

**An Investigation of Faculty Development Center Staff Evaluators Perceptions
of Processes, Theory, and Models Used in Evaluating Faculty Development**

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Thelma Mae Woodard
May 2013

Copyright © 2012 by Thelma Woodard
All rights reserved.

Dedication

The author Henry Miller said, “Dedication is not what others expect of you, it is what you can give to others.” Well, Mom, I made a decision to be right where I am today.....and yeah, I’m happy! So, I’m dedicating this dissertation to you.

My mother passed away while I endeavored to realize my educational dream, and get through my doctoral program. Anyone who knew Corine Woodard was certain she was irrefutably devoted to me and my older sister, Norma. Humble by nature, and sometimes unassuming, my Mom would transform into an unapologetic, prideful recipient in response to all compliments she received about her daughters. Mom would always respond to the compliments about her daughters with, “I know. You’re absolutely right.” When I asked Mom why she didn’t just say “thank you” in response to the compliments, she would say, “I just want them to know they got it right, and I wholeheartedly agree!” In her eyes, there was no one greater than the two daughters she raised to adulthood.

Mom was dedicated to living life and making the best of it. If I heard her say it once, I heard it a million times, “Happiness is what counts.” There is no bigger proponent of believing in your personal decisions about life and living it to the fullest than was my mother. She was always proud, and supportive, of my life decisions even when she preferred that I live the way she would have lived. I knew Corine was proud of me. After 93 years of living life her way, she transitioned into an energy that continues to transmit the same love for me, and devotion to me. I am grateful that I can still feel it.

Acknowledgements

To the members of my dissertation committee, I offer my sincere thanks to each of you.

Dr. Gary Skolits, thank you for your helpful comments and suggestions, and for providing guidance during this study. I feel fortunate to have had a published and distinguished faculty member as my committee chair.

Dr. Jennifer Morrow, thank you for your helpful comments and suggestions during this study. Thank you for your advice and leadership during my first year in the Evaluation, Statistics and Measurement program; it has left an indelible impression on me.

Dr. John Peters, thank you for your helpful comments and suggestions. I appreciate the extra time you took to review preliminary chapters of this study and provide feedback; and thank you for generously sharing your knowledge and giving me the opportunity to learn more about reflective practice.

Dr. Jennifer Richards, thank you for your helpful comments and suggestions. I especially thank you for helping me align the research questions with the purpose of this study.

To my colleagues, peers, and friends, I offer my heartfelt appreciation to each of you.

Dr. David Schumann and Dr. Taimi Olsen at the Tennessee Teaching and Learning Center (TTLC), thank you for the opportunity to work at TTLC. It was a wonderful experience for me and was undeniably a major part of my professional growth.

My dear colleagues, Ferlin McGaskey and Kelly Ellenburg, thank you for critically reflecting together with me about all sorts of educational topics and allowing me to develop my reflective practice; those conversations mean the world to me.

My dear colleagues, Jason Black and Jeral Kirwan, thank you both for your suggestions, ideas, answers to my statistics questions, and just being good friends.

Carol Price, for contributing your time, insight, and wisdom to this experience, and providing the guidance and valuable feedback I truly needed. Just know that your support and encouragement meant the world to me and I will never forget how much you helped me.

Billy Gibbs, for upgrading me on many, many levels during this process. You are generous.

To my family, I could not have made it this far without your support and love.

Dad, who has always been a stickler for setting goals and have the tenacity to reach them, I want to thank you for all the advice, prayers and love you have given me.

My sister Norma, who has been by my side since my first day on earth, thank you for growing, sharing, caring, and evolving with me. The love we have for each other is unparalleled.

My son, Drew, thank you for being my son. I enjoyed looking after you as a boy, and I now enjoy looking up to you as a man.

Abstract

The purpose of this research was to explore evaluator perceptions of the processes they engage in when evaluating faculty development programs; their use of program theory; and the evaluation models they use to design faculty development program evaluations. Due primarily to the fact that most faculty development centers were established during the last twenty years, (Wright, 2011), previous research has shown a lack of widespread knowledge of program evaluation, and its application to faculty development interventions. Studies indicate a gap in the literature exploring program evaluation in faculty development centers from the evaluator's standpoint. This study was interested in the processes evaluators engage in when they evaluate various programs in faculty development; the extent to which evaluators link evaluations to short, medium, and long term outcomes; and the use of standard evaluation models in faculty development.

A descriptive qualitative method utilizing semi-structured interview questions was used to collect data about the perceptions of evaluators concerning their experiences with program evaluation in faculty development. Five primary themes and one secondary theme emerged from the data analysis of information offered by each participant. The primary themes that emerged were: (a) purposeful selection of programs to evaluate, (b) careful planning of the evaluation methodology, (c) understanding that it takes considerable time to evaluate programs, (d) strategically linking program outcomes to center goals, and (d) implicit application of a utilization-focused approach to planning evaluation; a secondary theme emerged: (a) thinking critically about how to evaluate programs. These themes reflect consistent elements and patterns in participants' comments on programs they select to evaluate, methods used in program evaluation, and examples of how evaluation results can be used.

Implications of the findings are that program evaluators in the faculty development field should familiarize themselves with the evaluation processes, program theories, and evaluation models available in the literature and utilize them in faculty development evaluation. For future study, a descriptive case study of evaluation in faculty development programs is recommended. The suggested study design would include onsite observations, interviews with stakeholders, plus document review and analysis.

Table of Contents

Chapter 1 Introduction.....	1
Statement of the Problem.....	2
Purpose Statement.....	3
Overview Of Study Design and Conceptual Framework	3
Significance of the Study	4
Assumptions.....	5
Limitations	5
Definition of Terms.....	6
Chapter Summary	7
Chapter 2 Literature Review	8
Program Evaluation in Faculty Development.....	8
Program Evaluation Models	12
Professional Development Evaluation Models.....	17
Program Theory	20
Conceptual Framework.....	25
Chapter Summary	27
Chapter 3 Methods	29
Introduction.....	29
Research Questions	29
Study Design and Rationale.....	29
Participants.....	31
Centers and Participants.....	32
Sources of Data	40
Interviews.....	40
Data Analysis	43
Coding and Thematic Analysis	44
Establishing Trustworthiness	45
Worthy Topic	45
Rich Rigor	46
Sincerity	47
Resonance	48
Significant Contribution.....	49
Ethics.....	49
Meaningful Coherence.....	50
Chapter Summary	50
Chapter 4 Findings	52
Introduction.....	52
Research Questions	53
Chapter Summary	88
Chapter 5 Conclusions and Recommendations.....	89
Findings.....	89
Conclusions.....	90
Discussion	92
Implications for Practice	93

Recommendations for Future Research	96
Summary	97
References	99
Appendices	108
Appendix A Informed Consent Statement	109
Appendix B Letter to Participants	111
Appendix C Interview Protocol	112
Appendix D Research Audit Trail	113
Vita	115

List of Tables

Table 2.1 Posavac's Evaluation Models	15
Table 2.2 Patton's Types of Evaluation Approaches.....	16
Table 3.1 Demographics of Faculty Development Centers	36
Table 3.2 Research Questions Mapped with Interview Questions and Question Types	42
Table 4.1 Selection of Programs and Planning of the Evaluation	55
Table 4.2 Understanding That It Takes Considerable Time to Evaluate Programs.....	69
Table 4.3 Thinking Critically About How to Evaluate Programs	71
Table 4.4 Strategically Linking Program Outcomes To Center Goals	81
Table 4.5 Implicit Application Of A Utilization-Focused Approach To Planning Evaluation	87
Table 5.1 Faculty Development Logic Model Components	95

List of Figures

Figure 2.1 Logic Model	22
------------------------------	----

Chapter 1

Introduction

During the 1970s, scholars began to discuss the issue of program evaluation as a process in faculty development programs. Evaluation of such programs is now, of course, a requisite of measuring effectiveness of any formal or informal professional development intervention. Forty years ago, in an educational venue wherein outcome measures were becoming a greater factor in funding and support for any faculty development programs, identifying and assessing the evaluation needs of faculty development centers was a growing area of interest.

At the same time, there was a concern that faculty development centers might soon become extinct if the staff did not learn to evaluate their own programs, thereby demonstrating that they effectively manage their centers, resulting in more productive and engaged faculty members (Centra, 1976). Complicating the situation was the lack of widespread knowledge of program evaluation and its application to faculty development interventions. This lack of understanding of the use of program evaluation in faculty development remains a challenge. This ongoing challenge is due primarily to the fact that most faculty development centers established during the last twenty years, basically in the period from 1990 to 2010 (Wright, 2011).

For many faculty development centers the primary goals are a) to have participating instructors retain and apply ideas presented in center programming and b) to motivate change in instructors' behavior. The cascading result of realizing those goals will be to improve student retention rates, promote more effective student learning, and directly impact student outcomes (Sorcinelli, 2006).

Recent faculty development evaluation studies have focused on particular intervention strategies, and these studies have sometimes included evidence that the programs are having an impact on the participants or the institution, or both. However, these studies typically present vague ideas about program theory or evaluation frameworks (Dwyer, 1996; Metcalfe, Aitken, & Gaff, 2008; Stinson & Wilkinson, 2004). In addition, very few of the evaluation studies reported in the literature utilize an evaluation model (Kreber & Brook, 2001). Having a specific process for and model of how a professional development program is designed to work can help program staff members' work together and focus on those activities that are most important for program success.

Statement of the Problem

A review of faculty development literature reveals that much of the research on faculty development has been conducted quantitatively. Further, the literature review indicates that faculty development program evaluations have failed to address the evaluator's perspective in the evaluation process. Basically, the research ignores the role of the evaluator, and the evaluation process itself, in the evaluation of faculty development programs. Studies of evaluation in faculty development are sparse in the literature published since 1976, and many of the studies prior to 2011 relied on survey methods to investigate this topic (Centra, 1976; Nelson, 1979; Shore, 1976; Sorcinelli, 2006). Faculty development centers have mainly been engaged in evaluation of singular activities (Light, Calkins, Luna & Drane, 2008). However, recent literature does provide evaluation studies using specific evaluation models (Armstrong & Barsion, 2006; Cilliers & Herman, 2009) and program theory (Megdal, Engle, Pakenas, Albert, Peters, & Jordan, 2005).

More recent studies by Hines (2011) and Wright (2011) are indications of a recent shift in the research approach because they investigated or detailed factors contributing to “how” program evaluation is conducted at faculty development centers; however, very little is known about the evaluation process and methods for conducting program evaluation at a faculty development center from the evaluators’ perspective. Resulting studies indicate a gap in the literature exploring program evaluation in faculty development centers from the evaluator’s standpoint. Therefore, the significance of this study lies in its descriptive perspective and its potential ability to address a gap in the literature.

Purpose Statement

The purpose of this research was to explore evaluator perceptions of the processes they engage in when evaluating faculty development programs; their use of program theory; and the evaluation models they use to design faculty development program evaluations.

Overview Of Study Design and Conceptual Framework

This study utilized a descriptive qualitative methodology to answer three research questions:

1. What process do evaluators engage in when they evaluate various programs in faculty development?
2. To what extent and how do evaluators link evaluations to short-, medium-, and long-term outcomes?
3. To what extent and how do evaluators use the standard evaluation models in faculty development?

To answer the research questions, a semi-structured interview was undertaken to explore the experiences of evaluators in faculty development. Posavac’s (2011) evaluation planning process,

which explains how evaluators can plan evaluations, was used as the conceptual framework for this study.

This research was conducted in compliance with the policies established by the Institutional Review Board (IRB) of The University of Tennessee and the researcher subscribed to the principles and standards of professional ethics in all research, development, and related activities involving this study. The IRB review encompassed the research protocol, the informed consent document signed by participants, announcement used in recruiting participants, and other relevant documents (Appendix A). In carrying out its review, the IRB ensured that: (a) any risks to participants that may be incurred were warranted in relation to the anticipated benefits; (b) informed consent documents clearly conveyed the risks and the true nature of research; (c) announcements were not misleading; and (d) the selection of participants was equitable and justified.

Significance of the Study

This study offers three significant contributions for understanding evaluations of faculty development programs. First, evaluators of faculty development centers offer a rich and in-depth description of their experiences of the evaluation process. The results of this study thus provide insights for faculty developers interested in evaluating their programs; they also show how program theory and evaluation models are currently applied to evaluation in the field of faculty development. Second, the findings of this study provide a foundation for future studies and examination of factors related to program evaluation and its application to faculty development. Finally, qualitative methodology proposes a new approach for helping faculty developers understand the current process of program evaluation in the faculty development field as well as

offering insight about program theory and evaluation models that may increase the likelihood of establishing a successful evaluation design and implementation plan.

Assumptions

This study has been conducted based on the following assumptions:

- Responding evaluators were willing to share their true perceptions of the evaluation process.
- The interview questions were worded appropriately to generate responses relevant to the research questions.
- Participant responses were honest and representative of individuals' experiences with program evaluation.

Limitations

This study was conducted based on the following limitations:

- Faculty development centers employing staff designated to conduct program evaluations were the focus of data collection.
- Faculty development centers operating for at least five years were included in data collection.
- A descriptive qualitative study design precludes generalizing the findings.
- Only faculty development centers employing staff whose job responsibility included program evaluation were included in the study.
- This study is limited by the number of center evaluators willing to participate in an interview.

Definition of Terms

For the purposes of this study, the following definitions are being used. Terms will be further defined in the literature review to clarify application to the specifics of this research project, as appropriate.

Assessment: the measurement of one or more variables related to the current condition, ability, status, or knowledge of an individual or program (Wright, 2008). Assessment and evaluation is often used interchangeably; however, evaluation involves “discovering the value and worth of something, the purposes and methodologies are quite different” (Posavac, 2011, p. 3).

Evaluation Model: conception or approach or sometimes a method (e.g., naturalistic evaluation, goal-free evaluation) of doing evaluation (Scriven, 1991).

Faculty Development: programs that focus on the individual faculty member and or graduate student; used interchangeably in the literature to indicate educational and professional development.

Faculty Development Centers: centers that provide programs for faculty and/or graduate students focused on improving teaching in higher education classrooms; also referred to as Teaching and Learning Centers and Instructional Development Centers.

Program Theory: a set of assumptions about the relationships between the strategy and tactics a program has adopted and the social benefits it is expected to produce (Rossi, Freeman, & Lipsey, 1999, p. 98). Each program in an organization may operate under a different set of assumptions; for instance, the strategy and tactics utilized in faculty development in one-on-one consultations with faculty members are different from those utilized in a workshop, and outcomes and follow-up procedures are different. Program theories are different from program

planning and evaluation models. In program planning, the process of the program is detailed; evaluation models are used to analyze questions that the evaluation can answer. On the other hand, program theory examines the process of the program and its value.

Chapter Summary

Chapter 1 presents the introduction to this study and the development of program evaluation in faculty development. Since the development of program evaluation in faculty development, centers have mainly evaluated workshop activities. Studies of program evaluation are sparse and very little is known about how evaluation is conducted in faculty development. Assumptions related to this study are that the interview questions are worded appropriately and participants in this study are truthful. Often, assessment and evaluation are used interchangeably and program theory is often confused with program planning theory. This chapter provides definition of terms related to this study. Overall, the need for a study to describe experiences of program evaluators in faculty development centers was explored.

Chapter 2 reviews the literature on program evaluation in faculty development, program theory, and program evaluation models. Chapter 3 focuses on the research design and the methods for collecting and analyzing the data and includes information about the selection of participants as well as the method for establishing trustworthiness. Demographic data and findings are presented in Chapter 4, and Chapter 5 details implications for research and practice.

Chapter 2

Literature Review

The literature related to the research questions of this study includes a review of program evaluation conducted at faculty development centers in higher education, program evaluation models, and program theory literature. Stufflebeam (2001) defined an evaluation as “a study designed and conducted to assist some audience to assess an object’s merit and worth” (p. 11). Thus, program evaluation assesses the value of a specific program. The evaluation process is a set of activities that provides data about the program; specific methods used to gather and interpret data depend on the purpose of the evaluation, which can be formative or summative. Formative evaluation is conducted during the program to assess the implementation process or development of the program and provide data relative to the improvement of the program; summative evaluation is conducted at the end of the program to determine the program’s worth or effectiveness (Scriven, 1991).

Program Evaluation in Faculty Development

Most faculty development centers were established in the period from 1990 to 2010 (Wright, 2011). Faculty development centers design programs specific to the needs of their campus and which delivered through workshops, seminars, institutes, consultations, and observations; in addition to faculty development some centers also focus on instructional technology expertise, administer course evaluations, teach assistants, and supply grants to support faculty projects. The primary goals for many faculty development centers are to have instructors retain and apply program ideas presented and to change instructors’ behavior in order to promote more effective student learning and retention (Sorcinelli, 2006; Wright, 2011). The Professional and Organizational Development in Higher Education (POD Network) supports a

network of over one thousand faculty and teaching assistant developers, faculty, administrators, and consultants in higher education. The POD network website lists sixty-five faculty development centers with websites in the United States.

