EFFECTIVENESS PLANS AND PRACTICES IN PROGRAMMATIC ACCREDITATION: DIFFERENCES IN EVALUATIVE CULTURE IN NURSING AND ENGINEERING

by

Paul Richard Swift

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Education

MONTANA STATE UNIVERSITY
Bozeman, Montana

May 2024

DEDICATION

To my family, who inspire and drive so many of my efforts.

ACKNOWLEDGEMENTS

I am thankful for all of the faculty in the Adult and Higher Education program at Montana State University. I'm also thankful for my fellow students in the program: lots of shared laughs, frustration, and support from this wonderful group along the way. I am thankful for my committee in helping shepherd me through this process and see it to the end. Dr. Sweeney Windchief, you run the most effective Doctoral Thesis course; we still need to make our Yellowstone River float happen. Dr. Bryce Hughes, you asked me the hardest questions early in the process and this final product is so much better due to your input. Dr. Martha Cabell, you've been a wonderful colleague and having you on my committee eased a great deal of my anxiety; I appreciate your professional input on institutional effectiveness practices. Most especially, a big thanks to my chair, Dr. Carrie Myers. Your thoughtful feedback was essential in getting me to think deeper and more clearly about my topic and you were incredibly generous with your time. I would also like to acknowledge and thank the participants in my study—I'm grateful for you sharing your time, your practice, and your thoughts with me.

TABLE OF CONTENTS

TABLE OF CONTENTS CONTINUED

TABLE OF CONTENTS CONTINUED

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

ABSTRACT

One of the main methods of public accountability in higher education is the use of accreditation to measure quality and ensure continuous improvement. In recent decades, accrediting bodies have moved away from inputs-driven requirements to requirements that focus on outcomes. One of the major shifts has been towards requiring institutions and programs to report on their effectiveness. However, despite the significant commitment of resources towards accreditation, there is little research around the practices within accredited programs. The purpose of this grounded theory study was to better understand the practices of practitioners of program effectiveness in accredited programs of engineering and nursing at two separate institutions of higher education. The study found significant differences in the scope of effectiveness practice between the two disciplines, with engineering practitioners focusing primarily on the assessment of student outcomes while nursing practitioners focused on systems that comprehensively evaluated many different aspects of their program. The study further found that most of the practitioners had come to their positions as novices; this presented an opportunity to learn from these seasoned professionals and theorize best practices for the field that may contribute to improved effectiveness practices for programs that are programmatically accredited. Suggestions are also provided for the field of accreditation more broadly to help clarify terminology and expectations.

CHAPTER ONE

INTRODUCTION

While accreditation and program evaluation are central to the operations of higher education throughout the United States today, there is striking paucity of research on accreditation and the evaluative practice of program effectiveness required by most accrediting bodies. While institutional accreditation can seem removed from everyday practices at a departmental level, for those programs that carry programmatic accreditation, accreditation is very much at the forefront. However, despite the significant human and fiscal resources devoted to programmatic accreditation and evaluation for accredited programs, we have little empirical knowledge on what the practice of program effectiveness looks like within programmatically accredited units. This study is an attempt to fill—to some extent—that gap in knowledge.

Accreditation was first created as a voluntary process undertaken by institutions to help establish quality norms and distinguish between the different levels of education at the numerous institutions that were proliferating in the United States in the late 19th and early 20th century. Today, however, it has become a near necessity for any institution of higher education. Lacking regional accreditation, an institution is left without access to Title IV federal funding, the lack of which would be a death-knell for most institutions (Brittingham, 2009; Congressional Research Service, 2017; Gillen et al., 2010). Besides serving as the gatekeeper to federal funding, accreditation also bears the responsibility of the main mechanism for promoting constant quality improvement in institutions and has taken on the role of a de facto stamp of quality assurance despite many attempts at creating other means of evaluation and assessment in higher education (Harcelroad, 1980).

Within higher education, accreditation is an ongoing process of peer-review to ensure that an institution or program meets established quality standards (USDE). The Congressional Research Service classifies accreditation in higher education into three types: regional accreditation, national accreditation, and specialized or programmatic accreditation (2017). Regional accrediting agencies grant accreditation to entire institutions, national accrediting agencies accredit faith-based and private career institutions, and specialized or programmatic accrediting agencies accredit programs—typically ones housed within institutions that are also regionally accredited. Thus, it is not uncommon for larger institutions to have an over-arching regional accreditation while also having several programmatic accreditations for certain areas of academic study.

Like regional accreditation, programmatic accreditation—while ostensibly voluntary—is often essentially required of programs. The reasons vary. In some fields, programmatic accreditation is a requirement for students to be able to license and practice after graduation. In other fields, programmatic accreditation is preferred by employers, thus decreasing the value of a degree from a program that does not carry accreditation. Further, access to opportunities for continued education within a field are often limited if one has not graduated from a programmatically accredited institution (Overbay & Aaltonen, 2001). Thus, we see that accreditation—in its various types—is a necessary and multilayered component to the operations of virtually all institutions of higher education today.

