From: Daryle Gardner-Bonneau, Ph.D., Vice Chair

Re: HSIRB Project Number 15-08-18

This letter will serve as confirmation that your research project titled "Metaevaluation of School Counseling Program Evaluation" has been **approved** under the **expedited** category of review by the Human Subjects Institutional Review Board. The conditions and duration of this approval are specified in the Policies of Western Michigan University. You may now begin to implement the research as described in the application.

Please note: This research may **only** be conducted exactly in the form it was approved. You must seek specific board approval for any changes in this project (e.g., ***you must request a post approval change to enroll subjects beyond the number stated in your application under "Number of subjects you want to complete the study."*** Failure to obtain approval for changes will result in a protocol deviation. In addition, if there are any unanticipated adverse reactions or unanticipated events associated with the conduct of this research, you should immediately suspend the project and contact the Chair of the HSIRB for consultation.

Reapproval of the project is required if it extends beyond the termination date stated below.

The Board wishes you success in the pursuit of your research goals.

Approval Termination: August 27, 2016

1903 W. Michigan Ave., Kalamazoo, MI 49008-5456
PHONE: (269) 387-8293 FAX: (269) 387-8276
CAMPUS SITE: 251 W. Walwood Hall

Appendix E
Recruitment Email

To: Prospect interview participant

Subject: Research Participation Invitation

Dear (*Participant's name*),

My name is Saeed Almueed and I am a doctoral candidate working under the supervisions of Dr. Chris Coryn from the Interdisciplinary Ph.D. in Evaluation program at Western Michigan University. I am writing to invite you to participate in my research study titled "Metaevaluation of School Counseling Program Evaluation". You're eligible to be in this study because you participated in a study(s) aimed to evaluate school-based counseling programs and interventions. I obtained your contact information from the published article(s) or reports of your work.

This study is being conducted to gain insight into how program evaluation is being practiced in the field of school counseling. A semi-structured telephone interview is sought from those who have been part of the evaluation practice in the school counseling profession. Primarily, participation in this interview will be done to uncover the participants' perspectives and experiences in regard to program evaluations' strengths and weaknesses, methodology selection, and the participants' views about other major factors influencing program evaluation practice in the school counseling context.

There are no known risks or costs for participation other than approximately 30-45 minutes of your valuable time to share your views and experiences with us. There are no direct benefits of participation; however, your participation may positively contribute to the current and future practice of program evaluation within school counseling field. All personal identifiable information will be kept confidential and will not be included in the final project report.

Remember, this is completely voluntary. You can choose to be in the study or not. If you'd like to participate please click on the link below to electronically read the consent form and agree or disagree to participate in the study: <https://www.surveymonkey.com/s/DFSJ28T>

If you have any questions, please do not hesitate to contact me at saeed.m.almueed@wmich.edu or Dr. Chris Coryn at chris.coryn@wmich.edu.

* If you decide not to participate in this study, please ignore this invitation email or reply to the researcher to remove your email from the email list.

Thank you very much for your valuable time.

Sincerely,

Saeed M Almueed

Doctoral Candidate

The Interdisciplinary Ph.D. in Evaluation Program (IDPE), WMU

Appendix F

Overall Metaevaluation Ratings For All Evaluations

Title	Author/s	Year	Accuracy Standards								Overall Ratings
			A1	A2	A3	A4	A5	A6	A7	A8	
1. Closing the Gap: A Group Counseling Approach to Improve Test Performance of African-American Students	Bruce et al.	2009	Good	Poor	Poor	Very Good	Fair	Fair	Good	Poor	Fair
2. Using Culturally Competent Responsive Services to Improve Student Achievement and Behavior	Schellenberg et al.	2011	Very Good	Fair	Fair	Fair	Fair	Fair	Fair	Poor	Fair
3. Empowering Students: Using Data to Transform a Bullying Prevention and Intervention Program	Young et al.	2009	Good	Poor	Poor	Good	Fair	Fair	Good	Poor	Fair
4. A school intervention to increase pro-social behavior and improve academic performance of at-risk students	Kilian et al.	2011	Fair	Poor	Poor	Fair	Poor	Fair	Good	Poor	Poor
5. Impact Analysis and Mediation of Outcomes: The Going Places Program	Simons-Morton et al.	2005	Fair	Poor	Fair	Good	Poor	Good	Fair	Poor	Fair
6. Observed Reductions in School Bullying, Non-bullying Aggression, and Destructive Bystander Behavior: A Longitudinal Evaluation	Karin et al.	2009	Good	Poor	Good	Good	Fair	Fair	Fair	Poor	Fair
7. School Counselors Connecting the Dots Between Disruptive Classroom Behavior and Youth Self-Concept	Bidell et al.	2010	Good	Fair	Good	Good	Poor	Very Good	Good	Poor	Fair
8. Bully Busters Abbreviated: Evaluation of a Group-Based Bully Intervention and Prevention Program	Bell et al.	2010	Fair	Poor	Fair	Fair	Poor	Fair	Fair	Poor	Poor
9. Transitioning Hispanic Senior from High School to College	Marsico et al.	2009	Good	Fair	Fair	Good	Fair	Good	Good	Poor	Fair