In the late 1970s it was predicted that if faculty development centers did not soon learn to evaluate their programs and demonstrate that they successfully manage their center by educating the teaching faculty to produce more effective and satisfied faculty members, they might soon become extinct (Centra, 1976). Also during the 1970s, scholars began to identify and assess the evaluation needs of faculty development centers. Shore (1976) emphasized that centers should determine who would conduct evaluations, how often those evaluations will be conducted, and what would happen to the results. Centra (1978) and Nelson (1979) also drew attention to the need for program evaluation in faculty development and argued that evaluation is an essential indicator of program effectiveness. At around the same time, Hoyt and Howard (1978) differentiated among the three types of data useful in evaluation studies: that which describes how participants feel about the experience, that which describes behavioral change, and that which provides direct evidence of improvement in instructional effectiveness (p. 26).

Faculty development literature from the 1970s was scattered with descriptions of evaluations but lacked evidence about the impact of the programs on participants or their institutions (Gaff & Morstain, 1978). In the early 1980s, faculty developers began to implement evaluation methods for evaluating impact. Levinson-Rose and Menges (1981) reported that the strongest index of program effectiveness is impact on students and the weakest is self-reporting by participants attending faculty development programs. In addition, Van Note Chism and Szabo (1997) offered a method for conducting program evaluation in faculty development after they observed that evaluation could be performed at three levels: Level one determines how

satisfied participants are with the programs; level two determines the impact of the program on the participant's teaching practices and on their attitudes toward teaching; level three determines the impact of the program on the learning accomplished by the participants' students. Levinson-Rose and Menges (1981) found that most program evaluations conducted at faculty development centers only reflected superficial levels of experience and allowed only speculation about the cognitive and developmental experiences of program participants.

Even though Van Note Chism and Szabo (1997) offered an evaluation method for faculty development, the most common types of faculty development program evaluation currently found in the literature are needs assessments and post-workshop evaluations. For example, Milloy and Brooke (2004) described a faculty development needs assessment used to focus a center's programming and guide decision-making. Other studies described needs assessment studies that inform programming (Sorcinelli, 2002; Sorenson & Bothell, 2004; Travis, Hursh, Lankewitz, & Tang, 1996). Over time, research studies mainly described or reported on the evaluation of singular initiatives (Centra, 1976; Ferren & Mussell, 1987; Hines, 2011; Wright, 2011). For instance, faculty and teaching assistants who participated in one-on-one consultations were surveyed to assess the long-term value and effects of consultations (Jacobson, Wulff, Grooters, Edwards, & Freisem, 2009). The survey found that consultations contributed to the development of participants' teaching. This evaluation study is helpful for learning more about the ongoing effects of the consultation service, but it may not be applicable to other centers because not all development programs provide one-on-one consultations.

Other evaluations of singular activities found in the literature includes reporting on learning communities (Light, Calkins, Luna, & Drane, 2008), return on investment (Bothell & Henderson, 2004), teaching assistants (Way, Carlson, & Piliero, 2002), and Science,

Technology, Engineering and Mathematics (STEM) (Connolly & Millar, 2006). Light et al. (2008) demonstrated a positive relationship between key constructs of an extended model of teaching and learning using a mixed-method approach for a year-long development program designed for pre-tenure faculty. Bothell and Henderson's (2004) evaluation study indicated a positive return on investment based on student retention in a freshman seminar program taught by teachers involved in faculty development. Way et al. (2002) evaluated workshops designed for teaching assistants to determine transfer of training and reported that participants who attend training perceived that their peers who did not attend training scored lower on performance evaluations. Connolly and Millar (2006) evaluated faculty development workshops in the STEM field to determine ways of measuring the impact of workshops in order help workshop providers design better workshops and assist prospective participants in choosing appropriate workshop programs. All of the studies reported that making the connection between a workshop and changes in student learning is daunting and costly.

Many of the substantial faculty development evaluations have been conducted in medical and nursing schools (Metcalf, Aitken, & Gaff, 2008; Sarikaya, Kalaca, Yegen, & Cali, 2010; Stinson & Wilkinson, 2004; Sullivan, Lakoma, Billings, Peters, & Block, 2005). These evaluations have often gone farther than the satisfaction surveys found in other higher education journals and report the behavioral changes resulting from faculty development. For instance, Sullivan et al. (2005) evaluated the effectiveness of faculty development in a palliative care program at Harvard Medical School and found that preparation for teaching end-of-life care increased and respondents reported behavioral changes in patient care and teaching as a result of participation in the program.

The ultimate goal of teaching is subsequent learning, so the ultimate measure of effectiveness of a teaching workshop is the improvement in the participant's students' learning that can be attributed to the workshop (Wright, 2011). However, Levinson-Rose and Menges (1981) noted that workshops and seminars are probably the most frequent but least evaluated instructional improvement activity, which raises questions about the lack of literature describing the assessment of merit and worth of specific faculty development programs. Improvements in students' learning cannot be assumed to follow from their teachers' satisfaction with a workshop and may only be inferred indirectly from the changes in the teachers' instructional practices following workshop attendance (Felder & Brent, 2010). Additionally, according to limited studies, in the late 1990s faculty development centers lacked a systematic approach to evaluation (Hines, 2009). More recent literature has observed that program evaluation is an ongoing concern for faculty developers (Hines, 2009).

Faculty development literature now offers many recommendations on how to evaluate faculty development programs; however, there are few examples of how these centers actually conduct program evaluation. To date, the literature has provided a report of a singular event or recommendations about the need for program evaluation. More useful would be a description of the process they engage in when evaluating faculty development programs, the evaluation model used, or explanations of the program's theory of change. The following section discusses program evaluation models and program theory.

Program Evaluation Models

Evaluation models provide a method for designing and planning for the examination of desired outcomes, program delivery methods and measurement approaches (Patton, 2008). These models may be utilized by centers employing evaluators, although they may be

unpublished. Evaluation models can aid in systematically assessing the impact of faculty development programming on college campuses.

Early in the field of program evaluation, evaluation models were named after the individual who developed them. For instance, the Tyler (1942), Metfessel and Michaels (1973), and Hammond (1973) evaluation models were used to evaluate curriculum and educational programs. There are still a few evaluation models named for their developers, such as Scriven's goal-free evaluation model, Stufflebeam's CIPP evaluation model, and Guskey's (1999) and Kirkpatrick's (1994) professional development evaluation models, however, as the program evaluation field grows, numerous models have emerged based on alternative approaches.

Recently, the different types of evaluations point to different questions and focus on different purposes. Posavac (2011) provided a general overview of thirteen evaluations models with different approaches, types, and/or areas of focus. Table 2.1 provides the name and description of the six of the thirteen model listed by Posavac (2011). The other models not listed in the table include:

- Traditional
- Social science research
- Industrial inspection
- Black box evaluation
- Accountability
- Expert opinion
- Naturalistic or qualitative

The above models are not included in Table 2.1 because they either relate closely to research methods or they are models relevant to an industry unrelated to faculty development. For

example, the black box evaluation model listed above is a consumer model often used to evaluate manufactured objects. These models are not the only evaluation models available for use in planning program evaluations. Patton (2008) listed seventy-nine evaluation models with different approaches, types, and/or areas of focus. Of the evaluation models listed by Patton, many are appropriate for faculty development program evaluation (Table 2.2). Table 2.2 provides a list of nine evaluation models that include the evaluation focus, defining questions and approach to using each model in faculty development. Posavac (2011) and Patton (2008) provide fifteen evaluation models appropriate for use in evaluating faculty development programs, however, with the exception of Cilliers and Herman (2009), very few of the evaluation studies reported in faculty development literature have utilized an evaluation model.

Table 2.1 *Posavac's Evaluation Models*

<i>Model</i>	<i>Defining Questions or Approach</i>
Fiscal	Calculations of the financial investment needed to support the program and the return on that investment
Goal-free	Involves studying the program as administered, the staff, the clients, the setting, and the records to identify all the positive and negative impacts of the program
Improvement-focused	Program improvement is the focus rather than particular methodologies
Objectives-based	Emphasizes working with clearly stated program goals and objectives so that the degree to which such goals and objectives are achieved can be measured
Success Case	Provides detailed information from participants who have benefitted from the program
Traditional	Have resources been appropriately used to accomplish intended results? Key issue: Who is accountable to whom for what?

Table 2.2 *Patton's Types of Evaluation Approaches*

<i>Type of Evaluation</i>	<i>Defining Questions or Approach</i>
Accountability focus	Have resources been appropriately used to accomplish intended results?
Effectiveness focus	To what extent is the program effective in attaining its goals?
Effort focus	What are the inputs into the program in terms of number of personnel, staff/client ratios, and other descriptors of levels of activity and effort in the program?
Formative evaluation	How can the program be improved?
Improvement focus	Program improvement is the focus rather than particular methodologies
Mission focus	To what extent is the program or organization achieving its overall mission? How well do outcomes of departments or programs within an agency support the overall mission?
Program theory	Making explicit and testing the program's theory of change: What is the programs theory of change and to what extent do empirical findings support the theory in practice?
Theory of change approach	What are the linkages and connections between inputs, activities, immediate outcomes, intermediate outcomes and ultimate impacts?
Utilization-focused evaluation	What information is needed and wanted by primary intended users that will actually be used for program improvement and decision- making?

Professional Development Evaluation Models

Both the Guskey (1999) and Kirkpatrick (1994) models may be appropriate for evaluating faculty development because they are designed for evaluating employee training and professional development (Zepeda, 2008). A recent evaluation study used Kirkpatrick's professional development evaluation model to examine the impact of faculty development efforts in order to determine the level of impact a faculty development program had on faculty members' teaching practice over time (Cilliers & Herman, 2009).

Kirkpatrick's professional development evaluation model includes four levels of evaluation to identify a return on expectations: The first level addresses the degree to which participants react favorably to the training; the second addresses the degree to which participants acquire the necessary knowledge, skills, attitudes, confidence, and commitment based on their participation in a training event; the third addresses the degree to which participants apply what they learned during training once they are back on the job; the fourth addresses the degree to which targeted outcomes occur as a result of the training event and subsequent reinforcement. By utilizing interviews and surveys with program participants, the data showed high-level impact when rated according to Kirkpatrick's framework (Zepeda, 2008), including "changes to individual behavior and organizational practice, benefits to academics and perceived benefits to their students" (Bates, 2004).

Guskey's (1999) professional development evaluation model, which is popular in K-12 educational settings (Zepeda, 2008), is based primarily on the work of Kirkpatrick (1959), whose model is very popular in business training settings. Guskey's (1999) model involves five levels of planning and evaluation to improve professional development: participants' reactions, participants' learning, organization support and change, participants' use of new knowledge and

skills, and student learning outcomes. The levels are arranged hierarchically, with each succeeding level building on previous levels and becoming increasingly more complex. In other words, success at one level is necessary for success at the levels that follow (Guskey, 1999). The first level is the most common and simplest form of program planning and evaluation, and it is the level at which program developers have the most experience.

The first level is designed to capture the participants' reactions to the professional development experience. According to Guskey (1999), it is also the easiest type of information to gather and analyze. Assessment at this level would include addressing whether participants enjoyed the program, whether they thought their time was well spent, whether the materials made sense, and whether it was useful. In addition, other considerations would include whether the facilitator was helpful and knowledgeable and the refreshments and room temperature were acceptable. Usually at this level questionnaires are administered at the end of the session to determine initial satisfaction with the program. Data collected at this level is used to improve program design and delivery.

The second level focuses on measuring the knowledge, skills, and attitudes participants gained from their professional development experience. This level is designed to capture participants' learning. Assessment at this level would include focusing on collecting data on the extent to which faculty participants learned the intended material. Assessment at this level would include questioning faculty participants' reactions to the training experience to determine whether participants acquired the intended knowledge and skills; the new knowledge and skills of participants are also assessed. Information at level two can be obtained by using paper and pencil instruments, simulations, demonstrations, participant reflections, and participant

portfolios. Data collected at this level is used to improve program content, format, and organizations (Guskey, 1995).

The third level focuses on organizational characteristics and attributes necessary for success, two of the more complicated aspects of planning and evaluation, and assesses organization support and change. At the third level, assessment establishes the extent to which the organization supports professional development training. In addition, evaluation questions might focus on the impact faculty development had on the organization. Data gathered at this level would be derived from university records, follow-up meetings, questionnaires, structured interviews with administrators, and participant portfolios. The organization's advocacy, support, accommodation, facilitation, and recognition are measured, and data is used to document and improve organization support and inform future changes.

The fourth level focuses on whether participants are making use of their new knowledge and skills on the job. Assessment at this level would include questioning the extent to which faculty the skills learned during the course of the program transferred to the classroom by discovering whether participants applied their new knowledge and skills. Data at this level is obtained using questionnaires, structured interviews with participants and their supervisors, participants' portfolios, direct observation, and video and audiotapes. The degree and quality of implementation is measured, and the information is used to document and improve the implementation of program content.

Building on previous levels, the fifth level focuses on assessing the impact professional development has on students by asking whether the program affected student performance or achievement, if it influenced students' physical or emotional well being, whether students were more confident learners, and if student attendance improved. The fifth level is designed to

capture information on student learning outcomes, and thus assessment at this level would include determining the extent to which students learning change. To answer these questions, student records, questionnaires, structured interviews, and a portfolio of work samples are planned and implemented; in addition, cognitive, affective, and psychomotor student learning outcomes are measured. Data gathered at this level is used to focus and improve all aspects of program design, implementation, and follow-up and to demonstrate the overall impact of faculty development.

Both the Kirkpatrick and Guskey professional development evaluation models are more in-depth than the Van Note Chism and Szabo (1997) program evaluation model for faculty development. Whereas Van Note Chism and Szabo (1997) observed that evaluation of a faculty development program could be performed at three levels—satisfaction, impact of the program on teaching practices and attitudes toward teaching, and impact of the program on student learning—Kirkpatrick (1959) and Guskey (1999) increased the number of evaluation levels. However, none of these three models include a level devoted to program theory in which it is necessary to assess the evaluability of a program by constructing a logic model and showing how the elements of that program lead to the expected changes in program participants. The next section addresses the program theory and logic model literature.

Program Theory

Having a common model for a professional development program can help program staff work together and focus on those functions that are most important for program success. A well-developed and validated theory like the program theory evaluation approach can be used to develop propositions about ideal outcomes; the program is then compared to the theoretical model. Making explicit the underlying assumptions about how programs are expected to work,

describing the program theory, and then using theory to guide the evaluation can provide the tool necessary for documenting program impact. According to Patton (2008), “before undertaking an evaluation, the program should be clearly theoreticalized as some identifiable set of activities that are expected to lead to some identified outcomes...the linkage between those activities and outcomes should be both logical and testable” (p. 334).

Building a plausible model of how the program is meant to work can help directors; staff and other stakeholders identify the most important processes or outcomes on which to focus their measurement and attention. The logic model provides a hypothesis of how the program is supposed to work in order to achieve the desired results (Fig. 2.1). Descriptions and examples of the use of logic models are found in Frechtling (2007) and Knowlton and Phillips (2009). Variations of the logic model are identified different names: “chains of reasoning,” “theory of action,” “performance framework,” and “logical framework” (McLaughlin & Jordan, 1999).

According to Chen (1990), program theory distinguishes how the program is understood to work and how the program actually works, and the theory should be both prescriptive and descriptive. The logic model and these variations are all names for what evaluators term “program theory” (McLaughlin & Jordan, 1999). Another example is the development of a faculty development program theory based on the mission of the center or the development of a logic model for each program. Logic models are useful for identifying projects that are critical to attaining goals or which have inconsistent or implausible linkages among program elements. Figure 2.1 depicts the logic model components that include inputs, activities, outputs, and impacts. Each of these components can also help centers communicate underlying assumptions about how their programs are expected to work and then situate faculty development within the organization if stakeholders are included in the development.



Figure 2.1 Logic Model

The hypothesis is that if assumptions about contextual factors remain correct and the program uses these resources with these activities, then it will produce these short-term outcomes for identified participants who will use them, thus leading to longer term outcomes (Frechtling, 2007). In order to check the validity of a logic model, the program logic can be described as a linked hypothesis and mapped graphically with a series of “*if-then*” statements about the program. Using an earlier example, *if* faculty members participate in a learning community for one semester, *then* they will be more collegial with participating colleagues. The validity of a program theory can also be tested through pattern matching, which is the idea that theoretical patterns can be compared with the observed pattern, and a strong match between the two supports the theory. Yin (1989) described pattern matching as an important tool for the delivery and impact of a program.