Accreditation is also resource intensive (Barczykowski, 2018; Beno, 2004; Shibley & Volkwein, 2002; Willis, 1994; Woolston, 2012). Initial programmatic accreditation, under the best of circumstances, takes three years (Trifts, 2012). And the work has only just begun at that

point. Accredited institutions and programs are required to implement a plan of continuous program evaluation and improvement. This plan, often referred to as a plan for program effectiveness, has become the main channel through which accreditation attempts to ensure effective evaluation practices and overall improvement in academic programs. This ongoing requirement is paired with the periodic preparation of an accreditation self-study report that is followed by the hosting of on-site evaluators whose job is to verify and evaluate the program's self-study report. If successful, the program's accreditation is reaffirmed by the accrediting body. As can be inferred, the resources required for the enterprise of accreditation are significant (Alstete, 2007; Heriot et al., 2009). Within an institution, where there may be overlapping institutional and programmatic accreditation timelines and requirements, careful thought to resource allocation is necessary (Longenecker, 2012).

Added to the investment of human resources are considerable financial burdens. Just being accredited costs money in the form of member dues paid by institutions and programs to their various accrediting bodies. The costs of the site-visit for reaffirmation also lie with the institution or program. In addition, programmatic accreditation frequently imposes requirements as to the number or qualifications of faculty—adjustments made to meet the bar can increase the cost of a program. When factoring in all of the costs required—including the hiring of new faculty—one institution estimated that the annual additional cost of pursuing initial accreditation from AACSB, the body that accredits business programs, could top $2 million (Kelderman, 2009); Heriot et al. (2009) found that the average annual cost of maintaining AACSB accreditation was $400,000. In the current resource-scarce environment of higher education, these costs are anything but insignificant. As such, making informed decisions about the cost and

return on investment of accreditation at all levels is critical. For the accreditations deemed worthy of pursuit by an institution or program, it stands to reason that going about the process in the most efficient and effective way is paramount to ensuring ongoing institutional and programmatic financial well-being. The aim of this study is to help understand how practitioners' experience can help inform an understanding best practices around accreditation and program evaluation.

Statement of the Problem

For such an essential component of higher education, that carries a significant investment of both human and fiscal resources on programs and institutions, it is surprising that there is very little extant research on the practices involved in accreditation. One can find many "best practice" type lists to help programs as they embark on accreditation (Friedlander & MacDougall, 1990; Nichols, 2005; Wood, 2006); however, only rarely are they supported by empirical evidence suggesting their effectiveness. In one of the few empirical studies, Longenecker (2012) looked at the ways in which three universities aligned their regional accreditation practices with the multiple programmatic accreditation requirements they held. They found that, in these institutions, efforts were being made to align institutional and programmatic accreditation requirements for efficiency. They also found that having multiple programmatic accreditations helped foster a culture of assessment at the institutions writ large. Finally, the institutions studied found that the benefits of their accreditations outweighed the costs.

The theme around alignment of accreditation processes that Longenecker found was mirrored by Bresciani (2009). In their grounded-theory study of the characteristics used by

thirteen exemplary programs to evaluate their outcomes-based assessment practices (a process called meta-evaluation—evaluating the evaluation or assessment), they found that one practice frequently used was embedding a program or institution's practices into existing program review or accreditation processes. Ultimately, Bresciani's results found nine criteria used by these programs to evaluate the effectiveness of their assessment of student-learning outcomes. These criteria were, clear understanding of goals and expectations, collaboration, use of results, awards and recognition, resources, coordination of the process, flexibility, addressing barriers, and evaluating the process itself (Bresciani, 2009),

Bresciani's study, along with many others that take on the theme of program effectiveness, focuses specifically on assessing student-learning outcomes (Banta, 2002; Bresciani, 2006; Kuh, Kinzie, Schuh, & Whitt, 2005; Maki, 2012; Mentkowski, 2000; Suskie, 2004). While an integral part of the work done by institutions and programs in higher education, program effectiveness requirements for accreditation generally encompass more than student-learning outcomes. The review of institutional or program mission, program outcomes, curricular design, academic policies, faculty scholarship, fiscal resources, physical resources, human resources, and support services all fall under the larger umbrella of effectiveness practices in accreditation. And while a plan for program effectiveness would include the evaluation of all these aspects, there is little extant literature that outlines best practices for doing so, and that which exists is almost wholly theory based. The question of what constitutes effective practices in programmatic evaluation and accreditation in higher education remains largely unstudied.

Purpose of Study

The purpose of this grounded theory study was to understand the experience of practitioners carrying out program effectiveness in two academic disciplines that are programmatically accredited within larger institutions of higher education. The practitioners came from two R-1, doctoral granting research institutions, with practitioners in the fields of nursing and engineering being represented at each institution. These experiences were used to understand discipline-specific characteristics of effectiveness practice, as well as to theorize best practices in the implementation of program effectiveness for programmatically accredited academic units. These best practices are intended to assist practitioners of program effectiveness as they implement their plans withing their own programs.

Research Questions

Focusing on programs that are programmatically accredited, as opposed to looking at institutional accreditation, this study will be looking for practices that are both distinct and common to the programs in the study. The guiding research questions for the study are:

RQ 1) How do practitioners at two western, doctoral-granting R-1 institutions of higher education understand and experience their programs' formation and implementation of their program effectiveness plans?