10. Connectedness and Self-Regulation as Constructs of the Student Success Skills Program in Inner-City African American Elementary School Students	Lemberger et al.	2012	Good	Poor	Poor	Fair	Poor	Poor	Fair	Poor	Poor
11. Evaluation of a Career Development Skills Intervention With Adolescents Living in an Inner City	Turner et al.	2010	Good	Fair	Fair	Fair	Poor	Good	Good	Poor	Fair
12. Effects of School-Wide Positive Behavioral Interventions and Supports and Fidelity of Implementation on Problem Behavior in High Schools	Flannery et al.	2014	Very Good	Poor	Good	Good	Fair	Good	Good	Poor	Fair
13. Impact of a School-Based Dating Violence Prevention Program among Latino Teens: Randomized Controlled Effectiveness Trial	Jaycox et al.	2006	Good	Poor	Fair	Good	Poor	Good	Good	Poor	Fair
14. Empowering children with safety-skills: An evaluation of the Kidpower Everyday Safety-Skills Program	Brenick et al.	2014	Good	Poor	Fair	Excellent	Fair	Good	Good	Poor	Fair
15. An evaluation of Kornblum's body-based violence prevention curriculum for children	Hervey et al.	2006	Fair	Poor	Poor	Fair	Fair	Fair	Fair	Poor	Poor
16. Psychosocial Educational Groups for Students (PEGS): An Evaluation of the Treatment Effectiveness of a School-Based Behavioral Intervention Program	Rebecca et al.	2010	Good	Poor	Fair	Good	Poor	Good	Fair	Poor	Fair
17. Coping Power Dissemination Study: Intervention and Special Education Effects on Academic outcomes	Lochman et al.	2012	Very Good	Poor	Poor	Good	Poor	Good	Good	Poor	Fair

18. Effectiveness of the Surviving the Teens Suicide Prevention and Depression Awareness Program: An Impact Evaluation Utilizing a Comparison Group	Strunk et al.	2014	Good	Fair	Fair	Fair	Fair	Good	Fair	Poor	Fair
19. Using Problem-Based Learning with Victims of Bullying Behavior	Kimberly et al.	2006	Good	Fair	Poor	Fair	Fair	Fair	Good	Poor	Fair
20. Evaluation of a Bullying Prevention Program	Hallford et al.	2006	Good	Poor	Fair	Fair	Fair	Good	Fair	Poor	Fair
21. School-Based Prevention of Depressive Symptoms: A Randomized Controlled Study of the Effectiveness and Specificity of the Penn Resiliency Program	Gillham et al.	2007	Very Good	Fair	Good	Good	Poor	Very Good	Good	Poor	Fair
22. A School-Based Group Activity Therapy Intervention With At-Risk High School Students as It Relates to Their Moral Reasoning	Poane et al.	2008	Good	Fair	Fair	Fair	Fair	Fair	Good	Poor	Fair
23. Evaluation of a High School Peer Group Intervention for At-Risk Youth	Hyunsan et al.	2005	Good	Fair	Fair	Fair	Fair	Very Good	Fair	Poor	Fair
24. Evaluation of a Health Careers Program for Asian American and Pacific Islander High School Students	Yeh et al.	2012	Fair	Poor	Fair	Fair	Poor	Good	Fair	Poor	Fair
25. The Oregon First Step to Success Replication Initiative: Statewide Results of an Evaluation of the Program's impact	Walker et al.	2005	Fair	Poor	Poor	Fair	Poor	Fair	Fair	Poor	Poor
26. Building Skills for School Success: Improving the Academic and Social Competence of Students	Brigman et al.	2007	Good	Poor	Good	Fair	Poor	Good	Good	Poor	Fair
27. Closing the Loop: Incorporating Program Evaluation Into an Elementary School Career Day	Brown-Huston et al.	2014	Poor	Poor	Poor	Fair	Poor	Poor	Fair	Poor	Poor