Program theory can provide detailed information about the aspects of a program and identify what leads to success in order to replicate programs in new and varied settings (Hacsi, 2000); however, the program theory must be tested in order to fully accrue the benefits (Mark, 1990). According to Megdal et al. (2005), the problem with most educational interventions is that their efficacy is not firmly established, meaning that there is a lack of evidence that the intervention is effective. According to Weiss (2000), social science literature and the beliefs of program stakeholders are the two major sources of program theory. Both are credible sources as long as they are made explicit in terms of a specific program. One theory might develop based

on the beliefs of stakeholders associated with the program; for example, participation in a learning community enhances collegiality.

Another theory might be based on plausibility; for example, if teachers learn Bloom's taxonomy they will be able to design their own course using the taxonomy. Many faculty development program developers may assume that providing information to program participants will lead to a change in their knowledge, and increased knowledge will lead to positive change in behavior. This theory is the basis for many faculty development programs. With so much effort expended in providing information in order to change behavior, careful investigation of this theory is warranted. Studies have shown that information can lead to change in knowledge and attitudes but rarely to change in behavior (Hacsi, 2000; Mark, 1990; Weiss, 2000). Young (1987) highlighted the fact that many program evaluations in higher education focus on describing the process and products of programs and very few systematically look at program goals in advance of implementing and assessing a program.

Using a methodological framework to interpret data is a critical tool in evaluation (Chen, 1990; Rogers, 2007). In addition to using a framework to interpret data, a framework may be used to plan and design the evaluation. For instance Wright (2011) utilized a matrix in which key evaluation questions are listed and the methods for answering each question are described as a means to evaluate the center's impact on the university, faculty and student body. The Metcalfe et al. (2008) study compared the evaluations of an education intervention, a needs assessment, and community education program in genetics counseling and thus provides a glimpse of the program theory of professional development. This study emphasized the diverse applicability of program evaluation.

More recent literature has described the importance of evaluating impact. For instance, Sarikaya et al. (2010) used program evaluation to determine the impact of faculty development on medical educators from clinical disciplines. They found that participants modified their teaching activities and that the faculty training programs did have long-term impact on the faculty, learning environment, and culture at their institutions. Stinson and Wilkinson (2004) used Dwyer's (1996) logic model framework to plan, implement, and evaluate a clinical extern program, and they reported that the logic model allowed them to expand the program, increase the number of participants, and improve retention, while keeping patient safety a top priority. Armstrong and Barsion (2006) used an outcomes-logic-model approach to examine the impact on faculty members participating in a faculty development program. The researchers used structured telephone interviews and logic model pattern matching (McLaughlin & Jordan, 1999; Yin, 1989) as a tool to study how well the program mission and plan were implemented and whether outcomes had been attained. The researchers also examined outcomes such as greater use of active learning principles. Building and using a logic model provided a structure to examine the degree to which the desired outcomes, program delivery methods, and measurement approaches were aligned.

Felder and Brent (2010) described the design of an evaluation study to test the hypothesis that a three-day teaching workshop that had met once a year for 19 years met its goals. Institute goals were to improve participants' teaching effectiveness, promote engagement in scholarly teaching and educational scholarship, and motivate participants to engage in instructional development on their home campuses (Felder & Brent, 2010). A web-based survey was administered to alumni of the workshop to identify those factors in the workshop's structure and delivery that might have contributed to its success. Data were interpreted using adult motivation

theory and revealed that the workshop did “motivate its participants to adopt or increase their use of proven teaching strategies known to correlate with improved student learning; made them more student-centered, scholarly, and reflective in their teaching practice; and induced many of them to engage in instructional development and educational scholarship” (Felder & Brent, 2010, p. 4).

Conceptual Framework

This study will address the experiences of program evaluators in faculty development by using Posavac’s (2011) evaluation planning process, which explains how evaluators can plan evaluation. Posavac’s (2011) evaluation planning process is a popular process in the evaluation field and is a conceptual framework used to train novice evaluators about the evaluation planning process. The process involves several steps.

The first step in the evaluation planning process is to review evaluation models. Posavac (2011) found that “one value of thinking about models lies in developing an appreciation of the range of questions evaluators can consider. The specific questions being addressed by an evaluation or the specific aspects of the program setting often make one or another of the models especially useful” (p. 24). The second step in the evaluation planning process identifies the program to evaluate. Evaluators should obtain a complete program description because it “makes a difference whether an evaluation is proposed for a new program or a well-established, whether the program is locally managed or offered at many sites and whether the program theory is well articulated or based on common sense” (p. 30). The third step identifies relevant stakeholders; stakeholders include program personnel, program sponsors, and program clients. The fourth step in the process includes ascertains information needs. The evaluator should answer the following questions: “Who wants the evaluation? What should be the focus of the

evaluation? Why is the evaluation wanted? When is the evaluation wanted? What resources are available?” (p. 31). In addition, it is necessary to assess the evaluability of the program by constructing a logic model showing how the elements of the program lead to the expected changes in the program’s participants.

After the information needs are identified, the evaluator should examine the relevant literature before designing or developing new instruments so as to provide “a picture of the methodological, political, and practical difficulties” of in the field (p. 36). Evaluators should:

Keep several key questions in mind: In what ways is the program evaluated similar to the program being considered for evaluation? What research designs were used? What research designs were used? Can some of the measures of the outcome criteria be adapted? How reliable and valid were the measures? What statistical analyses were used? Is there a consensus among the reports? Are there are conflicting findings; are these conflicts due to different approaches to sampling, design, or interpretation? What issues were not addressed? (Posavac, 2011, p. 38).

The evaluation methodology should be planned, giving careful thought to how the evaluation is to be reported to stakeholders. Finally, the evaluation plan is written.

Although several evaluation planning process models exist in the literature, Posavac’s (2011) model is an appropriate choice for this study. The steps described above provide a framework for analyzing the process evaluators engage in when they evaluate various programs in faculty development (i.e. workshops, one-on-one consultations, and learning communities). Posavac’s (2011) evaluation planning process model is particularly useful in this study as each

faculty development program has different goals related to a theory of change; therefore, activities, objectives, and processes are possibly quite different on every college campus. The model offers a systematic way to examine the process in which evaluators engage.

While ideally all evaluators in this study would evaluate programs according to the Posavac's (2011) model, evaluators may not necessarily utilize each step. Evaluators may engage or emphasize various steps and in different order by programs. However, the model still provides an appropriate lens through which to view this study and the proposed research questions.

Chapter Summary

Although most faculty development professionals value the importance of monitoring their programs' impact, systematic evaluation is still not common and often relies on inference measures such as extent of participation and satisfaction (Kreber & Brook, 2001). Evaluation models, outcome logic models, and program theory are rarely reported in evaluation of faculty development literature. Very few describe alignment of focus (i.e., consultations, workshops) with the desired change; the intervention strategies used to bring about such change, as well as the mission of the center and university exist in the literature. However, the theory about how and why the program works or should work must be explicit so that other centers may benefit.

Suggestions for improved program evaluation in faculty development include selecting appropriate data sources, stakeholder involvement, and including context in the evaluation (Wright, 2011). Improvements or impacts relevant to the theory of change are often implicit in the evaluation studies reported in the literature. However, explicit program theories of professional development are absent from medical literature. Recent faculty development evaluation studies have focused on particular intervention strategies and include evidence that

these programs are having an impact on their participants and/or institutions, but they do not offer a definitive model of the program theory or evaluation framework (Armstrong & Barsion, 2006; Cilliers & Herman, 2009; Felder & Brent 2010; Krebera & Brook 2001; Wright, 2011).

Chapter 3

Methods

Introduction

This chapter introduces and describes the methods and procedures that are used to conduct this study, including the methods used to explore evaluator perceptions of the processes they engage in when evaluating faculty development programs; their use of program theory; and the evaluation models they use to design faculty development program evaluations. Three research questions identify the objectives of the study.

Research Questions

Three research questions guided this study:

1. What process do evaluators engage in when they evaluate various programs in faculty development?
2. To what extent and how do evaluators link evaluations to short-, medium-, and long-term outcomes?
3. To what extent and how do evaluators use standard evaluation models in faculty development?

Study Design and Rationale

Hines (2011) indicated that an in-depth study using qualitative methodology is necessary to gain further insight into assessment at faculty development centers. Hence, the focus of this study is to understand and describe the evaluation processes and methods used by evaluators involved in faculty development. Qualitative research allows the researcher, as well as readers and participants, to interpret results and advance understanding. Qualitative research is also interpretative research (Creswell, 2007); therefore, this study on perceptions of evaluation in

faculty development centers does not stipulate an approved method for how program evaluation should be conducted in faculty development. Instead, it is a study of evaluators' perceptions of how they conduct evaluation of faculty development programs. Using a qualitative method allows for the exploration and description of the evaluation processes at faculty development programs and takes into account the program theories and models that have been applied. The topic of program evaluation, including the research questions raised, lends itself to a qualitative approach.

Maxwell (1996) outlined five areas that are suitable in qualitative studies, and each area is consistent with the overall focus of this study. The first area involves understanding the participants' perspectives; a review of faculty development reveals that much of the research on faculty development did not investigate program evaluation from the perspective of the evaluators' experiences, and this study will explore those perceptions. Second, qualitative research augments understanding the context of the participants' experiences and how this context affects their actions. In addition, the literature often describes evaluations of particular intervention strategies but does not help us understand the context of the participants; this study will illuminate the evaluation process within the context of faculty development. Third, the researcher can discover new or unplanned occurrences that can lead to new theories or areas for future study. Fourth, qualitative studies assist in understanding the process of how the events and actions of the participants take place; this study focuses on the process rather than the outcome, hence the need for a qualitative design. Last, qualitative research is useful in developing causal explanations or logic for how events and different process can lead to particular results.

This study employed a qualitative descriptive design and, while exploring the phenomenon of program evaluation in faculty development centers that employ evaluators, the researcher gained an understanding of the process the participants engage in when they evaluate programs in faculty development from knowledgeable representatives of each organization.

Participants

Criterion sampling is a purposive method for picking all cases that meet some criterion (Patton, 1990) and was used to find participants for this study. Eligible participants met two criteria: (a) responsible for evaluation and assessment in faculty development, and (b) employed by a center established for five years or longer. Using a purposive criterion sampling strategy, a search of the sixty-five faculty development centers found on the Professional and Organizational Development in Higher Education (POD) website revealed only seven faculty development centers employing program evaluators. All seven centers were invited to participate in this study.

The evaluators were initially contacted by email (Appendix B), which conveyed the nature of this research project, how each potential participant was identified, and the confidentiality of the research. Three evaluators immediately responded by email and agreed to participate. Another evaluator who responded did not meet the criteria for this study. Three potential participants did not respond to any of three email invitations to participate in this study. The third email attempt was followed by a telephone call that left a short message describing the study and inviting each of the three evaluators to participate in this study; none of these contacts elicited a response.

Program evaluation in faculty development is a new phenomenon, and it is uncommon to find a large number of full-time employed evaluators at faculty development centers. Due to the

low participation response, snowball sampling was used to identify additional participants engaged in conducting evaluation and assessment in faculty development. The snowball sampling method requires identifying people who may know anyone who could be a good source (Patton, 1990). Each of the participating evaluators was asked if he or she knew of other evaluators in faculty development; this yielded five evaluators who were invited by email to participate in this study. Four responded immediately by email and agreed; one participant did not respond, and there was no response to a follow-up telephone call.

Centers and Participants

Eight program evaluators, from eight different faculty development centers, participated in this study. Descriptions of each center and each participant's background are provided. All of the centers and evaluators were assigned pseudonyms to protect their anonymity. Therefore, some details about the evaluators were not included in the description. For example, providing educational disciplines for each of the evaluators may identify participants due to the small number of faculty development evaluators. The demographics of the faculty development centers, and data regarding the number of students attending the university, institution type, and the number of programs and staff supported by the centers was also identified.

Centers

Center A is situated in a public university in the Western Pacific region of the United States. The university has an enrollment of over 35,000 students, of whom 64% are undergraduates and 36% are graduate students. The Center employs approximately 24 staff members and promotes 10 instructional development programs, as well as a variety of smaller short-term programs. The center serves approximately 4,000 faculty members, graduate student instructors, graduate students (including non-teaching), postdoctoral faculty and administrators

each year. The mission of the Center is to support the instructional mission of the university and enhance teaching and learning opportunities on campus.

Center B is situated in a public university, also in the Western Pacific region of the United States. The university has an enrollment of over 37,000 students, 69% of whom are undergraduates and 31% graduate students. The Center employs approximately 15 staff members, promotes major programs, and serves approximately 4,000 faculty members, instructors; graduate teaching assistants, graduate students, and postdocs each year. The mission of the Center is to promote teaching practice grounded in pedagogical scholarship.

Center C is situated in a private research university within the Midwestern region of the United States. The university has an enrollment of over 16,000 students, divided between 51% undergraduate and 49% graduate students. The Center employs approximately 16 staff members, promotes 14 instructional development programs, and offers many other programs addressing undergraduate student teaching and learning. The mission of the Center is to provide professional development experiences that enhance the academic culture and encourage innovation and excellence in teaching.

Center D is situated in a public research university within the Eastern region of the United States. The university has an enrollment of close to 23,000 students, of whom 68% are undergraduates and 36% are graduate students. The Center employs approximately nine staff members, promotes eight instructional development programs, and serves tenured and non-tenured faculty members, instructors, graduate teaching assistants, graduate students, and postdocs each year. The mission of the Center is to implement teaching innovations and best practices and promote understanding of the learning process on the university campus.

Center E is situated in a public research university in the Midwest. It has an enrollment of over 60,000 students consisting of 89% undergraduates and 20% graduate students. The Center employs approximately nine staff members, promotes 13 instructional development programs and is available to anyone who teaches at the university, approximately 7,500 people; however, the Center actually serves about 2,000 individuals each year. The mission of the Center is to promote academic excellence on the university campus and contribute to local, national, and international conversations about faculty development.

Center F is situated in private liberal arts university in the South. The university has an enrollment of about 6,000 students, with 88% of them being undergraduates and 12% being graduate students. The Center employs approximately five staff members, promotes 10 instructional development programs, and serves approximately 700 faculty members and instructors each year. The mission of the Center is to promote excellence and innovation in teaching on the university campus.

Center G is situated in a public comprehensive university within the Eastern region of the United States. The university has an enrollment of about 20,000 students, consisting of 91% undergraduate and 9% graduate students. The Center employs approximately five staff members, promotes 15 instructional development programs, and serves approximately 1,700 instructional faculty members each year. The mission of Center is to assist all the university's teachers to excel.

Center H is situated in a private comprehensive university in the South. The university has an enrollment of approximately 13,000 students, almost evenly divided between undergraduate students (52%) and graduate students (49%). The Center employs approximately eight staff members, promotes eight instructional development programs, and serves

approximately 1,200 instructional faculty members each year. The mission of the Center is to promote understanding of the teaching and learning process, cultivate dialogue about teaching and learning, and create and disseminate best practices in teaching and learning.

Table 4.1 below, provides a snapshot of the demographics of each faculty development center, and data regarding the number of students attending the university, institution type, and the number of programs and staff supported by the centers.

Table 3.1 *Demographics of Faculty Development Centers*

Center	Years	Public or	Number of	Number of	Number of	Type
		Private	Programs	Students	Staff	
A	>30	Public	>10	>40,000	>20	Research
B	>30	Public	>20	>30,000	>20	Research
C	>10	Private	>10	>10,000	>15	Research
D	>20	Public	>8	>20,000	<10	Research
E	>20	Public	>9	>50,000	<10	Research
F	<15	Private	>15	>5,000	<5	Liberal Arts
G	<20	Public	>15	>20,000	<8	Comprehensive
H	>20	Private	>6	>13,000	<8	Comprehensive

Evaluators in this study are responsible for evaluating programs at their particular centers. However, because there are so few centers with budgets large enough to hire dedicated evaluators, six of the eight evaluators have additional faculty development responsibilities. The average length of time and experience among participants is eight years. Five of the eight evaluators indicated they were members of Professional and Organizational Development (POD), three were members of International Society for the Scholarship of Teaching and Learning (ISSOTL), two were members of American Educational Research Association (AERA) and one evaluator is a member of Association of American Colleges and Universities (AACU), Association for the Study of Higher Education (ASHE), American Psychological Association (APA) and American Evaluation Association (AEA).

Participants

Allison

Allison started working in faculty development part-time as a graduate student in 1999 and full-time since 2003. She has a Ph.D. and is a member of POD; in the past she has also been a member of AEA and ASHE and APA. By the time of this study, Allison had conducted “action-oriented” assessments with academic units. She has helped units think about types of data they might want to collect to inform their understanding of the student learning experience and helped them determine how they might want to make curricular changes as a result. Allison’s first step has been to review the goals of the program she is preparing to evaluate in order to determine which measures should be used for collecting data.