RQ 2) How has the context of institution and discipline influenced practitioners' program effectiveness practices?

RQ 3) How do these practitioners' experiences and knowledge contribute to an understanding of generalizable (Ragin & Becker, 1992; Stake, 2006; Yin, 2003) best practice in program effectiveness?

Significance of Study

While currently there exists some theoretical literature around program evaluation and accreditation in higher education, the lack of empirical studies is glaring. It is hoped that this study will be an initial foray into research around the practices related to programmatic accreditation, specifically as it relates to the ongoing requirement for program effectiveness. The results of this study have implications for both institutions and programs. Since the fiscal and human costs of accreditation are not insignificant, how to carry out programmatic accreditation and program evaluation in the most effective manner should be of interest. This study will be of particular importance for programs that are, or are seeking to become, programmatically accredited but may also be of interest to those working in institutional accreditation as well. Due to the overlap of program effectiveness planning and general evaluative practices, there may be theoretical implications for effective program evaluation as well.

Conceptual Framework

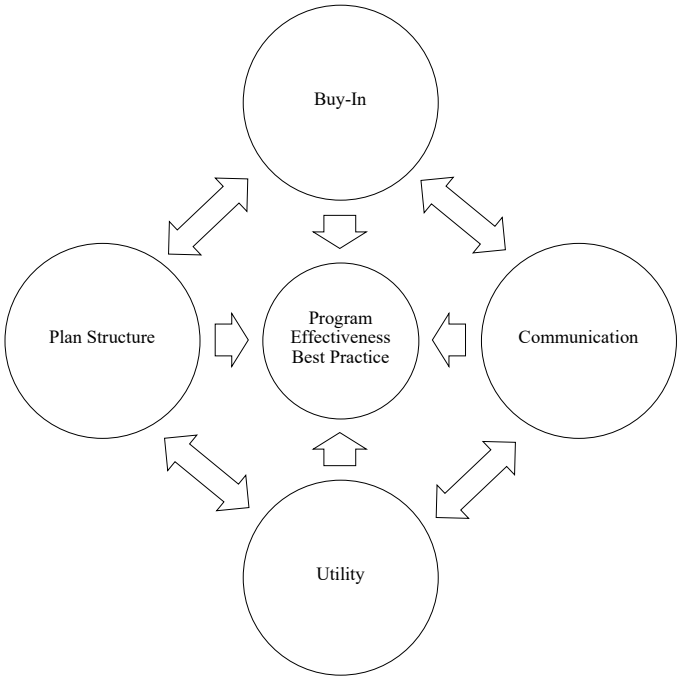

Figure 1. Initial Framework for Best Practices in Program Effectiveness Planning

Frameworks for evaluation are often represented as a cyclical, iterative process. Something as familiar as the Plan-Do-Study-Act (PDSA) associated with W. Edwards Deming is a good example of such a cyclical approach. It is assumed for this study that some kind of process like this is most likely incorporated into the program effectiveness plans within most accredited programs. The framework here, instead, focuses on constructs seen as critical to establishing and implementing best practices in program effectiveness. The literature points out several barriers to effective evaluative practices in higher education, and the constructs outlined here are emblematic of those barriers which can also serve as opportunities to establish good practice.

These constructs are not stratified aspects of best practices in program effectiveness—they interact with each other in complex ways. For example, good communication leads to better buy-in of the parties involved in the process; an effective plan structure facilitates utility of findings. The constructs are not hierarchical, nor are they cyclical but interact with each other at all points (Figure 1).

Constructs

Buy-In

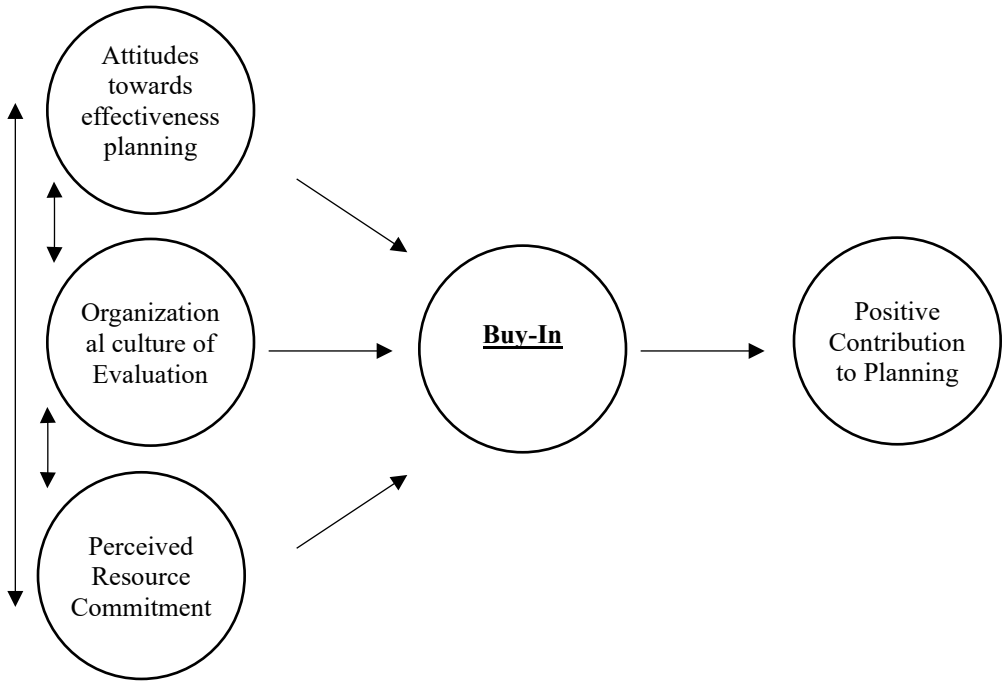

Figure 2. Theory of Positive Contributions to Planning, adapted from Ajzen, 1991.