28. Evaluation of a Truancy Diversion Program at Nine At-Risk Middle Schools	Haight et al.	2014	Fair	Poor	Fair	Fair	Poor	Fair	Fair	Poor	Poor
29. Investigation of the Effectiveness of a School-Based Suicide Education Program Using Three Methodological Approaches	Cigularov et al.	2008	Good	Good	Good	Fair	Fair	Very Good	Good	Poor	Fair
30. Evaluation of California's in-school tobacco use prevention education (TUPE) activities using a nested school-longitudinal design, 2003-2004 and 2005-2006	Park et al	2010	Good	Fair	Poor	Good	Poor	Good	Fair	Poor	Fair
31. Effects of an integrated prevention program on Urban youth transitioning into high school	Valerie et al.	2008	Very Good	Fair	Poor	Very Good	Poor	Very Good	Fair	Poor	Fair
32. Evaluating the effectiveness of a reading remediation program in a public school setting	Downing et al.	2009	Good	Fair	Poor	Good	Fair	Fair	Good	Poor	Fair
33. At-risk ninth-grade students: A psycho-educational group approach to increase study skills and grade point averages	Kayler et al.	2009	Very Good	Fair	Poor	Good	Fair	Fair	Fair	Poor	Fair
34. Utilizing staff perceptions to guide and shape future program planning	Sherwood et al.	2010	Good	Poor	Poor	Good	Fair	Poor	Fair	Poor	Fair
35. Connections Through Clubs: Collaboration and Coordination of a School wide Program	Logan et al.	2008	Good	Poor	Poor	Good	Poor	Fair	Fair	Poor	Fair
36. Improving Reading Fluency and Comprehension Among Elementary Students: Evaluation of a School Remedial Reading Program	Hausheer et al.	2011	Fair	Poor	Poor	Fair	Poor	Fair	Fair	Poor	Poor
37. Effective Counseling Strategies for Supporting Long-Term Suspended Students	Johnson et al.	2006	Good	Poor	Poor	Fair	Fair	Fair	Fair	Poor	Fair

38. Evaluation of an Innovative Approach to Improving Middle School Students' Academic Achievement	Poynton et al.	2006	Very Good	Fair	Fair	Fair	Poor	Very Good	Fair	Poor	Fair
39. Bring out the brilliance: a counseling intervention for underachieving students	Berger et al.	2013	Good	Poor	Poor	Very Good	Fair	Good	Fair	Poor	Fair
40. Using Service Learning to Achieve Middle School Comprehensive Guidance Program Goals	Stott et al.	2005	Fair	Poor	Poor	Fair	Fair	Fair	Fair	Poor	Poor
41. Nebraska School Counseling State Evaluation	Carey et al.	2010	Very Good	Poor	Poor	Very Good	Fair	Fair	Good	Poor	Fair
42. Proactive Schools: Year Two Evaluation of the Impact of the Bridges Curriculum on Middle and High School Students	Carey et al.	2006	Fair	Poor	Poor	Poor	Poor	Fair	Poor	Poor	Poor
43. A Statewide Evaluation of the Outcomes of the Implementation of ASCA National Model School Counseling Programs in Utah High Schools	Carey et al.	2012	Good	Fair	Fair	Very Good	Fair	Fair	Good	Poor	Fair
44. Reducing Levels of Elementary School Violence with Peer Mediation	Schellenberg et al.	2007	Good	Fair	Poor	Fair	Poor	Good	Fair	Poor	Fair
45. Preliminary Evaluation of the Impact of Proactive Schools Curriculum	Carey et al.	2005	Fair	Poor	Poor	Poor	Poor	Poor	Fair	Poor	Poor
46. An Evaluation of Utah's comprehensive guidance program	Nelson et al.	2007	Fair	Poor	Poor	Fair	Poor	Good	Poor	Poor	Poor
47. Evaluating a Small-Group Counseling Program-A Model for Program Planning and Improvement in the Elementary Setting	Bostick et al.	2009	Good	Fair	Fair	Excellent	Fair	Good	Fair	Poor	Fair

48. Integrating Academic Interventions into Small Group Counseling in Elementary School	Steen et al.	2007	Good	Fair	Poor	Fair	Fair	Fair	Fair	Poor	Fair
49. Supporting Academic Improvement among Eighth Graders at Risk for Retention	Mason et al.	2009	Fair	Poor	Poor	Fair	Fair	Fair	Fair	Poor	Poor
50. Utah Comprehensive Counseling and Guidance Program Evaluation Report	Carey et al.	2010	Fair	Poor	Poor	Fair	Fair	Fair	Fair	Poor	Poor
51. An Investigation of Comprehensive School Counseling Programs and Academic Achievement in Washington State Middle Schools	Sink et al.	2008	Fair	Poor	Poor	Very Good	Fair	Fair	Good	Poor	Fair
52. Closing the Achievement Gap of Latina/Latino Students: A School Counseling Response	Leon et al.	2011	Good	Poor	Fair	Very Good	Fair	Very Good	Fair	Poor	Fair
53. Paving the Road to College: How School Counselors Help Students Succeed	Lapan et al.	2010	Very Good	Poor	Poor	Very Good	Fair	Poor	Fair	Poor	Fair
54. Evaluation of an Individualized Counseling Approach as Part of a Multicomponent School-Based Program to Prevent Weight-Related Problems among Adolescent Girls	Flattum et al.	2011	Fair	Poor	Fair	Good	Fair	Very Good	Fair	Poor	Fair