Ann

Ann evaluated faculty development programs sponsored by her center since 2003 with the aid of a research assistant helping her evaluate programs. She has a Ph.D. and is a member

of POD. At the time of the study, she had also evaluated grant-funded projects and was the principal investigator or co-principal investigator on projects related to teaching and learning. She had worked with faculty members to develop the evaluation plan for grant submission. She also helped faculty analyze data and works on projects that outside the center, as well as projects that involved collaboration with other faculty. Her approach to evaluation included: determining the goals of the program, deciding the evaluation is going to be formative or summative, determining the available resources and identifying the stakeholders and, after these are established, designing the data collection process.

Connie

Connie was the newest to the faculty development field of the participants, having evaluating faculty development programs for approximately two years prior to this study. She is a doctoral student and a member of POD, AERA and ISSOTL. She has evaluated consultations, workshops, institutes, orientation, and teaching certificate programs, and she had evaluated all of the programs at her center.

Dave

Dave has been in faculty development since 1999 and involved in evaluating faculty development programs for seven years prior to this study. He has a Ph.D. and is a member of POD, ISSOTL, and AACU. He has evaluated workshops, faculty development programs, and initiatives sponsored by the center. Dave recommended that evaluators of faculty development programs think about how the evaluation is going to be used, determine the goals of the evaluation and, after the data is collected, share the information and use the evaluation results to make decisions.

Erin

Erin started in faculty development nine years ago as a graduate student in 2003 and was later hired to evaluate programs in 2008; she has a Ph.D. and is a member of AERA. At the time of the study, she had evaluated faculty development programs, worked with faculty who are trying to obtain grants, helped academic units evaluate programs, directed evaluation of instructional learning, and assisted faculty members with formative and summative evaluations for teaching improvement.

Jennifer

Jennifer had been in faculty development for approximately eight years by the time this study was conducted. She worked for the center as graduate student while getting her PhD and was later hired to evaluate the programs at the center. She has a Ph.D. and is a member of POD. In addition to evaluating programs, orientations, consultations, and educational grants, she has directed several faculty development programs. When evaluating programs, Jennifer's process has been to determine the evaluation question of interest and how the evaluation is going to be used, define the scope of the evaluation project, compare benchmarking information, and then establish how the evaluation information will be shared with internal and external stakeholders.

Stephanie

At the time of the current study, Stephanie had been in faculty development for 20 years and working to help people understand program evaluation and assessment about ten years. In addition, she has conducted external evaluations for faculty development centers at other universities. She has a Ph.D. and is a member of POD. She has evaluated consultations, workshops, institutes, and intensives. In preparing to evaluate the programs at her Center,

Stephanie has determined the goals for undertaking the evaluation and considered the best approach to accessing evaluation data.

Tiffany

At the time the study was conducted, Tiffany had been in the faculty development field for 10 years, evaluating programs for the previous six years. She has a Ph.D. and is a member of POD. She has evaluated faculty development workshops, summer institutes, and symposia. Tiffany advised making the evaluation a part of the program planning process and, when evaluating her center's programs, she decided on the best data collection method for assessing impact before collecting data.

Sources of Data

Data collection in qualitative descriptive studies is typically directed toward discovering the “*who, what, and where* of experiences” (Sandelowski, 2000, p. 338); thus, data collection techniques can include moderately structured open-ended individual interviews and examination of documentation. To answer the three questions that guided this study, data were collected from telephone interviews. Informed consent was discussed with each participant before data collection occurred.

Interviews

A qualitative descriptive design was used for this study; therefore, interviews were an appropriate method for gathering the necessary data. Merriam (1992) explained “interviewing is necessary when we cannot observe behavior, feelings, or how people interpret the world around them” (p. 72). In addressing the primary purpose of the interview, Patton (1990) stated, “We interview people to find out from them those things we cannot observe... we cannot observe feelings, thoughts and intentions... we cannot observe how people have organized the world and

the meanings they attach to what goes on in the world” (p. 196). Those who live and experience program evaluation in faculty development are best suited to relate the process of evaluation in faculty development centers, and the best way for that type of data to be gathered is through an interview.

Once the participant agreed to be interviewed, he or she was emailed the informed consent form to be signed, scanned, and returned. A time for the telephone interview was established and, before that time, each participant was asked if there were any questions or concerns about the research or the informed consent form. They were then asked to mail the signed hard-copy informed consent form to: Thelma Woodard, College of Education, Health, & Human Sciences 1122 Volunteer Blvd., A503 Bailey Education Complex Knoxville, TN 37996. The telephone interviews (Appendix C) lasted approximately 30-40 minutes.

Kvale (1996) noted that there is a set of themes and points covered in a semi-structured interview that allows openness in the order of questions and opportunity for follow-up questions that provide in-depth answers. Interview questions can be categorized into five types (Merriam, 2009): experience/behavior, opinion/value, feeling, knowledge, and background/demographics. The following table depicts these five interview question types to show the variety of questions that were used in the semi-structured interview format utilized in this study (Table 3.1).

Table 3.2 *Research Questions Mapped with Interview Questions and Question Types*

Interview Questions	Question Type
RQ 1. What process do evaluators engage in when they evaluate various programs in faculty development?	
1. What has your experience in evaluation and assessment in faculty development been like?	Experience/Behavior
2. What types of programs do you evaluate or assess?	Knowledge
3. What needs to be done/what steps do you take when you are contemplating the evaluation of a program?	Experience/Behavior
10. If other centers wanted to start evaluating their programs, what advice would you give them?	Opinion/Value
RQ 2. To what extent and how do evaluators link evaluations to short, medium and long term outcomes?	
4. How do you evaluate the impact of the programs at your center?	Experience/Behavior
5. In what circumstances do you use logic models in your work?	Experience/Behavior
6. What are the advantages or disadvantages of documenting outcomes?	Opinion/Value
8. How familiar are you with the idea of program theory?	Experience/Behavior
9. How do you link your evaluations to the short, medium and long term outcomes of your centers goals and mission?	Experience/Behavior
RQ 3. To what extent and how do evaluators use evaluation models (from the literature) in faculty development?	
7. What evaluation or assessment models do you use in your work?	Experience/Behavior
9. How do you link your evaluations to the short, medium and long term outcomes of your centers goals and mission?	Experience/Behavior

Note: RQ = Research Question

Two faculty development experts at the University of Tennessee validated the interview questions and protocol. Interviews were tape-recorded, and recordings and transcriptions locked in a secure cabinet at the researcher's home. It is very important that each interview was tape recorded in order to accurately reflect interviewees' responses and to ensure an accurate transcription of the interview. Confidentiality was maintained by changing all descriptive information that might allow a third party to identify any of the participants.

Documents

The participants were asked to provide documents in order to supplement the insights arising from the interviews. It was posited that documents provided by evaluators would help identify how outcomes are documented and aid in answering each of the three research questions of this study. Evaluation plans and reports, instructional consultation documents, and survey questionnaires were requested during each interview. Three of the eight evaluators provided survey questionnaires; one evaluator referred to a written chapter about the center's evaluation process, two evaluators provided annual reports. No evaluation plans or reports were provided; therefore documents were not used in this study.

Data Analysis

Qualitative content analysis evaluated verbal and visual data, allowing for summarization of the informational contents of the data (Sandelowski, 2000). Interview transcripts were analyzed using the constant comparative method, defined by Merriam (1998) as the constant comparative method, beginning with a particular incident from an interview and comparing it with another incident in the same set of data or in another set.

Coding and Thematic Analysis

Ryan and Bernard (2003) provided examples of techniques that qualitative researchers may use when identifying themes. These techniques include identifying repetitions found in the data where “topics occur and reoccur or are recurring regularities” and identifying similarities and differences (p. 89). Ryan and Bernard (2003) referred to these techniques as the constant comparison method, which “involves searching for similarities and differences by making systematic comparisons across units of data” (Ryan & Bernard, 2003, p. 91).

First, the transcripts were read looking for responses relevant to the research questions. Notes and comments were written in the margins of the transcript pages; codes were generated from the data and then systematically applied content analysis was reflexive and interactive while the treatment of the data was continuously modified to accommodate the new data and insights about those data. After completing the coding of the data, the researcher read through and examined the transcripts to make sense of what was said by the participants as a group. Next, the codes were grouped using analytic coding, which refers to grouping coded data based on reflection and interpretation of meaning (Merriam, 2009). These groups were then compiled into a separate group with similar categories. The data was inventoried and organized using Microsoft Word for retrieval and manipulation.

In this study, these constant comparisons led to tentative categories that were then compared to one another. This process was followed and repeated for each interview until a general theme emerged. Each interview transcript was compared to the others identify various themes and create an understanding of the data. Creswell (2007) noted that a researcher must sort through all of the data, form categories of information, and attach codes to them. This

creates the basis for the themes and the assessment that the qualitative researcher is attempting to build.

Two additional techniques used in this study included “cutting and sorting” and “word lists and key words in context” (Ryan & Bernard, 2003, p. 94-96). Cutting and sorting involves “identifying quotes or expressions that seem somehow important and then arranging the quotes/expressions into piles of things that go together” (p. 94). The researcher cut out each quote, making sure to maintain some of the context in which it occurred, and pasted the material on a small index card. On the back of each card, the researcher wrote down the quotes reference, and then laid out the quotes randomly on the floor, sorted them into groups of similar quotes, and then named each group and the themes. Word lists and key words in context involve generating a list of words used by the participants in order to better understand their references and meaning.

Establishing Trustworthiness

Tracy (2010) established eight big-tent criteria for effective qualitative research: “high quality qualitative methodological research is marked by (a) worthy topic, (b) rich rigor, (c) sincerity, (d) credibility, (e) resonance, (f) significant contribution, (g) ethics, and (h) meaningful coherence” (Tracy, 2010, p. 837).

Worthy topic

The first key marker for establishing trustworthiness is that the research is relevant, timely, significant, or interesting (Tracy, 2010). This study offers insight into evaluation of faculty development programs. Evaluators participating in this study offered in-depth descriptions of their experiences with program evaluation, and these descriptions provided details of processes for other faculty development staff interested in evaluating their own

programs. This study also provides information about how program theory and evaluation models may be applied to the faculty development field. In addition, this study provides a foundation for future examination of factors related to program evaluation and its application to professional development. Lastly, the results of this study may aid in helping faculty development centers establish successful evaluation designs and implementation plans.

Rich Rigor

The second key marker for establishing trustworthiness is rich description and explanation (Tracy, 2010). Richness is generated through a “requisite variety” of theoretical constructs, data sources, contexts, and samples (p. 16). Requisite variety refers to the need for an instrument to be as complex, flexible, and multifaceted as the phenomena being studied. Applying the concept of requisite variety to qualitative rigor suggests that a researcher with “a head full of theories” and “a case full of abundant data” is best prepared to “see nuance and complexity” (Tracy, 2010, p. 16). Accordingly, a richly rigorous qualitative scholar is equipped to make the most appropriate choices about samples and contexts for studying specific issues.

Preparation for conducting this study took place within the environs of a land-grant level-three research university, evaluation, statistics and measurement program. Further, experiences as a research assistant with the Institute for Assessment and Evaluation and the Tennessee Teaching and Learning Center offered opportunities for application of education to evaluation of educational programs. Both education and professional experiences have contributed to the knowledge, skills and ability to conduct this study.

Rigor also provides face validity, which is whether a study appears, on its face, to be reasonable and appropriate. Researchers should implement due diligence, exercising appropriate time, effort, care, and thoroughness. Researchers should also make sure the content and sample

are appropriate for the goals of the study, and use appropriate procedures in terms of interviewing practices and procedures (Tracy, 2010). Two faculty development experts at the University of Tennessee validated the interview questions and protocol; this review added face validity and thoroughness appropriate for this study.

Rigor is also judged by the care and practice of data collection and analysis procedures. In terms of interviewing rigor in the course of this study, eight interviews lasting approximately 30-40 minutes each are appropriate given the sample, goals of this study, and types of questions asked during the interviews. The eight interview transcripts from this study provide a valuable contribution because the data intersects with rigorous line-by-line data analyses as detailed in this chapter. Readers are provided an explanation about the process by which the raw data were transformed and organized into this research. Rigorous analysis is marked by transparency regarding the process of sorting, choosing, and organizing data (Tracy, 2010).

Sincerity

The third marker of trustworthiness is marked by honesty and transparency about the researcher's biases and goals (Tracy, 2010). My biases stem from my experience working in the faculty development field. I have conducted assessments and evaluations of faculty development programs and attempted to build evaluation capacity during three years at the Tennessee Teaching and Learning Center at the University of Tennessee. I have helped staff at the Center learn evaluation terms and develop and use logic models for programs. I have also discussed their interest in evaluating their programs and how those evaluations have impacted the center's growth. In addition, I conducted a logic model, program theory workshop at the Professional and Organizational Development in Higher Education (POD) conference in 2010 and, based on

the lack of logic model knowledge and application of program theory in program evaluation, I determined the need for this exploratory and descriptive study.

Before planning this study, I reflected on and assessed my biases and motivations. My motivations are twofold: (a) to meet the requirements of my graduate program, and (b) to contribute to the program evaluation and faculty development field. Researchers should also practice self-reflexivity and be transparent about the research process (Tracy, 2010); thus, a physical research audit trail (Appendix D) provides clear documentation of all research decisions and activities.

Credibility

Qualitative research assumes that if two or more sources of data or types of data are collected, the conclusion is more credible because different data often yield different results (Tracy, 2010). In this study, the participants may have espoused very different values in interviews than the values they have enacted in contextual interactions; however, both sets of data are equally true. “Crystallization” is a term used in qualitative research to encourage researchers to gather multiple types of data in order to be open to understanding a problem in-depth rather than seeking to “validate singular truth” (p. 844). A member check, a process of obtaining regular feedback from another researcher (O’Donoghue & Punch, 2003), was used to enhance the trustworthiness of this study. The researcher employed regular member checks throughout this process with a researcher of faculty development to discuss potential biases as well as triangulation and situation of data.

Resonance

Transferability and generalizability in qualitative research refer to a study’s potential to have value across a variety of contexts or situations (Tracy, 2010). Both transferability and

generalizability are characteristics of resonance, the fifth key marker of trustworthiness.

Transferability is achieved when readers of the study believe the research overlaps with their own situation. Tracy (2010) noted, “qualitative researchers seek resonance not because they desire to generalize across cases, but rather because they aim to generalize within them” (p. 845). Generalization in qualitative research comes from “taking small instances and placing them within the larger frame” (p. 845). Accordingly, this qualitative study achieves resonance across various populations interested in program evaluation or faculty development even though the data is acquired from a unique population.

Significant Contribution

The sixth key marker is judged by the significance of the study’s contribution (Tracy, 2010). This study will extend the knowledge of program evaluation in the faculty development field, it will help improve evaluation and assessment practices of a variety of faculty development programs and, lastly, this study is significant because of its methodological design. According to Tracy (2010), significance is established “through engaging research methodology in a new, creative or insightful way” (p. 846). The few existing studies on evaluation in faculty development have been examined exclusively quantitatively, and this study is a significant contribution because it introduces and explicates a qualitative strategy for studying the problem. Knowledge generated through qualitative methods can still transfer and be useful in other settings, populations, or circumstances.

Ethics

The seventh key marker for establishing trustworthiness is procedural ethics. Procedural ethics are codified by the Institutional Review Board (IRB), including mandates such as do no harm, avoid deception, negotiate informed consent, and ensure privacy and

confidentiality. Procedural ethics also advise that research participants have a right to know the nature and potential consequences of the research and understand that their participation is voluntary. Such procedures not only attend to ethics, but also lead to more credible data (Tracy, 2010, p. 847).

This study followed University of Tennessee ethical guidelines in completing and gaining approval of the IRB board. All participants were informed of the purpose of the study, how their information would be used, as well as their rights in participating or not participating in this study. Participants were advised that a pseudonym would be used and that names would not be used in publications resulting from the study. They were informed that all recordings and related notes would be stored in a locked and secure file cabinet accessible only to the researcher. Consent forms (Appendix A) were explained and signed by all participants before data collection began.

Meaningful Coherence

The final key marker for establishing trustworthiness is that the study be coherent and accomplish its purpose and goals. The study should “hang well together,” the reviewed literature should “situate the findings,” and the findings should “attend to the state research questions and the conclusions and implications meaningfully interconnect with the literature and data presented” (Tracy, 2010, p. 848). To attend to meaningful coherence throughout this process, I employed regular reviews with the research overseer to connect the research design, data collection analysis, and findings.

Chapter Summary

Because the focus of this study is to understand and describe the evaluation processes and methods used by evaluators involved in faculty development, using a qualitative research method

allows for exploration into those perceptions. In addition, it is uncommon to find large numbers of program evaluator in faculty development. For this reason, criterion and snowball sampling was utilized to identify and invite evaluators to participate in this study.

Interviewing evaluators who live and experience program evaluation in faculty development is a method used in this study to assist in understanding the process of how evaluation takes place. Documentation of program evaluation would also assist in understanding the evaluation process and were requested from the participants; however, no evaluation plans or reports were provided. To aid in designing a high quality research study, Tracy's (2010) eight big-tent criteria for excellent research was detailed and provided the framework for establishing trustworthiness. In summary, the purpose of this chapter was to describe the study design and rationale for the use of a descriptive qualitative study. Next, Chapter 4 provides the findings related to the research questions of this study.