It is well documented that the attitudes of particularly faculty (Miller, 2013; Rickards & Stitt-Bergh, 2016; Walton, 2022; Welsh & Metcalf, 2003b)—but also administration and staff (Walton, 2022; Welsh & Metcalf, 2003b)—are important factors in the success of effectiveness

practices and the assessment of programs. Faculty who have an outcomes-focused approach to teaching and learning are more likely to be supportive of assessment practices (Ewell, 2009; LoCascio, 2010; Longenecker, 2012; Welsh & Metcalf, 2003b). Attitudes of administration towards effectiveness planning also shape faculty and staff conceptualization and support of the process (Walton, 2022). Understanding and guiding faculty, staff, and administrative attitudes is central to successful effectiveness planning (Bresciani, 2009; Longenecker, 2012; Miller, 2013; Walton, 2022).

The overall organization, whether that be the institution or the program, also lies somewhere on a scale of a developed culture of evaluation. The importance of where a program lies on that scale is also significant to garnering overall buy-in towards program effectiveness efforts (Miller, 2013). The greater the culture of evaluation, the greater the buy-in that exists for effectiveness planning and implementation. Another important aspect in gaining support around effectiveness practices in the literature is that of appropriate commitment of resources (Lynch, 2018; Walton, 2022). This commitment could take the form of human resources, time, or money.

The theory of planned behavior (TPB) has been shown as a valid predictor of behavior (Ajzen, 1991). The theory makes use of a combination of attitude, perceived behavioral control, and a subjective norm variable to form a construct of intent that is a good predictor of behavior (Ajzen, 1991). For this study, the construct of buy-in replaces that of intent in the original theory as a good predictor of positive contributions to effectiveness planning in administration, faculty, and staff. Viewed through the lens of TPB, attitudes towards effectiveness and assessment practices would substitute for the general “attitudes” measure of TPB. Perception of resource commitment fills in for perceived behavioral control. The existing culture of evaluation takes the

place of the subjective norm variable. Finally, these three factors interact to form the construct of buy-in, predicting success in effectiveness planning (Figure 2).

Communication

Another construct important to the research on effectiveness practice is communication (Bresciani, 2009; Walton, 2022). In Patton's Utilization-Focused Evaluation (U-FE) model, effective communication is a critical aspect of success throughout the process (Patton & Campbell-Patton, 2022). One can find in phenomenological studies on accreditation processes many critiques of the communication at work (Bush, 2016; Walton, 2022). Poor communication is a large dissatisfier and a source of confusion and poor assessment outcomes (Walton, 2022). Criticism of effectiveness plans are often that once they are made they seem to "disappear" (Walton, 2022).

Communication theories abound with differing ontological, epistemological, and axiological roots. Communication within organizations has distinct features (Baker, 2007). Organizational communication theory discusses levels of communication: interpersonal communication, group level communication, organizational level communication, inter-organizational level communication, and mass communication (Baker, 2007). These levels of communication serve various functions. Baker (2007) outlines these functions as being: compliance-gaining, leading, motivation, and influencing; sense-making; problem-solving and decision-making; and conflict management, negotiating, and bargaining. When discussing communication within an organizational setting it's important to be cognizant of both the level at which the communication is happening as well as the function it seems to be serving.

Plan Structure

The structure that the plan for program effectiveness takes is important. Here, structure means, what does the plan itself look like? How is it broken down into its component pieces? Things that should be included in any plan are things like evaluation processes, timelines, parties included, outcomes, steps taken, etc. The plan structure should also somehow outline how evidence is collected and where it resides. There are differing approaches to the structure of the effectiveness plans. On the one hand, an effectiveness plan could be structured according to some sort of evaluative theory or framework (Hanna et al., 2016; Ingersoll & Sauter, 1998; Oermann, 2017; Saewert, 2013). Another common approach in higher education is to structure the effectiveness plan around accreditation standards (Bresciani, 2009; Longenecker, 2012; Oermann, 2017).

Utility

As the name would suggest, U-FE puts a central importance on the use of the evaluation. As accreditors are increasingly looking for not only the existence of an effectiveness plan, but documented evidence of its use, this focus on utilization in U-FE can be helpful to an accreditation effectiveness plan. Thus, in a U-FE approach to effectiveness planning, use is considered at every juncture (Patton & Campbell-Patton, 2022). Further, a U-FE approach stresses that the simple construction of the effectiveness plan is not the end goal—acting on the findings in the effectiveness plan—again, use—is the goal.