Chapter 4

Findings

Introduction

The primary focus of this study was to explore faculty development center evaluators' perceptions of the processes they engage in when evaluating faculty development programs; their use of program theory; and the evaluation models they use to design faculty development program evaluations. Three questions were used to guide the semi-structured interviews: (a) what processes do evaluators engage in when they evaluate various programs in faculty development? (b) To what extent and how do evaluators link evaluations to short-, medium-, and long-term outcomes? (c) To what extent and how do evaluators use the standard evaluation models in faculty development?

Five primary themes and one secondary theme emerged from the data analysis of information offered by each participant. These themes reflect consistent elements and patterns in participants' comments on programs they select to evaluate, methods used in program evaluation, and examples of how evaluation results can be used. Emergent themes are developed and described in detail below.

Findings are described and grouped by research questions in alignment with the purposes of this study, which are to explore the evaluation processes, theories, and models used in faculty development. Then, the resulting themes are presented. Re-presentation of data and evidence of that theme is provided in table format. Re-presentation of data means that the raw data from the interview transcripts are "re-presented" within the context of the research questions (Sandelowski, 2000). Evidence tables depict quotes excerpted verbatim from the pages of transcribed text as evidence for the theme (Bazeley, 2009). A quote from each of the participants

powerfully illustrates evidence for each theme, and conveys frequency of occurrence of the theme. In addition, participants' quoted response provide a clear chain of evidence for discussions concerning implications, research, and practice undertaken in Chapter 5.

Research Questions

In accordance with the research questions, the data from participants report on their evaluation process, use of program theory, and evaluation models are re-presented here in the findings. In qualitative research, themes originate from the participants, programs, investigation, literature, or interpretation (Constas, 1992). The themes for the first research questions evolved from the evaluation process (Posavac & Carey, 2007; Posavac, 2011) found in the literature a priori. A priori describes a theme is established from the literature before data collection (Constas, 1992). Posavac's (2011) evaluation process provided categories for possible theme identification to answer Research Question 1.

In addition, themes originated from the participants in this study; the comparative method was used to identify these themes (Ryan & Bernard, 2003). The comparative method included taking pairs of comments reported by different participants and asking, "How is this comment different or similar to the other participant comments?" If a particular theme was present in both comments, the next question was, "Is there any difference in the way the theme was articulated in both comments?" These comparisons provide a method for analyzing the data and identification of themes. The research process included particular care in describing how evaluators talked about the theme and in showing the context in which the participants discussed it.

Overall, the primary themes that emerged are: (a) purposeful selection of programs to evaluate, (b) careful planning of the evaluation methodology, (c) understanding that it takes

considerable time to evaluate programs, (d) strategically linking program outcomes to center goals, and (e) implicit application of a utilization-focused approach to planning evaluation.

Themes in this study are labeled *primary* if all eight participants commented on the topic and *secondary* if at least six participants commented on the topic. The secondary theme that emerged is (a) thinking critically about how to evaluate programs.

Research Question 1: What process do evaluators engage in when they evaluate various programs in faculty development?

Posavac (2011) identified eight steps in the evaluation planning process: (a) review evaluation models, (b) identify the program to evaluate, (c) identify relevant stakeholders, (d) identify information needs, (e) plan the evaluation, (f) examine the literature, (g) plan the methodology, and (h) write the evaluation plan. The evaluators who participated in this study do not follow Posavac's (2011) evaluation planning process exactly. One evaluator typically followed six of the eight steps, one evaluator followed five, three evaluators followed four, one evaluator followed three, and two evaluators followed two of the steps. All of the evaluators identified the program to evaluate and planned the methodology. Table 4.1 below, depicts the steps followed in their evaluation planning process and provides the evidence for the two primary themes *Purposeful Selection Of Programs To Evaluate*, and *Careful Planning Of The Evaluation Methodology*.

Table 4.1 *Selection of Programs and Planning of the Evaluation: Evaluators Following Posavac (2011) Evaluation Planning Process*

	Allison	Ann	Connie	Dave	Erin	Jenn	Steph	Tiffany
Select evaluation model								
Identify program	X	X	X	X	X	X	X	X
Identify stakeholders		X		X				
Identify information needs		X	X	X			X	X
Plan evaluation		X	X	X			X	X
Examine literature	X	X						
Plan methodology	X	X	X	X	X	X	X	X
Write evaluation plan								

Primary Theme 1: *Purposeful Selection of Programs to Evaluate*

Selecting a program to evaluate is a part of the process that all these evaluators engaged in when they evaluated programs in faculty development. Eight out of the eight evaluators interviewed discussed the selection of a program or programs as being a part of their evaluation process. Selected programs were primarily dependent upon the responsibilities of the evaluator. Some of these evaluators only evaluated center programs that included workshops, consultations, institutes, and other short-term programs sponsored by the center. Other evaluators have written evaluation plans for grant proposals, have been responsible for running student evaluation systems, and have worked with academic units to create psychometric assessments in addition to evaluating center programming.

Each of the evaluators described their evaluation process differently depending primarily on the particular type of programs they have evaluated. If the evaluator reported working with educational or foundation grants, he or she spoke about the goals of the program and utilization of the evaluation results. The participants also spoke about the evaluation process and stakeholder identification, in addition to the selection of a program to evaluate. For example, because Erin evaluated Department of Education and Gates Foundation grants as well as "...the academic center grants or [grants from] one of the technology units" that are within her university, she mentioned selecting those specific programs for evaluation. Likewise, the programs Ann selected to evaluate were based on her role as an evaluator for programs sponsored outside of her center. Ann stated:

Basically I work on some projects that are grant funded that I'm the PI on or co-PI, and these are projects related to teaching and learning. I also work with faculty who are putting in NIH training

grants for graduate students, and I'm working with them on developing their evaluation plan. I feel like I'm on 100 grants and proposals because they want to put down a name and qualifications so that the proposal will get funded.

Those evaluators who were not responsible for evaluating grants for academic units talked in terms of selecting a major program in addition to evaluating all of the workshops conducted by the center. Programs discussed by the evaluators included workshops, consultations, and semester programs. Ann also mentioned selecting a yearlong faculty development program for junior pre-tenure faculty for evaluation. She described how she used the *Approaches to Teaching Inventory* to measure whether teachers were changing from somewhat teacher-focused approaches to more student-centered approaches.

Along with Ann, Jennifer and Tiffany were also evaluators responsible for directing and evaluating their programs. Jennifer described the program she selected to evaluate:

One program is our professional development program for graduate students and, when we began it, we started with pre- and post-testing. It's perceptual in terms of the grad students' level of preparedness to meet what we designed the program to meet. Then we followed up on that working on faculty at our School of Education to look at faculty positions and postdoc positions to see how well, again it's perceptual, but how well they perceived the program either preparing them for the job market or preparing them for their position.

Tiffany said:

We've really put a lot of work into our workshop series and defining what we expect people to get out of that, so we're thinking more sort of longitudinally about assessing the impact of workshops, especially since it involves a lot of human capital to put together a workshop series—a good one. I want to make sure we're really getting something out of it.

I'm trying to look at where either we spend a lot of human capital, or spend a lot of dollars, so we spent a fair amount of money on faculty orientation. We have a weeklong series of events called May symposium, so we take a big picture look at that, you know, how does this event feel; what does the campus think about it; is it servicing people's needs; do we need to make a shift in sort of the big boat that's kind of floating along.

Stephanie, Dave and Connie spoke about evaluating all of the programs offered by their centers. They have not been involved in evaluating academic units or university grant programs. Stephanie described the programs as follows:

We evaluate all of our individual consultations that we do with the faculty and TAs. We evaluate all of our programs, which include workshops. It also includes more long-term programs; like we have an Institute we do regularly, probably about every six or seven times a year, so that is extensive. We have other programs that are also long; the longer that we do assessment of we also have a year-

long learning community that we do assessment of a variety of programs.

Connie stated that she evaluates all of her centers' consultation services, workshops and events. She said, "I evaluate our junior faculty program, which includes the fellows program that is a yearlong commitment, so we look at the participation in that, as well as the feedback."

Connie also expressed her desire to move past the mere collection of participation data from the database or satisfaction surveys. She wants to look at trends and evaluate impact:

Well, it's kind of two-sided to me, so first is the participation numbers, and then, for instance, like the teaching certificate program, it's been going on for years. Those participation numbers are more comprehensive because we're looking at people perhaps over a couple of years, where in shorter programs or events or consultations I'm just looking at a semester, for instance; but either way, I'm gathering up participation data. Then, the opposite of that is brainstorming about how we can get the best feedback we can for our different programs, and so in the past we primarily relied on surveys using survey monkey and that kind of thing. We revised our strategy this past fall and tried to consolidate. We changed the questions that we were even asking and are now hoping to launch several focus groups during the summer and in the fall to get a better look at the long-term impact that participating in one of our program has had on an individual's teaching.

Dave also noted that he evaluates all of his centers' programs. However, he did not mention efforts to determine impact. His focus was on evaluating programs for feedback:

What we do [is] sort of evaluate individual programs routinely, for example, a workshop. We do a lot of small group work, so when the small group ends at the end of the semester, we gather feedback about what was helpful about it [and] what could be changed so that it could be more helpful: [We ask], Did it help you meet your goals? Those are basically the kinds of questions we tend to ask, and then my center does targeted deep evaluation of one program, initiative, or issue each year.

Primary Theme 2: *Careful Planning of the Evaluation Methodology*

In addition to selecting a program to evaluate, these evaluators talked about planning methodologies. Plan evaluation methodology is another of the steps in Posavacs' (2011) eight-step evaluation process. Six of the eight evaluators interviewed commented that they were planning methods to provide a more in-depth look at how faculty members use their services, what the outcomes are, and the impact on teaching. These evaluators have been using qualitative methods—annual surveys, validated measures, and planning evaluation strategies—that help evaluate impact.

Although evaluators in this study talked about various evaluation strategies and methods, they reported that collecting participation data was a common method of evaluating their programs. All of the evaluators reported looking at self-reports from faculty attending one of their programs. They also reported trying to design and implement methods that assess the outcomes and impact of their programs. For example, Connie digitized participant self-report

data store on paper going back to 1986 in order to look at overtime, to “make more sense about the participation data,” and to determine trends of effectiveness and impact. When talking about designing methods to assess outcomes, Connie said “it’s been a moving target” in some ways:

When I joined the center, I took a look at what assessment had been going on before I began here, and primarily we were working with just participation data, you know, how many consultations did we do? How many times does someone come to an event? Did we have overlap in those numbers? Since I’ve been here we’ve tried to get a more comprehensive look at ways we can assess the work that we do beyond just participation numbers. So, in the past we primarily relied on surveys using survey monkey and that kind of thing. We revised our strategy this past fall and tried to consolidate. We changed the questions that we were even asking and are hoping to launch several focus groups during the summer and in the fall to get a better look at the long-term impact participating in one of our program has had on an individual’s teaching.

All of the evaluators emphasized the importance of advanced planning of evaluation methods and data collection. Jennifer commented that “informally we are talking about doing a center-wide self-study, which we haven’t done for about ten years, but [that] is still very much in the planning stages, but we haven’t done anything with that.” The most mentioned component of planning evaluations was the process of thinking about evaluation methods and data collection. Jennifer indicated that being strategic when planning an evaluation is important:

I found it really useful to think about the program input, what my goals were...in terms of thinking about what are the inputs, what are the outcomes, what are our big goals?...I was having a conversation with a colleague the other day who has been running a program for about eight years and finds her feedback really consistent... I know it's working and it's working well, and the program is meeting the outcomes and, in terms of time investment, are there better things to assess now?

In general these evaluators have spent time thinking about what and how to evaluate their programs: What evaluation questions to ask, how they can get the best feedback, what the goals of the evaluation are, and what are the interests of stakeholders? For instance, Ann stated that an important part of planning includes determining whether the data will be meaningful:

I guess I usually think about what the goals are of the program so I can outline the evaluation with the goals, and then I think about whether the evaluation is going to be a formative evaluation or summative evaluation, and then I also think about the resources that are available because that will affect the scope of what can be done, and then I might think about the stakeholders and who they are, what outcomes and what data are going to be meaningful to them so that I can collect data that will be meaningful and then determine how much will be qualitative and how much will be quantitative and things like that. But I think two key points are always making sure the goals or the objectives are firmly

established because sometimes people have programs and they actually have objectives but they haven't articulated them, and it's really hard to plan a good evaluation if you don't know what those things are.

Allison also mentioned having limited resources and simply focusing the evaluation project on what is important to the evaluator and the center:

If there are limited resources, which actually there always are, think about the need for the evaluation and what types of data are most important to collect given that need. For example, a new faculty developer may not want to evaluate all of the programs at once but perhaps choose one that is strategically important, and that might be strategically important because there are a lot of resources that go into it or because it aligns with the university's strategic plan...and it is often very important to think about how it gets shared and how it might inform curricular change.

Tiffany, Stephanie, and Dave also described planning the methods for assessing impact. Tiffany spoke explicitly about making evaluation planning necessary in the process; she explained that the center staff shifted thinking so they could be mindful about program evaluation. She recommended that evaluators think about making evaluation a part of the program planning process and decide the best data collection method for assessing impact:

We collect the kind of information that will really help us either make the program better or decide whether or not a program is worthwhile. And I think that early on in my career we were sort

are just happy to have programs, and it would be nice to have some data to let us know how those programs are going, but now what I think we do is, as we are developing a program, we're actually mindful of the kind of evidence that we would like to be able to collect to measure the impact or the efficacy of the program and, so rather than assessment as the last thing we do, I think we're starting to move that into one of the first things we do. It's part of the planning process, and when we start shifting our process that way we can be a lot more mindful—let's put that on our radar, let's make sure that happens as opposed to sort of the of the way of working where you're just happy to get things done and then you're onto the next thing. Then three or four years down the road you realize, huh, I haven't really collected any assessment data. So for us, I think it's really forcing us to put assessment on the table during the planning process.

Regarding thinking about the planning methods, Stephanie noted that “through a number of different measures, [she looks at] satisfaction surveys but it doesn't really get an impact so [she does] an annual survey [and] a random sample of a quarter of the people who used [them] over the year.” Stephanie also had an interesting take on aligning data collection with evaluation questions:

Like with any assessment, I think to really figure out what questions we really have about it, I mean what our goals are in doing the assessment and what we're hoping to find out so that

defining our goals is probably the first big step and then finding out different ways to actually be able to get those data. We do that a number of ways, both through asking participants immediately after but also retrospectively so we do an annual survey of people who participated in program so that we can see impact has been beyond when they leave the room. We also do it through a lot of simple, I hate to call it “bean counting,” but we have a somewhat sophisticated database for keeping track of everything. The main question is, what do we really need to find out and what are the different ways we can get the data that we need to be able to answer those questions?

Dave stated that his center focuses on participation data and uses the data for feedback on how their center is affecting the campus:

We may develop a qualitative survey rather than just a check “yes/no” or short answer sort just thinking about who are we going to give it to [and] how are we going to help them do it so that we can get a reasonable response rate and reasonable information and then gathering it up and then analyzing it and sharing it and using it to make decisions.

Dave went on to say that time is a factor in his approach to planning and evaluation:

My experience first was the realization that perhaps we should evaluate our work, but it took me a few years to get there beyond

sort of what a colleague used to call smile sheets, you know the end of workshop evaluations you know you flip through as you walk back to your office and pretty much be done with. So, thinking about systematic evaluation, it's hard and it takes time that I don't have much of *and* it's very important.

Most of the evaluators expressed concern about the amount of time it takes to plan an evaluation that may provide meaningful information about the impact their programs are having on their respective campuses. For instance, Connie mentioned that determining the impacts the center has and the evaluation process has been a “moving target.” She went on to say:

It is a struggle. I think that it's difficult sometimes to have people give you meaningful impacts immediately after they've attended a service or attended an event or came through a program. Like sometimes those impacts take a couple of years, you know, to happen, and so it's a struggle of either keeping in contact with those people, or then sometimes they're like later on, “Oh, I learned that from you all? I didn't know that. I remember picking it up from somewhere. Oh yeah, it was that workshop.” And so it's difficult to figure out the best way to do that [evaluate and measure impact] and we haven't yet.

Tiffany also spoke about how time-consuming it is to link evaluation planning to their mission and to determine impact:

I think it's sort of mapping, what we have to look at here is our mission and our goals and our outcomes, and here are all the programs, and here's how they link to all of the aspects of our mission, and here is the assessment data that we have but right now. It's all very disconnected. I mean, I think I know; I mean I can look at it myself and I can see all of the connections because I'm immersed in it, but I'm not so sure anyone else will be able to see all of those things. But again, that's a really time-consuming process. But we've realized over the past year, you know, we had some external consultants come in and do some work with us, and they, of course, highlighted that this is an area where we need to work on.