Unlike the type of formal evaluations that U-FE typically addresses which tend to have a set ending with a formal report, however, accreditation effectiveness plans are meant to be ongoing. We should not allow this to deter us from applying the valuable aspects of U-FE into an

effectiveness plan, though. Indeed, in some ways the ongoing requirement of the effectiveness plan enhances U-FE and addresses a deficiency in the field of evaluation for accountability in which it has arisen (Rickards & Stitt-Bergh, 2016). Hence, the focus for an improvement paradigm approach, as suggested by Ewell (2008), needs to be on how to create a plan that is *ongoing*. For that to happen, the plan needs to be fully integrated into the program with thoughtful points of feedback. Ewell (2009) echoes the importance of integrating the plan. It also needs to be evaluated itself in a process called metaevaluation (Bresciani, 2009).

In an ongoing system of evaluation, what Ewell (2008) has referred to as an improvement paradigm, there is a focus on building a "culture of evidence" (2008, p. 10). One of the benefits of this ongoing type of approach, as opposed to a more summative approach, is that evidence can be gathered over time, tracking progress over areas of intended improvement. There must be ongoing feedback loops built into the system as well, allowing stakeholders multiple chances to evaluate improvement initiatives and progress. The hope is that integrating the assessment this deeply within the program will illuminate the effectiveness of program assessment to the program and encourage its continued use. In essence, the hope is to ultimately develop and cultivate an enduring culture of assessment and learning within the program.

Research Methodology

The grounded theory methodology employed used multiple bounded case studies of two programs each at two separate institutions to better understand practitioner experiences in program effectiveness and, through a cross-case analysis of the cases, theorize best practices in program effectiveness. Participants were criterion-based, with practitioners of program effectiveness chosen from two separate institutions. Within each of these institutions,

practitioners were chosen from the field of nursing and the field of engineering, resulting in a total of five cases. Theoretically, the sampling approach was based on the notion of institutional theory proposed by Meyer and Rowan (1977), and the idea that the field of program evaluation, as it relates to programmatic accreditation, has matured to look similar—what institutional theory would label "isomorphic"—independent of the field being evaluated. Essentially, institutional theory applied to program effectiveness practices would stipulate that programmatic accreditation is an institutionalized approach to hold the larger institution (in the theoretical sense of institution) of higher education—and the academic sub-unit being accredited—accountable in a ceremonial way. The sampling approach chosen allowed a test of that theoretical proposition by providing for both separate discipline contexts as well as overarching institutional (in the educational sense of institution) environments. The hope was that identifying these cases purposefully in this way would allow for any discipline or institution-specific influences on effectiveness practices to be illuminated, thus placing those aspects of program effectiveness that are more universal (i.e., isomorphic) in relief to be able to examine more thoroughly. That is, this approach allowed for what Yin (2003) has called "analytic generalizability" using "replication logic."

Operational Definitions

1. *Institutional Accreditation*—The goal of accreditation is to ensure that institutions of higher education meet acceptable levels of quality. Accreditation in the United States involves non-governmental entities (accrediting organizations) as well as federal and state government agencies (these three entities are formally known as the Triad). Accreditation's quality assurance function is one of the three main elements of

oversight governing the Higher Education Act's (HEA's) federal student aid programs. In order for students to receive federal student aid from the U.S. Department of Education (Department) for postsecondary study, the institution must be accredited by a "nationally recognized" accrediting agency (USDE).

2. *Programmatic Accreditation*— Specialized or programmatic accreditation normally applies to programs, departments, or schools that are parts of an institution. The accredited unit may be as large as a college or school within a university or as small as a curriculum within a discipline. Most of the specialized or programmatic agencies review units within an institution of higher education that is accredited by an institutional accrediting agency (USDE)
3. *Assessment*—Yarbrough, et al. (2011) define assessment as “The determination of relative or absolute position on some variable of interest based on qualitative and/or quantitative evidence” (p. 283)
4. *Evaluation*—Yarbrough, et al. (2011) define evaluation as a “Systematic investigation of the value, importance, or significance of something or someone along defined dimensions (e.g., a program, project, or specific program or project component)” (p. 287).
5. *Program*—ABET, the accrediting body for engineering programs, defines a program as “an integrated, organized experience that culminates in the awarding of a degree. The program will have program educational objectives, student outcomes, a curriculum, faculty, and facilities” (2024a).

6. *Effectiveness Plan*—The Southern Association of Colleges and Schools Commission on Colleges (SACSCOC)—the first regional accrediting body to require effectiveness planning in their standards—defines the plan for institutional effectiveness as engaging in "ongoing, comprehensive, and integrated research-based planning and evaluation processes that (a) focus on institutional quality and effectiveness and (b) incorporate a systematic review of institutional goals and outcomes consistent with its mission" (2017, p. 19). By substituting "program" for institution in the above statement, one has a good definition of program effectiveness planning. Please note that not all bodies use this term consistently, but this is the operational definition that will be used for this study.
7. *Standard*—Yarbrough, et al. (2011) define standard as "A principle commonly agreed to by experts in the conduct and use of evaluation, that when implemented will lead to greater evaluation quality" (p. 292). Accreditation requirements are often referred to as standards; another common term is criterion.
8. *Metaevaluation*—Yarbrough, et al. (2011) define metaevaluation as the "Evaluation of an evaluation" (p. 289).
9. *Design*—Yarbrough, et al. (2011) define design as "A plan for conducting an evaluation; e.g., data collection schedule, report schedules, questions to be addressed, analysis plan, management plan, etc. Designs may be either preordinate or emergent" (p. 286).