Regarding an answer to Research Question 1, three primary themes and one secondary theme emerged during data analysis. The content analysis revealed a pattern of all participants reporting purposefulness in selecting programs. The selected programs were grant programs, and long- and short-term programs. Those patterns make "purposeful selection of programs to evaluate" a primary theme of the faculty development program evaluation experience. These evaluators also reported wanting to move away from just collecting participation data. The content analysis revealed a pattern of all participants reporting purposefulness in planning methods for assessing impact of programs rather than relying on descriptive data. Those patterns make "careful planning of the evaluation methodology" a primary theme of the faculty development program evaluation experience.

Primary Theme 3: *Understanding That It Takes Considerable Time To Evaluate Programs*

Evaluators also reported that it takes considerable “time” to plan and evaluate their programs effectively. Table 4.2 below, provides the results of the content analysis that revealed a pattern of participants reporting “time” as a major challenge in program evaluation; many also characterized the challenge as a disadvantage of conducting program evaluation. Those patterns make *Understanding That It Takes Considerable Time To Evaluate Programs* a primary theme of the faculty development program evaluation experience.

Table 4.2 *Understanding That It Takes Considerable Time To Evaluate Programs*

Understanding That It Takes Considerable Time to Evaluate Programs		Quote from Participant
Erin	It's not brain surgery but it's time-consuming	
Stephanie	The challenge of any kind of assessment is the time it takes to do it	
Connie	Sometimes those impacts take a couple of years you know to happen	
Ann	The downside of evaluation is it takes time and money	
Tiffany	That's a really time-consuming process	
Allison	There are a lot of resources and time that go into it	
Jennifer	The program is meeting the outcomes and in terms of time investment are there better things to assess now	
Dave	It's hard and it takes time that I don't have much of and very important	

Secondary Theme: *Thinking Critically About How To Evaluate Programs*

When speaking about the process evaluators engage in when they evaluate various programs in faculty development, these evaluators also reported “thinking” critically about how to evaluate their programs. Table 4.3 below, provides the results of the content analysis that revealed a pattern of participants reporting thinking about evaluation questions, feedback, goals of the evaluation, and stakeholders. Those patterns make “thinking critically about how to evaluate programs” a secondary theme of the faculty development program evaluation experience.

Table 4.3 *Thinking Critically About How to Evaluate Programs*

Thinking Critically	
About How to	Quote from Participant
Evaluate Programs	
Stephanie	I think to really figure out what questions we really have about it I mean what our goals are in doing the assessment
Connie	Brainstorming how we can get the best feedback that we can for our different programs
Ann	I usually think about what the goals are... I also think about the resources that are available
Tiffany	We can be a lot more mindful, let's put that on our radar, let's make sure that happens
Allison	Think about the need for evaluation and what types of data are most important
Jennifer	In terms of thinking about, you know what are the inputs? What are the outcomes? What are our big goals? We spend a lot of time in the center
Dave	Just thinking about who are we going to give it to

Research Question 2: To what extent and how do evaluators link evaluations to short-, medium-, and long-term outcomes?

To link program processes to short-, medium-, and long-term outcomes, program evaluators use logic models and program theory (Hacsi, 2000; Knowlton & Phillips, 2009). Both help identify underlying assumptions about how programs are expected to work, aid in evaluation planning, identify stakeholders, and provide pictorial representation of the program. If program theory is used, it can also help develop propositions about ideal outcomes, and those propositions can be compared to the theoretical model used in the evaluation process (Chen, 1990).

Primary Theme 4: Strategically Linking Program Outcomes To Center Goals

Although most of the evaluators participating in this study were not familiar with program theory, and only one reported using program theory in her work, they all reported that they were seeking ways to link outcomes to the center goals and mission. Of the eight evaluators, only Allison was familiar with program theory but stated that she does not use it formally in evaluation projects at her center and explained, “We don’t use it in a formal way. We often use it in our own work and with our clients in terms of thinking through conceptual basics about the theory that informs an activity or when thinking through the conceptual basis of an intervention; I think [that is] the way we informally use program theory.”

During the interview, Connie indicated unfamiliarity with program theory; however, while doing the interview she looked up the term, read the description, and said, “I’m reading the description of that now. Yeah, it looks like that is exactly what we do, we just don’t call it that. Yes, I’m looking at the logic models, and you know how it is sometimes in offices, you do what

you do and you don't always use the appropriate term; you know, we have our own words for what we do."

The other six evaluators were unfamiliar with the term *program theory*. However, although they had not used program theory, they were evaluating impact by using alternative data collection methods such as focus groups, interviews, and random sample surveys. Dave commented that he tries to link [the center] programs to our mission, so if we are assessing our programs, we're in some ways assessing if it fits the mission." Jennifer commented:

[We are] looking at patterns over time. We do a lot of qualitative end-of-program reports where people are writing answers to two or three questions about their experience with the program, and we look for patterns and trends in those; [however,] we haven't done as much work on how to gather the information that we get about different programs at the center level in terms of short-, medium-, and long-term goals. We are trying to do more of that.

Tiffany provided an alternative evaluation approach:

We do an annual survey, a random sample of a quarter of the people who used us over the year on just any general services, so those questions really do get at what did you get from our programs, how have you used them, and try to measure like how is it actually getting into the classroom we each year; [however,] you know, that's where I really think we're missing some real work right now. I think we know we need to do that, but I'm not so sure

we're being as overt as we can. This is an area where we need to work on."

Regarding the use of logic models, five of the eight evaluators said they were familiar with logic models. Typically, if the center received foundation money, or the evaluator worked on grant applications with academic units, the participants were familiar and had utilized logic models. Erin said logic models are not used because people are often confused by some pieces of the logic model, saying that, "people often get confused by the difference between outputs and outcomes." Jennifer indicated that she uses a modified logic model for faculty development programs:

I was having a conversation with a colleague the other day that has been running a program for about eight years and finds her feedback really consistent... I know it's working and it's working well, and the program is meeting the outcomes and, in terms of time investment, are there better things to assess now? I found it really useful to think about the program input, what my goals were, [although] I have not used it exclusively in other programs but implicitly in terms of thinking about what are the inputs, what are the outcomes, what are our big goals. We tend to take a backwards design approach to things, but I think it's similar to a logic model in terms of thinking out or thinking through what our goals are, and how do we design programs to meet those goals, and how we assess along the way what we're doing, but they are not logic models per se.

Two of the eight evaluators were not at all familiar with logic models and how they are used. In order to link program processes to program outcomes, these evaluators have been conducting follow-up surveys, interviewing program participants, and selecting to focus evaluation on one program at a time. Stephanie talked about assessing program outcomes by surveying participants who had used consultation services overall several years:

So, the last few years we're targeting specifically consultations because we wanted to see what the impact of that was, so we did a second big survey that went out to people that use consultations that had more specific questions for that...the challenge with measuring this, which is the challenge of any kind of assessment, is the time it takes to do it, and most centers are overworked as it is and adding that level of assessment on can seem overwhelming to people.

Connie concluded that her center has not done a very good job of assessing outcomes and stated that evaluators are "trying to go beyond just the survey, but it has been a difficult to assess impact" and link outcomes to center goals. However, she remarked that they are moving in the right direction:

Here we go through a fairly lengthy and thorough process of all of our programs and all of our services, so those are divided up and assigned to different staff members here at our center so everyone doesn't do everything. Some people are just tasked with four or five different things. We assess against what we laid out in a vision document and we compare, well, did we reach the number

of individuals that we hoped to. Does it appear as if we are kind of reaching the goals that we intended with this program? ...the positive outcome would be we would know what types of work to continue doing...it is a struggle. I think that it's difficult sometimes to have people give you meaningful impacts immediately after they've attended a service or attended an event or [come] through a program, sometimes those impacts take a couple of years, you know, to happen.

Tiffany reported that she can assess outcomes by sight, but she was not sure how to be systematic about mapping outcomes for others to see:

I think it's sort of mapping we just have to look at, here is our mission and our goals and our outcomes, and here are all the programs, and here's how they link to all of the aspects of our mission, and here is the assessment data that we have, but right now it's all very disconnected. I can look at it myself, and I can see all of the connections because I'm immersed in it, but I'm not so sure anyone else will be able to see all of those things. But again, that's a really time-consuming process, but we've realized over the past year—you know, we had some external consultants come in and do some work with us, and they, of course, highlighted that this is an area where we need to work on. So, I think what we're going to do is just use basic assessment strategies—mapping—you know, kind of like content mapping,

mapping your curriculum, you know, we're sort of rolling out a curriculum we can look at ourselves, like an academic unit: here are our courses, here are the outcomes that we have for our graduates (so to speak, our faculty), and so...where does the data—where does it all align, and where are we missing connecting? That's what we're keen to find out, you know, where [are] the big gaps in what we're doing and let it be driven by the mission as opposed to being driven by the programs. You know that's where I really think we're missing some real work right now. I think we know we need to [link to long-term outcomes], but I'm not so sure we're being as overt as we can, and so in fact this summer I'm going to have a graduate student come in from the assessment program to work with us on that exact thing.

Although program theory is not used to link outcomes to center goals, four of the eight evaluators said that they link outcomes to center goals in various ways. Allison said, “We do link [short medium and long-term outcomes] to the mission in the sense that we support evidence-based teaching and learning practices and, so given that, it is important for our work to be evidence-based.” In addition, Ann commented, “I think where we can we try and make a link to long-term outcomes, especially when we're looking at programs and we've got some money and resources to fund an evaluation.”

Whereas Allison talked about supporting evidence-based practices, and Ann talked about funding outcomes based evaluations, Stephanie and Dave spoke in terms of aligning outcomes with the mission of the center. Stephanie said:

We try to define the outcomes we measure, having those be guided by our goals and mission...we have, as part of our mission, changing the culture for teaching and learning at the university, which helps us define what outcomes we're even measuring, so that is directly linked to that mission.

Dave stated that the mission of his center is very similar to that of Stephanie's description of the mission at her center, and he explained:

When we do the big assessment project each year, usually what we're trying to do is evaluate something that is central to our mission. Sometimes it's something new, or sometimes is something we're not sure is worth the effort; maybe it's because it seems over time to be a little more peripheral to our mission, so we try to link all what we do back to our mission. The other thing that I haven't said yet that I meant to say earlier is we try to keep, I don't know what the right word is, we try to keep in mind that direct outcomes from our program are how faculty teach and how they think about teaching and learning, and so we don't spend a lot of time trying to evaluate student learning that is produced from our programs. That seems like a great goal, it just seems a step too far for what we are evaluating, and we have some degree of hope that our faculty are changing in ways that are evidence-based, that we can assume that evidence holds and, if faculty are using some techniques that help students learn more, odds are the students are

learning more, rather than having to trail all the way to student learning to determine whether there [are] outcomes from our programs.

Stephanie and Dave spoke in terms of aligning outcomes with the mission of the center. Jennifer and Erin, however, did not talk about their centers' mission, but instead about linking outcomes to center goals. Erin noted:

An easy term to use would be from a marketing or PR perspective, but it's not marketing in the sense that it is a consumer-driven model, more that it is...educating the people across campus of what my goals are and the goals of the office of instructional development, and the goals of the college, which [are] truly to help enhance instruction and support faculty, and is not being based on some rubric of what should be standardized milestones for each department, or each course, or each discipline, but rather helping them define what they want their students to learn, and how they want to teach it. So it's—on a good day it's really fun and exciting, and on a bad day it feels like I'm knocking my head against the wall. But I do want to say I enjoy my job.

Jennifer also talked about linking outcomes to center goals, but indicated the focus is on assessment rather than evaluation due to the programmatic structure of her center:

We looked at it in the past as, I think, to evaluate. We do a lot of assessment of individual programs, including workshop series that function as a program and, we have, I think, thought of those in the

past as things that coexist. We haven't done as much work on how to gather the information that we get about different programs at the center level about short-, medium-, and long-term goals. We are trying to do more of that. To this point we've used it largely, I think, to shape the development of individual programs, both the short- and the long-term.

Only two evaluators reported being familiar with program theory, and none of the evaluators reported using logic models as a way to link short-, medium-, and long-term outcomes. As a result, the theme "linking outcomes" emerged. All of the evaluators suggested that they seek to link outcomes to inputs, but did not know to call that effort "program theory." Table 4.4 below provides the comments made by each of the participants re-presented how evaluations can be used, and these patterns make "strategic linking of program outcomes to center goals" a primary theme of the faculty development program evaluation experience, providing evidence for the theme *Strategically Linking Program Outcomes to Center Goals*.

Table 4.4 *Strategically Linking Program Outcomes to Center Goals*

Theme:	Quote from Participant
Strategically Linking Program Outcomes to Center Goals	
Stephanie	We try to define the outcomes we measure having those be guided by our goals and mission
Connie	Does it appear as if we are kind of reaching the goals that we intended with this program?
Ann	Where we can we try and make a link to long-term outcomes
Tiffany	We know we need to do that [link to long-term outcomes]
Allison	We do link it [short medium and long-term outcomes] to the mission
Erin	Educating the people across campus of what my goals are and the goals of the office of instructional development, and the goals of the college which is truly to help enhance instruction and support faculty
Jennifer	Haven't done as much work on how to gather the information that we get about different programs at the center level about short medium and long-term goals
Dave	Usually what we're trying to do is evaluate something that is central to our mission

Research Question 3: To what extent and how do evaluators use the standard evaluation models in faculty development?

Seven of the eight evaluators indicated that they did not use an evaluation model when evaluating faculty development programs. Allison was the only evaluator in this study who used evaluation models in her work. She had used both Kirkpatrick's evaluation model and the utilization-focused evaluation model to evaluate programs at her center. She described her knowledge of and experience with these evaluation models:

To be honest, I wasn't formally trained in the evaluation models. I was formally trained in my academic discipline and learned about evaluation models on the job, but probably the one that we use is a modified version of Kirkpatrick's model. It's modified to [help] think about participation in addition to the other levels of evaluation. The second model that informs our evaluation practice is Patton's utilization focused approach. We really hope to think about the utilization of the evaluation findings or the assessment findings rather than just have them sit in a filing cabinet or a pile on the floor...I can see a lot of advantages, so some of the advantages would be helping to communicate the value of our work to internal audiences for example to others at the university. A second advantage would be helping to document what's effective for the faculty development community nationally and internationally, and then a third key purpose is to help us continually improve our service work to the university.

Primary Theme 5: *Implicit Application Of A Utilization-Focused Approach To Planning Evaluation*

The other seven evaluators in this study reiterate Allison's comments on the utilization focused evaluation model. Although seven of the eight evaluators indicated that they did not use an evaluation model, they all reported having a utilization-focused approach for conducting evaluation. For instance, Erin said that her evaluation reports provide "a feedback mechanism":

Have in place a feedback mechanism for whatever it is you learn from your research and your data collecting. Don't just be data miners, you know, where you just going in mining anything you can get out of any data but actually find a way that you can get faculty administration [or] whoever to actually do something and inform their practices based on the data.

Dave stated that program evaluation is important because it can be used "to improve what we're doing":

I think we first think about what kind of information do we need, so is this for our own purposes, to improve what we're doing—is it to report to someone else about something?—so first we have to determine that because I do not want to do an assessment or evaluation or what anyone call it unless I know why we're doing it. Then we think about what the goals are for this thing we're evaluating, and then we try to develop an evaluation system process, or whatever, that's going to give us relevant evidence

about things, and then we have to do some of the grunt work of—you know—getting together the information...who's this information going to go to? What kinds of feedback are we going to get when we do our deeper evaluations of a single program? We typically don't—it's typically not just a survey; it may be some sort of interview, or we may develop a qualitative survey rather than just a "check yes/no" or short answer...[Just] thinking about who are we going to give it to, how are we going to help them do it, so that we can get a reasonable response rate and reasonable information, and then gathering it up, and then analyzing it and sharing it and using it to make decisions.

Stephanie said that evaluations "can help us track impact that is spread through the university and in the departments, so one of the things we look at is, you know, we can start seeing whether there is clearly some word-of-mouth going on. That's another way we're looking to see kind of impacts within the culture of the institution." She added:

We do it through a number of different measures. We do it through the standard that everybody does to seek immediate satisfaction surveys, but it doesn't really get an impact, so we do an annual survey of...we do a random sample of a quarter of the people who used us over the year on just any general services, so those questions really do get at, What did you get from our programs, how have you used them? And, try to measure like how is it

actually getting into the classroom. We each year then also target a specific service, so the last few years we're targeting specifically consultations because we wanted to see what the impact of that was. So, we did a second big survey that went out to people that use consultations, that one had more specific questions. One of the things that's been useful for what I was calling the "bean counting" earlier, which people tend to be dismissive of, but they can help us track impact that is spread through the university in the departments...one of the things we look at is, you know, we can start seeing whether there is clearly some word-of-mouth going on because suddenly we're getting a lot of people from a certain department coming to things, and that's another way we're looking to see kind of impacts within the culture of the institution

Connie said that the evaluations at her Center were used "to make what we do more clear to the teaching community here at the university." Ann said evaluation reporting could help a faculty development center thrive:

I think if you're looking at sustainability of programs, if you don't have data you can't make a case for programs to continue being funded or someone can't go to bat for you. For example, my office comes under the office of the Provost, and the assistant Provost has been able to go to bat for us several times for sustaining programs because we've been able to give them really good data.

Tiffany stated that evaluations aid in “validating the role of the center” and commented further on the nature of center survival:

On a campus like this, where there are limited resources, and if history continues to repeat itself, we will maybe see other centers closing due to these kinds of constraints, and so for just sort of pure survival of a center, I think that it’s important to say, You know what? You gave us this money, and here’s in fact what we do with it, and this is the impact that it had.

Jennifer also stated that program evaluation is important:

To have outcomes to report on teaching and learning center[s], particularly at research universities, can be in tenuous positions, so it’s good to be able to say to our vice Provost, look at this year’s impact of the programs, and they can also give us some feedback on what we need to improve or change or take strategic initiatives in a different way.

Although only one evaluator mentioned Patton’s Utilization evaluation model by name, the content analysis revealed a pattern of all of the participants reporting how evaluations can be used. These patterns make “how evaluations can be used” a primary theme of the faculty development program evaluation experience. Table 4.5 below, provides the comments made by each of the participants re-presented to state how evaluations can be used and depict evidence for the theme *Implicit Application Of A Utilization-Focused Approach To Planning Evaluation*.

Table 4.5 *Implicit Application Of A Utilization-Focused Approach To Planning Evaluation*

Theme: Implicit Application of A Utilization-Focused Approach To Planning Evaluation	Quote from Participant
Stephanie	Can help us track impact that is spread through the University
Connie	We could make what we do more clear to the teaching community
Ann	If you don't have data you can't make a case for programs to continue being funded or someone can't go to bat for you
Tiffany	Well there's positive advantages that range from sort of validating the role of the center
Allison	Advantages would be helping to communicate the value of our work to internal audiences
Erin	Have in place a feedback mechanism
Jennifer	They can also give us some feedback on what we need to improve or change or take strategic initiatives in a different way
Dave	So is this for our own purposes to improve what we're doing is it to report to someone else about something

Chapter Summary

Overall, the primary themes that emerged are: (a) purposeful selection of programs to evaluate, (b) careful planning of the evaluation methodology, (c) understanding that it takes considerable time to evaluate programs, (d) strategically linking program outcomes to center goals, and (e) implicit application of a utilization-focused approach to planning evaluation. The secondary theme that emerged is (a) thinking critically about how to evaluate programs. A table with evidence for each theme was provided. Evidence Table 4.1 was provided for the first two themes originating from the Posavac (2011) conceptual framework. All other themes originated from the interviews with the participants of this study (Constas, 1992).

Themes revealed that participants in this study are purposeful in planning methods for assessing impact. Although they reported that conducting program evaluation is a challenge because it takes considerable time, they want to collect more than program participation data. In fact, participants reported spending considerable time thinking critically about how to evaluate their programs. In addition, they look for methods to link faculty development program outcomes to center goals and ways to utilize evaluation results. Overall, these themes reflect consistent elements and patterns in participants' comments on the processes they engage in when evaluating faculty development programs, their use of program theory, and the evaluation models they use to design faculty development program evaluations. Chapter 5 provides the conclusions, discussion, implications, and recommendations related to the goals of this study.

Chapter 5

Conclusions and Recommendations

This research effort sought to explore evaluator perceptions of the processes they engage in when evaluating faculty development programs; their use of program theory; and the evaluation models they use to design faculty development program evaluations. The findings of the study were presented in Chapter 4. Conclusions drawn from the findings presented in Chapter 4 are grouped by research question and discussed below. Implications for practice and recommendations for future research follow.

Findings

Question One: What process do evaluators engage in when they evaluate various programs in faculty development?

Finding One: Based on the findings of this study, many faculty development evaluators do not use a formalized evaluation process; however, they do engage in various evaluation strategies such as participation data and self-reports from faculty who have been through their programs.

Finding Two: Participating evaluators spend time thinking about how to move beyond collecting participation data. They are trying to design and implement evaluation processes that can assess the impact of their programs.

Finding Three: The evaluators do not follow all of Posavacs' (2011) steps in the evaluation planning process. For example, identifying which programs to evaluate and planning evaluation methodologies were the most common steps evaluators reported utilizing in their evaluation planning. However, less often reported were (a)

identification of relevant stakeholders, (b) identifying information needs, (c) planning evaluations, (d) examination of available literature, and e) writing evaluation plans.

Question Two: To what extent and how do evaluators link evaluations to short- medium-, and long-term outcomes?

Finding One: Based on the findings of this study, although these evaluators are thinking critically about ways to link program outcomes to center goals, they do not have procedures or methodologies for achieving that connection.

Finding Two: The evaluators do not use program theories or logic models in their evaluation work to link program evaluations to short-, medium-, and long-term goals of their centers. Instead, they are using qualitative methods, annual surveys, and validated measures and looking for strategies that will help them develop processes that will evaluate impact.

Question Three: To what extent and how do evaluators use the standard evaluation models in faculty development?

Finding One: Based on the findings of this study, the evaluators are not familiar with the many evaluation models in the extant literature that has been designed to help guide and inform evaluators in creating and implementing evaluation plans. Most evaluators do not use evaluation models in their work; however, all of the evaluators who participated in this study identified beneficial uses for conducting evaluations.

Conclusions

All the participants in this study reported that they have selected programs to evaluate purposefully. The programs they reported selecting are grant programs and long- and short-term programs such as workshops, consultations, and institutes. While these evaluators also reported

a desire to move away from simply collecting participation data, participation and satisfaction information is the most common type of data they have collected. Only two evaluators reported being familiar with program theory, and none of the evaluators reported using logic models as a way to link short-, medium-, and long-term outcomes. However, all of the evaluators reported that they are seeking ways to link program inputs to program outcomes with recognizing that this effort is termed “program theory.”

Although only one evaluator mentioned Patton’s utilization evaluation model by name, all of the participants reported that they have used evaluations to communicate programming results. In addition to using evaluations to communicate results, they also suggested that evaluation results could be used to garner campus-wide support and funding. When speaking about the process evaluators engage in when they evaluate various programs in faculty development, these evaluators also described “thinking” critically about how to evaluate their programs. Evaluators also acknowledged that it takes considerable time to plan and evaluate their programs effectively. Thus, while reporting that they spend considerable time planning and conducting evaluations, these evaluators have not availed themselves of the evaluation models or evaluation planning processes found in the literature that could organize and simplify their efforts. Without the aid of established models and planning steps, the evaluation process can consume more time and effort than necessary.

The evaluators in this study did not report involving stakeholders in evaluation planning, nor that they link program activities to program outcomes as a way to align program evaluations with center goals. However, it has been demonstrated that evaluators can improve program evaluation in faculty development if they select appropriate data sources, involve stakeholders, and provide evaluation context (Wright, 2011).

Discussion

As a review, five primary themes emerged: (a) purposeful selection of programs to evaluate, (b) careful planning of the evaluation methodology, (c) understanding that it takes considerable time to evaluate programs, (d) strategically linking program outcomes to center goals, and (e) implicit application of a utilization-focused approach to planning evaluation. In addition, a secondary theme emerged: (a) thinking critically about how to evaluate programs.

Both formative and summative program evaluation assesses the value of a specific program (Stufflebeam, 2001). Whereas, formative evaluation is conducted during the program to assess the implementation process or development of the program and provide data relative to the improvement of the program; summative evaluation is conducted at the end of the program to determine the program's worth or effectiveness (Scriven, 1991).

Evaluators in this study evaluated center programs that included workshops, consultations, institutes, other short-term programs sponsored by the center, and grant proposals; however, these evaluators described their efforts to conduct summative evaluations of their programs. For instance, a sample representative comment was, "...we're thinking more...about assessing the impact of workshops, especially since it involves a lot of human capital to put together a workshop series." Or expressed differently, "...at the end of the semester, we gather feedback about what was helpful about it [and] what could be changed so that it could be more helpful." Most of the evaluators interviewed commented that they were planning methods to provide a more in-depth look at how faculty members use their services, what the outcomes are, and the impact on teaching. The evaluators in this study rarely mentioned formative evaluations.

Centra (1978) and Nelson (1979) emphasized the need for program evaluation in faculty development and argued that evaluation is an essential indicator of program effectiveness. In

addition, Shore (1976) recommended that centers determine who conducts evaluations, how often they are conducted, and what will happen to the results. In this study, a search of the sixty-five faculty development centers found on the Professional and Organizational Development in Higher Education (POD) website revealed only seven faculty development centers employing program evaluators. Program evaluation in faculty development is a new phenomenon, and it is uncommon to find a large number of full-time employed evaluators at faculty development centers. In addition, because there are so few centers with budgets large enough to hire dedicated evaluators, in most cases evaluating center programs is a small part of many other responsibilities held by staff program evaluators.

Implications for Practice

The results of this study concerning faculty development center evaluators' perception of their program evaluation processes and use of program theories and evaluation models suggest that program evaluators in the faculty development field should familiarize themselves with the evaluation processes, program theories, and evaluation models available in the literature and utilize them in faculty development evaluation. For example, the Posavac (2011) evaluation process was used as the framework for this study because it offers a comprehensive evaluation planning process, encompassing the steps recommended by many of the evaluation planning processes covered in previous research. The first implication is that faculty development program evaluations should utilize the program evaluations reviewed in the literature, which can aid faculty development evaluators to evaluate their programs.

Utilizing program theories and logic models in faculty development evaluation may also help evaluators plan and implement evaluations, conduct the evaluations, and communicate evaluation results. Table 5.1 below, represents examples of possible components of faculty

development logic. Logic model components include inputs, outputs, and outcomes (Fig. 2.1). Inputs are the resources, contributions, and investments that go into the faculty development program. Outputs are the activities, services, events, and products that reach faculty members who participate. Outputs in faculty development may include any of the services faculty development centers offer. Outcomes are the results or changes for the center, faculty, academic units, and/or the university. Outcomes are therefore referred to as indicators of change in individuals. Examples of questions to ask when determining outcomes are “How will we know that we have met our goals?” and “What will be the indicator(s)?” Outcomes may be determined based on a center’s mission and or the goals of the program.

Table 5.1 *Faculty Development Logic Model Components*

Inputs	1. Assessing the needs of the faculty or groups 2. Materials, equipment and facilities 3. Staff 4. Costs of producing program 5. Time in faculty schedules
Outputs	1. Presentations 2. Workshops 3. Demonstrations 4. Action research 5. Faculty inquiry groups consultations
Outcomes	1. New knowledge and skills 2. Change in practice 3. Change in academic unit or university culture 4. Change in student learning change in student engagement

The second implication is that the logic model can facilitate the (a) identification of evaluation questions, (b) alignment between the need for evaluations and the performance of those evaluations, and (c) planning and implementation of evaluations. By assisting in the selection of the most appropriate evaluation model, the logic model contributes to the processes involved in conducting the evaluation and communication results. Also, since the evaluators in this study had spent considerable time thinking critically about how to link their program goals and outcomes to their centers' goals and missions, they can understand program theory and use logic models in the evaluation planning process in terms of the unique mission of each center. As indicated in the literature review conducted for this study, in practice, a variety of evaluation models (e.g. objectives-based, program theory, utilization-focused) are demonstrated to have appropriate design and flexibility to meet the needs of evaluators in planning the evaluation process.

Recommendations for Future Research

This study was undertaken to add to the existing research of program evaluation in faculty development. A descriptive case study of evaluation in faculty development is recommended. The suggested study design would include onsite observations, interviews with stakeholders, plus document review and analysis. The small number of faculty development centers with designated program evaluators on staff limited this study and until more centers employ evaluators, future research on evaluation processes in faculty development will likely be limited. However, by gathering many evaluation stories in the suggested investigation, a greater understanding of faculty development evaluation issues in a broader context would be gained.

Another recommendation for future research is for assessing the program evaluation needs of faculty developers. The majority of faculty development centers do not have designated staff members conducting program evaluation. Findings from a needs assessment study utilizing a survey created with a focus on program evaluation training and resource needs in faculty development would provide the field with strategic initiatives, areas of improvement, and focus in the field.

Summary

This study offers three key contributions to the understanding of program evaluation in faculty development based on the rich and in-depth description of their experiences with program evaluation in faculty development programs provided by evaluators of faculty development centers. The first contribution of this study is that it provides insights into evaluation in faculty development and helps those interested in evaluating their programs understand more about the evaluation process. It also points to implications for how program theory and evaluation models may be applied in the field of faculty development. Second, the findings of this study can provide a foundation for future studies to examine factors related to program evaluation and its application to professional development. Further, qualitative methodology is demonstrated as a viable approach that can increase the likelihood of establishing a systematic evaluation design and implementation plan.

In 1976 John Centra said that faculty development centers needed to learn to evaluate their programs and demonstrate that they manage their centers efficiently, effectively educate faculty, and produce more effective and satisfied faculty members. Centra (1976) also argued that program evaluation is an essential indicator of program effectiveness. Not only is it important to designate individuals to conduct center evaluations (Shore, 1976), this study has

demonstrated that program evaluation knowledge and skills are necessary for successful faculty development.

References

- Anfara, V. A., Brown, K. M., Mangione, T. L. (2002). Qualitative analysis on stage: Making the research process more public. *Educational Researcher*, 31(7), 28-38.
- Armstrong, E. G., & Barsion, S. J. (2006). Using an outcomes-logic model approach to evaluate a faculty development program for medical educators. *Academic Medicine*, 81(5), 483-488.
- Bates, R. (2004). A critical analysis of evaluation practice: The Kirkpatrick model and the principle of beneficence. *Evaluation and Program Planning*, 27, 341-347.
- Bazeley, P. (2009). Analyzing qualitative data: More than identifying themes. *Malaysian Journal of Qualitative Research*, 2, 6-22.
- Bothell, T. W., & Henderson, T. (2004). Evaluating the return on investment of faculty development. In C. Wehlburg & S. Chadwick-Blossey (Eds.) *To improve the academy: Resources for faculty, instructional, and organizational development*, (Vol. 22, pp. 52-70). Bolton, MA: Anker.
- Centra, J. A. (1976). Faculty development practices in U.S. colleges and universities. Princeton, N.J.: Educational Testing Service.
- Centra, J. A. (1978). Faculty development in higher education. *Teachers College Record*, 80, 188-201.
- Chen, H. (1990). Issues in constructing program theory. *New Directions for Program Evaluation*, 47, 7-18.
- Cilliers, F. J., & Herman, N. (2009). Impact of an educational development programme on teaching practice of academics at a research at a research-intensive university. *International Journal for Academic Development*, 15(3), 253-267.

- Connolly, M. R., & Millar, S. B. (2006). Using workshop to improve instruction in STEM courses. *Metropolitan Universities Journal*, 17(4), 53-65.
- Constas, M. A. (1992). Qualitative analysis as a public event: The documentation of category development procedures. *American Educational Research Journal*, 29(2), 253-266.
- Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches*. Thousand Oaks, CA: Sage.
- Dwyer, J. (1996). Applying program logic in program planning and evaluation. *Public Health and Epidemiology Report Ontario*, 7(2), 38-46.
- Felder, R. M., & Brent, R. (2010). The national effective teaching institute: Assessment of impact and implications for faculty development. *Journal of Engineering Education*, 99(2), 121-134.
- Ferren, A., & Mussell, K. (1987). Strengthening faculty development programs through evaluation. In J. Kurfiss (Ed.), *To improve the academy: Resources for faculty, instructional, and organizational development*, (Vol. 6, pp. 133-143). Stillwater, OK: New Forums Press.
- Frechtling, J. A. (2007). *Logic modeling methods in programs evaluation*. San Francisco CA: Jossey-Bass.
- Gaff, J. G., & Morstain, B. R. (1978). Evaluating the outcomes. In J. G. Gaff (Ed.), *Institutional renewal through improvement of teaching*. *New Directions for Higher Education*, (Vol. 24, pp. 73-83).
- Guskey, T. R. (1995). Professional development in education: In search of the optimal mix. In T. Guskey and M. Huberman (Eds.), *Professional Development in Education: New Paradigms and Practices* (pp. 114-131) New York: Teachers College Press.