10. *Accountability*—Yarbrough, et al. (2011) define accountability as a "Demonstrated responsibility for the use of resources, activities, or decisions made in the course of a program and/or its evaluation" (p. 283).

Assumptions, Limitations, and Delimitations

Assumptions

1. Most institutions of higher education in the United States take part in various forms of accreditation.
2. There are aspects of program effectiveness planning required by programmatic accreditation that are common to most programs, regardless of their academic focus.
3. While there are uniquenesses to setting at every institution, there will be aspects of program effectiveness planning experiences that are common to most practitioners in the field.
4. The cases in the study are engaged in an honest effort to carry out the best plan for program effectiveness that they can.
5. The practitioners interviewed were able to self-reflect and honestly assess their practices around program effectiveness and accreditation.
6. Motivations and emotions around accreditation are mixed; accreditation is often essentially required, on the one hand, but also can be seen as serving important functions around quality improvement. Different people have different viewpoints they bring when discussing accreditation.

Limitations

1. The study looks only at cases at two different institutions in two different fields, limiting the scope. This study was looking for what Yin (2003) calls "replication logic," which differs from "sampling logic."
2. The focus of this study was on only two programmatic accrediting bodies—ABET and the Commission on Collegiate Nursing Education (CCNE). There are 96 U.S.-recognized institutional and programmatic accrediting bodies that could be studied. It is possible, or even likely, that such a study might provide different results.

Delimitations

1. While there are discipline-specific aspects of accreditation that are no doubt important, this study is focused primarily on what the practices of effectiveness in engineering and nursing tell us about effectiveness practices generally. That is, there will not be a heavy focus on curricular dictates or faculty requirements that are specific to the discipline; instead the focus will be more on the structure of the accreditation process and how they may be similar or different across the disciplines.
2. Environment and context no doubt play a role in most aspects of program operation at any institution of higher education. While noted where relevant, this study did not seek an exhaustive sample of institutional location, type, and size.

Chapter Summary

While accreditation is a fundamental aspect of higher education there is little research on practitioners' experience of the process and the best practices in implementing the ongoing

program effectiveness plan as part of the accreditation process. In the current resource-starved higher education environment, finding ways to efficiently carry out the important work of program evaluation is critical. It's important not only for ensuring constant quality assessment and improvement within academic units and the institution at large, but also in fulfilling accreditation requirements at both the program and institutional level. In Chapter Two I will discuss the history of accreditation and program evaluation in U.S. higher education. I will also undertake a review of the literature pertinent to this study.

CHAPTER TWO

LITERATURE REVIEW

The purpose of this grounded theory study was to understand the experience of practitioners carrying out program effectiveness plans within the context of programmatic accreditation. These practitioners were stationed within two larger institutions that also have institutional accreditation requirements. Information gained from these experiences was used to theorize best practices in program effectiveness planning in the context of programmatically accredited academic units. The following literature review will primarily be focused on evaluation and assessment as it relates to, primarily, the programmatic level. As such, the review does not include issues at the level of student and classroom. Such important topics as theories of student learning assessment, information included in the National Survey of Student Engagement, and program evaluation related to external funding agencies are not covered.

Accreditation Background

There are four main types of accreditation in the U.S. today: regional, national faith-related, national career-related, and specialized or programmatic (Hegji, 2017; Eaton, 2015). A search on the Council for Higher Education Accreditation's (CHEA) database of colleges, universities, and higher education institutions recognized by U.S. accrediting organizations turns up over 8,000 results; a search of accredited programs finds over 30,000. The database also lists 96 U.S.-recognized institutional and programmatic accrediting organizations (Council for Higher Education Accreditation, n.d.). The large number of accredited institutions highlights the fact that while all accreditation is nominally voluntary, lack of accreditation, which would bar institutions

from receiving Title IV funding from the federal government, would cause most institutions to collapse (Brittingham, 2009; Gillen et al., 2010; Hegji, 2017). We also see in these numbers that, in addition to their institutional accreditation, many institutions house departments and programs that are programmatically accredited, adding layers of complexity to these institutions' overall accreditation pictures and the fiscal and human resources built in to support them.

Various organizations have outlined the purposes that accreditation serves, with some slight differences amongst them. The U.S. Department of Education (USDE) lists the functions of accreditation as follows:

- Assess the quality of academic programs at institutions of higher education
- Create a culture of continuous improvement of academic quality at colleges and universities and stimulate a general raising of standards among educational institutions
- Involve faculty and staff comprehensively in institutional evaluation and planning (USDE)

CHEA lists the roles of accreditation as: assuring quality, granting access to federal and state funds, engendering private sector confidence, and easing transfer (Eaton, 2015). Accreditation serves all these purposes, often with varying degrees of success.