- Guskey, T. R. (1999). New perspectives on evaluation professional development. Paper presented at the Annual Meeting of the American Educational Research Association Montreal, Quebec, Canada April 19-23, 1999.
- Hacsi, T. A (2000). Using program theory to replicate successful programs. *New Directions for Evaluation*, 87, 71-78.
- Hammond, R. (1973). Evaluation at the local level. In B. R. Worthen & J. R. Sansers (Eds.), *Educational evaluation: Theory and Practice*. Belmont, CA: Wadsworth.
- Hines, S. R. (2011). How mature teaching and learning centers evaluate their services. In J. E. Miller & J. E. Grocci (Eds.), *To Improve the Academy*, 30, 277-290.
- Hines, S. R. (2009). Investigating faculty development program assessment practices: What's being done and how it can be improved? *Journal of Faculty Development*, 23(3), 5-10.
- Hoyt, D. P., & Howard, G. S. (1978). The evaluation of faculty development programmes. *Research in Higher Education*, 8, 25-39.
- Jacobson, W., Wulff, D. H., Grooters, S., Edwards, P. M., & Freisem, K. (2009). Reported long-term value and effects of teaching center consultations. In L. B. Nilson & J. E. Miller (Eds.), *To improve the academy: Resources for faculty, instructional, and organizational development*, Vol 27 (pp. 223-246). San Francisco: Jossey-Bass.
- Johnson, R. B. (1997). Examining the validity structure of qualitative research. *Education*, 118(2), 282-292.
- Kirkpatrick, D. L., & Kirkpatrick, J. D. (1994). *Evaluating training programs: The four levels*. San Francisco: Berrett-Koehler.
- Knowlton, L. W., & Phillips, C. C. (2009). *The logic model guidebook: Better strategies for great results*. Thousand Oaks, CA: Sage.

Krebera, C., & Brook, P. (2001). Impact evaluation of educational development programmes.

The International Journal for Academic Development, 6(2), 96 - 108.

Kvale, S. (1996). *InterViews: An introduction to qualitative research interviews*.

Thousand Oaks, CA: Sage.

Levinson-Rose, J., & Menges, R. J. (1981). Improving college teaching: A critical review of research. *Review of Educational Research*, 51, 403-434.

Liamputtong, P. (2009). Qualitative data analysis: conceptual and practical considerations.

Health Promotion Journal Australia, 20(2), 133-9.

Light, G., Calkins, S., Luna, M., & Drane, D. (2008). Assessing the impact of faculty development programs on faculty approaches to teaching. *International Journal of Teaching and Learning in Higher Education*, 20(2), 168-181.

Mark, M. M. (1990). From program theory to tests of program theory. *New Direction for Program Evaluation*, 47, 37-51.

Maxwell, J. A. (1996). *Qualitative research design: An interactive approach*. Thousand Oaks CA: Sage.

Maxwell, J. A. (2005). *Qualitative research design: An interactive approach*. Second edition. Thousand Oaks CA; Sage.

McLaughlin, J. A., & Jordan, G. B. (1999). Logic models: A tool for telling your program's performance story. *Evaluation and Program Planning*, 22, 65-72.

Megdal, L., Engle, V., Pakenas, L., Albert, S., Peters, J., & Jordan, G. (2005). Using program logic model analysis to evaluate and better deliver what works. In *Proceedings of 2005 European Council for and Energy Efficient Economy Summer Study*. France.

- Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco: Jossey-Bass.
- Metcalfe, S. A., Aitken, M. A., & Gaff, C. L. (2008). The importance of program evaluation: How can it be applied to diverse genetics education settings? *Journal of Genetics Counseling*, 17, 170-179. doi 10.1007/a10897-007-9138-8
- Metfessel, N., & Michael, W. (1967). A paradigm involving multiples criterion measures fir the evaluation of the effectiveness of school programs. *Educational and Psychological Measurement*, 27, 931-943.
- Milloy, P. M., & Brooke, C. (2004). Beyond bean counting: Making faculty development needs assessment more meaningful. In C. Wehlburg & S. Chadwick-Blossey (Eds.) *To improve the academy: Resources for faculty, instructional, and organizational development*, (Vol. 22, pp. 71-92). Bolton, MA: Anker.
- Nelson W. C. (1979). Faculty development: Key issues for effectiveness. *Forum*, 2(1), 1-4.
- O'Donoghue, T., & Punch, K. (2003). *Qualitative educational research in action: Doing and reflecting*. Routledge Falmer.
- Palomba, C. A., & Banta, T. W. (1999). *Assessment essentials: Planning, implementing, and improving assessment in higher education*. San Francisco, CA: Jossey-Bass.
- Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newbury Park, CA: Sage.
- Patton, M. Q. (2008). *Utilization focuses evaluation*. Thousand Oaks, CA: Sage.
- Posavac, E. J., & Carey, R. G. (2007). *Program evaluation: Methods and case studies*. Upper Saddle River, NJ: Pearson.

- Ryan, G. W., & Bernard, H. R. (2003). Techniques to identify themes. *Field Methods*, 15(1), 85-109.
- Rossi, P. H., Freeman, H. E., & Lispey, M.V. (1999). *Evaluation: A systematic approach*. Thousand Oaks, CA: Sage
- Saldana, J. (2009). *The coding manual for qualitative researchers*. London: Sage Publications
- Sandelowski, M. (2000). Whatever happened to qualitative description? *Research in Nursing and Health* v23, 334-340.
- Sarikaya, O., Kalaca, S., Yegen, B.C., & Cali, S. (2010). The impact of a faculty development program: Evaluation based on the self-assessment of medical educators from preclinical educators from preclinical and clinical disciplines. *Advances in Physiology Education*, 34, 34-50.
- Scriven, M. (1991). *Evaluation Thesaurus*. Sage Newbury Park, CA.
- Shore, B. M. (1976). Success and failure of formal teaching improvement efforts in higher education, *Higher Education Bulletin*, 4, 123-138.
- Sorcinelli, M.D. (2002). Ten principles of good practice in creating and sustaining teaching and learning centers. In K. H. Gillispie, L. R. Hilsen, & E. C. Wadsworth (Eds.). *A guide to faculty development: Practical advice, examples and resources* (pp. 9-23). Bolton, MA: Anker.
- Sorcinelli, M. D. (2006). *Creating the future of faculty development: Learning from the past, understanding the present*. Bolton, MA: Anker.
- Sorenson, D. L., & Bothell, T. W. (2004). Triangulating faculty needs for the assessment of student learning. In C. Wehlburg & S. Chadwick-Blossey (Eds.), *To improve the*

- academy: Resources for faculty, instructional, and organizational development*, (Vol. 22, pp. 23-40). Bolton, MA: Anker.
- Stinson, S., & Wilkinson, C. (2004). Creating a successful clinical extern program using a program planning logic model. *Journal for Nurses in Staff Development*, 20(3), 140-144.
- Stufflebeam, D. (2001). Evaluation models. *New Direction for Evaluation*, 2001(89). DOI:10.1002/ev.2
- Sullivan, A. M., Lakoma, M. D., Billings, J. A., Peters, A. S., & Block, S. D. (2005). Teaching and learning end-of-life care: Evaluation of a faculty development program in palliative care. *Academic Medicine*, 80(7), 657-668.
- Tracy, S. J. (2010). Qualitative quality: Eight big tent criteria for excellent qualitative research. *Qualitative Inquiry*, 16, 837-851.
- Travis, J. E., Hursh, D., Lankewitz, G., & Tang, L. (1996). Monitoring the pulse of the faculty: Needs assessment in faculty development programs. In L. Richlin & D. DeZure (Eds.), *To improve the academy: Resources for faculty, instructional, and organizational development*, (Vol. 22, pp. 23-40). Stillwater, OK: New Forums Press.
- Tyler, R. W. (1942). General statement on evaluation. *Journal of Educational Research*, 35, 492-501.
- Van Note Chism, N., & Szabo, B. S. (1997). How faculty development programs evaluate their services. *Journal of Staff, Program and Organizational Development*, 15(12), 55-62.
- Wright, M. C. (2011). Measuring a teaching center's effectiveness. In C. E. Cook & M. Kaplan (Eds.), *Advancing the culture of teaching on campus: How a teaching center can make a difference*, (pp. 38-49). Sterling, VA: Stylus Publishing.

- Wright, R. J. (2008). *Educational Assessment: Tests and Measurements in the age of accountability*. Thousand Oaks, CA: Sage Publications.
- Way, D. G., Carlson, V. M., & Piliero, S. C. (2002). Evaluating teaching workshops: Beyond the satisfaction survey. In D. Lieberman & C. M. Wehlburg (Eds.), *To improve the academy: Resources for faculty, instructional, and organizational development*, Vol. 20 (pp. 94-106). Bolton, MA: Anker.
- Weiss, C. H. (2000). Which links in which theories shall we evaluate? *New Directions for Evaluation*, 87, 35-45.
- Wergin, J. F. (1977). Evaluating faculty development programs. *New Directions in Higher Education*, 17, 57-76.
- Young, R. E. (1987). Evaluating faculty development programs: Program goals first. In J. F. Wergin & I. A. Braskamp (Eds.), *New Directions for Institutional Research*, (Vol. 56, pp. 71-82). doi:10.1002/ir.37019875609
- Zepeda, S. J. (2008). *Professional development: What works*. Larchmont, NY: Eye on Education.
- Yin, R. K. (1989). *Case study research: Design and methods*. Newbury Park, CA: Sage Publications.
- CA: Sage.

Appendices

Appendix A

Informed Consent Statement

Dear Prospective Participant,

You have been identified as an evaluator or assessor in a faculty development center and you are invited to participate in a study regarding evaluation conducted in faculty development centers. This research is being conducted by Thelma Woodard, a fourth year Ph.D. candidate from the department of Evaluation Statistics and Measurement at the University of Tennessee, Knoxville.

This study will offer three main contributions to the understanding of evaluation and assessment in faculty development. First, directors and staff of faculty development centers will have a rich and in-depth description of evaluator's experiences with evaluation in faculty development centers. Second, the findings can provide a foundation for future studies to examine factors related to evaluation and assessment.

If you volunteer to participate in this study, I would ask you to complete one interview about your experience as an evaluator or assessor in faculty development centers. Interviews will last approximately 30 to 40 minutes and will be audio recorded. If you opt not to be recorded, you may still participate.

All interviews will be audio recorded with digital technology. Recording will be uploaded to the computer of the principal investigator. Recordings will be transcribed. Your name will not be used in publications resulting from the study; instead a pseudonym will be used. The name of the university will also be a pseudonym to further aid in privacy concerns. All recordings and related notes will be stored in a locked and secure file cabinet. Only the researcher will have access to the information you provide.

Although it is not perceived that you will directly benefit from participating in this study, the data generated from your interview, along with others, will be used to inform those in our field about the evaluation and assessment in faculty development centers.

In rare cases, a research study will be evaluated by an oversight agency, such as the University of Tennessee Institutional Review Board or the federal Office for Human Research Protections that would require that I share the information collected from you. If this happens, the information would only be used to determine if this study was conducted properly and adequately protected your rights as a participant.

If you have questions at any time about the study or the procedures, (or you experience adverse effects as a result of participating in this study,) you may contact the researcher at

Thelma Woodard,
College of Education, Health, & Human Sciences
1122 Volunteer Blvd., A503 Bailey Education Complex

Knoxville, TN 37996
865-247-5827

If you have questions about your rights as a participant, contact the Office of Research Compliance Officer at (865) 974-3466.

_____ Participant's initials

Consent:

By signing this consent form, I am indicating that I have read and understand the information provided above and I agree to participate in this study. I understand that my participation is voluntary and I am free to withdraw at any time. I understand that I can still participate in this study, even if I do not wish to be recorded.

☐ *I agree to be audio-taped*

☐ *I do not want to be audio-taped*

Participant's Name _____

Participant's Signature _____

Date _____

Appendix B

Letter to Participants

Subject: Seeking help for study about faculty development evaluation

Greetings,

My name is Thelma Woodard. I am currently a Ph.D. candidate in Educational Psychology under the direction of Dr. Gary Skolits at the University of Tennessee-Knoxville (UT). I am presently conducting my dissertation study on program evaluators' perceptions of faculty development evaluation. I would like to ask for your voluntary participation in my research.

Over the next few weeks, I am seeking to interview program evaluators in faculty development about their perceptions and experiences. Accordingly, I would enjoy the opportunity to talk with you. Participation takes the form of completing one telephone interview that will last approximately 30 to 40 minutes and can be scheduled at your convenience. **At your earliest convenience, please reply to this email and let me know if you would be willing to participate in this study.**

Despite what I am sure is a very hectic schedule, I hope that you will be able to take the time to share your insights with me. Also, feel free to contact me should you have any questions about my research or the interview process. You can also reach Dr. Gary Skolits, my dissertation committee chair, for support of my research either by email at gskolits@utk.edu or by phone at (865) 974-2777.

Thank you for your time and I look forward to hearing from you soon!

Respectfully,

Thelma Woodard
Doctoral Candidate

Appendix C

Interview Protocol

Introduction

Thank you for agreeing to participate in this interview. This interview should take approximately 30 to 40 minutes. I will be recording this interview. Do you have any questions regarding the procedure or the IRB informed consent form you signed? Thank you. Let's begin the interview.

Background Questions:

1. How long have you held a position of responsibility in faculty development?
2. How many people do you serve?
3. What is your student population?
4. Is your institution public or private?
5. In what professional associations are you a member?

Interview Questions:

1. What has your experience in evaluation and assessment in faculty development been like?
2. What types of programs do you evaluate or assess?
3. What needs to be done/what steps do you take when you are contemplating the evaluation of a program?
4. How do you evaluate the impact of the programs at your center?
5. In what circumstances do you use logic models in your work?
6. What are the advantages or disadvantages of documenting outcomes?
7. What evaluation or assessment models do you use in your work?
8. How familiar are you with the idea of program theory?
9. How do you link your evaluations to the short, medium and long term outcomes of your centers goals and mission?
10. If other centers wanted to start evaluating their programs, what advice would you give them?

Final Questions:

After all of this discussion about evaluation and assessment in faculty development, what concluding or final thoughts do you have? And, Do you have any evaluation reports and annual report you can share?

Appendix D

Research Audit Trail

A physical audit trail documents the stages of a research study and reflects the key research methodology decisions (Yin, 1989). The physical audit trail for *An Investigation of Faculty Development Center Staff Evaluators Perceptions of Processes, Theory, and Models Used in Evaluating Faculty Development Programs* study is as follows:

The research proposal: Based on this research problem, a proposal was developed and submitted to my dissertation committee for approval. This proposal included an introductory chapter, a literature review and methodology chapter. An abbreviated proposal was submitted to The University of Tennessee, Knoxville Institutional Review Board for approval.

Reviewing the literature: An in-depth review of the program evaluation in faculty development literature was undertaken. Despite decades of research in this area, the literature review highlighted a gap in the literature exploring program evaluation in faculty development centers from the evaluator's standpoint and utilization of qualitative methods in the study of program evaluation in faculty development was underused.

Designing a research framework: The next step involved designing a research framework to support the collection of evidence. Since the literature lacks evidence from the evaluators' personal experience, perspective and practice, a descriptive qualitative approach. A descriptive qualitative method utilizing semi-structured interview questions was selected as an appropriate research strategy. In addition, Posavac's (2011) evaluation planning process which explains how evaluators can plan evaluations was used as the conceptual framework for this study.

The interview schedule: The semi-structured interview was the source of descriptive

qualitative evidence. Based on issues identified in the literature and in defining the research problem, an initial interview protocol was prepared. This was pre-tested in a pilot interview to determine the participants understanding of questions and the depth of the research inquiry, and was subsequently refined.

Selection of interview participants: In order to achieve breadth and depth of coverage all evaluators of faculty development programs, which differed in a number of respects, listed on the POD website were chosen to participate in this study. The participants selected had in-depth knowledge of program evaluation in faculty development. Through purposive and snowball sampling, participants were identified and asked to participate in the study.

Evidence collection: In total eight semi-structured interviews were conducted. These interviews lasted between 30 and 45 minutes and were recorded and transcribed. The interview transcriptions were used in developing the study's themes.

Managing and analyzing the evidence: A qualitative descriptive approach was used to analyze the narrative from the transcripts. Through reflection on these narratives, the data was coded into key themes. A thematic evidence table was created to depict quotes excerpted verbatim from the pages of transcribed text as evidence for each theme.

Vita

Thelma Woodard was born in Midland, Texas. She worked at Southwestern Bell Telephone Company for twelve years before resigning to pursue a college degree. She completed an Associate Degree in General Studies at Midland College, a Bachelors Degree in Speech Communication at The University of Texas at the Permian Basin, and a Masters Degree in Communication Studies at Texas Tech University. After working several years in the nonprofit industry, she was awarded the National Urban Fellowship, completed a Masters of Public Administration at Baruch College, and subsequently was accepted into the Ph.D. program at the University of Tennessee Knoxville. Her doctoral program concentration is Educational Psychology and Research and her major is in Evaluation, Statistics, and Measurement. During her four years of full-time studies, she worked as a graduate assistant in the Department of Educational Psychology with the Institute for Assessment and Evaluation and the Tennessee Teaching and Learning Center. Thelma will continue working in higher education with a focus on evaluation and assessment of educational programs.