Accreditation did not always take on these roles, however. Indeed, if someone were to have imagined the purposes served by accreditation and were then to create a system to fulfill those purposes, they most likely would not have designed the system we have (Brittingham, 2009; Gillen et al., 2010). By comparison, the U.S. system of oversight at the post-secondary level is unique in the world in its decentralization and complexity (Brittingham, 2009). At the institutional accreditation level, the U.S. has seven regional accrediting bodies that oversee

accreditation at degree-granting institutions of higher education. While there has been a push for standardization amongst the regional accrediting bodies (Brittingham, 2009; CHEA; Crow, 2009; Gillen et al., 2010), the current situation still puts institutions with similar missions but different accreditors in the position where they are meeting differing standards. While choice of accrediting body was once determined solely by geographic location, it is now allowable for an institution to seek accreditation by regional accreditors that would have once been considered outside their traditional region. To this day, support for accreditation remains controversial (Alstete, 2007; Bardo, 2009; Birnbaum, 2000; Bruhn, 1993; Nichols, 1995; Scriven, 2000; Swing & Coogan, 2010; Welsh & Metcalf, 2003a). There is no question, however, that it remains one of the primary drivers of systematic attempts to evaluate and improve institutions and programs (Alstete, 2007; Challa et al., 2005; Heriot et al., 2009; Kuh & Ewell, 2010; Longenecker, 2012; Oden, 2009; Volkwein et al., 2007; Welsh & Metcalf, 2003a; Welsh & Metcalf, 2003b).

For a variety of reasons, programmatic accreditation—also purportedly voluntary—is in many cases essentially required of programs. In many fields, licensure or requirement for practice require graduation from a program that carried programmatic accreditation. In some fields without the licensure or practice requirements, graduates of accredited programs are preferred by employers or by graduate programs, making a program without programmatic accreditation less valuable to students (Kelderman, 2009; Overbay & Aaltonen, 2001; Pavalkis & Kelley, 2016). And so, there are many different factors to consider for institutional and programmatic leaders when considering programmatic accreditation.

At the level of programmatic accreditation, there has long been cries to stem the proliferation of the various programmatic accrediting bodies. As far back as 1924, institutional

leaders have been complaining of the burden of numerous accreditations while also questioning accreditation's value (Pinkham, 1955). However, there is no clear indicator of what would serve as substitute in the lack of regional and programmatic accreditation (Overbay & Aaltonen, 2001). As it stands now, accreditation remains a fundamental aspect of higher education operations and is neglected to the detriment of the institution and programs that are accredited or could benefit from accreditation.

One of the main reasons that accreditation has been able to retain a foothold in the workings of higher education has been its adoption of evaluation and assessment practices. While primarily instigated by calls for accountability after accreditation took on the role of gatekeeper for federal funding, state governments have also mostly accepted accreditation's role in evaluating higher educational pursuits and assessing outcomes. As will be shown, this appropriation of the accountability mantle has resulted in a system that tries to serve both evaluation for improvement and evaluation for accountability (Ewell, 2008)—a sometimes tenuous balancing act.

The combination of state and federal oversight, together with accreditation, is often referred to as "the Triad" and did not coalesce at once. There were, rather, specific historical points where different aspects merged to form what we now consider the foundations of accreditation. As Dickeson (2009) stated, "the structure of accreditation is more historical than logical" (p. 6). In the next section I'll trace the roots of the history of accreditation, and then the history of program evaluation, and show how they came together in the mid-1980s to establish, roughly, the shape of accreditation and program evaluation as it exists in higher education today.

The next section traces the important events that have shaped the landscape of accreditation, evaluation, and accountability in higher education as it roughly stands today.

Historical Events in Accreditation

The historical threads that have led to the current state of accreditation in the United States were driven by social, political, economic realities. Unwrapping these threads helps us to understand the unique—and somewhat incongruent—shape of the current system. These threads include the proliferation of new educational institutions that came with westward expansion in the latter 19th and early 20th century, the growth of professional education, the accreditation requirement for the receipt of Title IV federal funding, and events surrounding the reauthorization of the Higher Education Act in 1992.

Proliferation of Educational Institutions

The first significant historical event affecting accreditation—indeed, probably the one most directly important for its beginnings—was the proliferation of educational institutions in the United States from the latter half of the 19th century to the beginning of the 20th century. The main drivers of this proliferation were the expansion of Western settlers into indigenous lands, the elaboration of primarily denominational institutions in existing states, and the passing of the Morrill Act of 1862 followed, eventually, by the Hatch Act of 1887 and the second Morrill Act of 1890 (Geiger, 1995; Johnson, 1981)—it is worth noting that tribal colleges and universities were not established as land grant institutions until the Equity in Educational Land-Grant Status Act was signed into law by Bill Clinton in 1994. There were also fundamental changes regarding the nature of knowledge and learning that played out during this time that helped to drive the

formation of new types of institutions (Gruber, 1975). These forces combined to form, towards the end of this period, a collegiate environment that looks much more contemporary than that which had existed until the middle of the 19th century.

The rise of the industrial revolution granted new importance to types of learning that had not found much credence under the classical curriculum that was predominant at institutions of higher education in the early history of the colonies and United States. While the Morrill Act of 1862 explicitly referenced these changes in its call for studies in agriculture and the mechanical arts, the act was alluding to changes in the educational landscape that were already taking place with the variety of new institutions that were being formed. Indeed, looking at lists of the types of educational institutions being founded at the time one sees that many of the designations make little sense to a contemporary eye. There were not necessarily the clear demarcations of primary, secondary, and post-secondary education we have come to take for granted today (Thelin, 2019). And so, the growth included not only schools focusing on agriculture and mechanical arts as the Morrill Act outlined, but also a growth in professional schools, seminaries, normal schools, vocational and technical schools, as well as separate schools for women and newly freed African Americans (Geiger, 1995; Gruber, 1975; Thelin, 2019).

It is within this context of not only an increase in the sheer numbers of institutions, but the elaboration of varying types of institutions, that the call for standardization began to grow in the late 19th and early 20th century. Another important factor in this call for standardization reflected the fact that the proliferation of these institutions had largely preceded the primary and secondary educational infrastructures necessary to support them (Johnson, 1981). Due to this, nearly all of the newly founded institutions offered some form of remedial education for

underprepared students (Johnson, 1981). The lack of prepared students with no agreed upon standards for college entrance, combined with varying definitions of what the composition of a collegiate course of study should be, led to drastic differences based on region and institution (Hawkins, 2007; Johnson, 1981).

Many different associations and organizations stepped in to attempt to bring clarity and standardization to the field. While the Carnegie Foundation, American Association of Universities, U.S. Bureau of Education, American Association of Colleges, American Council on Education, and others made varying attempts to fill this role (Boyd, 1973; Harcleroad, 1980; Hawkins, 2007; Shawen, 1983), it ultimately was the accrediting bodies—both regional and programmatic—that stepped in as the primary groups establishing standards and norms in higher education at the turn of the 20th century (Brubacher & Rudy, 2017; Harcleroad, 1980; Hawkins, 2007; Shawen, 1983). This historical peculiarity has resulted in a system of non-governmental oversight of post-secondary education that is largely unique to the United States (Brittingham, 2009; Pinkham, 1955). This system, while bringing some standardization to the chaos of educational institutions in the late 19th century, was also decentralized enough to allow for the diversity that exists in higher education in the U.S. with institutions of varying scopes and missions (Brittingham, 2008; Brittingham, 2009; Eaton, 2007).

The Growth of Professional Education

A sub-point in the history of proliferating institutions listed above is the particular growth of institutions offering professional education. One example of this was in the field of medicine, which also eventually saw the earliest attempt at programmatic accreditation that stemmed from the American Medical Association (AMA). Originally founded in 1847 to promote “scientific

advancement, standards for medical education, launching a program of medical ethics, [and] improved public health" (American Medical Association) the AMA first began categorizing medical schools in 1905 (Eaton, 2012; Harcleroad, 1980). Facing many of the same challenges present in higher education at large at this time, and concerned about the lack of standards, the AMA turned to the Carnegie Foundation for the Advancement of Teaching for help with what became the Flexner Report of 1910, the result of visits to 155 medical schools in the United States and Canada (Cooke et al., 2010).

This report had a tremendous influence on medical education. It led to the closure of many of the more dubious medical schools in existence at the time (Pavlakis & Kelley, 2016; American Medical Association). It also accelerated and, in essence, codified the change to a reliance on science-based education in the medical field; as such, the argument in the Flexner Report was that medical schools ought to be housed in the places that conduct science-based research: universities (Flexner, 1910). Many of the recommendations in the report are still standard practice today (Cooke et al., 2010). Not long after the publication of the report—in 1927—the AMA Council on Medical Education first published its list of approved medical schools, essentially starting medical accreditation, the first example of programmatic accreditation in the United States (American Medical Association).

This model of programmatic accreditation that was established by the AMA is one that has influenced the formation of other programmatic accreditations. We see that programmatic accreditation in other fields—engineering, business, law, nursing, architecture—just like Flexner had suggested for the medical field, takes place within the context of larger institutions of learning. This context of institutions of higher education with regional accreditation housing

several programs with their own accreditation is important to understanding the entire context of accreditation, evaluation, and assessment as it exists today in higher education. It also has led to cries from institutional leaders about the overwhelming nature of requirements from the various accrediting bodies housed within their organizations. Indeed, as early as 1924 the National Association of State Universities leaders were complaining about "the rapid increase in the number and variety of organizations which have undertaken to standardize procedures and policies in one or another branch of higher education" (as cited in Pinkham, 1955, p. 67). This sentiment was shared broadly (Harcleroad, 1980; Hawkins, 2007). In fact, in 1949 there was a joint committee formed by the Association of Land-Grant Colleges, Association of Urban Universities, Association of American Universities, and Association of American Colleges to help curb the influences of accreditation on higher education (Pinkham, 1955).

Accreditation as Requirement for Federal Funding

It is within this context that we turn our eyes to the third important event in the history of accreditation in the United States, which is the tying of federal funding to an institution's accreditation status. This link was first established with the Veteran's Readjustment Act of 1952. The act was an attempt to stem complaints that many of those taking advantage of the funding provided by GI Bill of 1944 for continuing education were attending colleges and universities of questionable quality (Gillen et al., 2010). With some exceptions, the Veteran's Readjustment Act explicitly tied the issuing of federal funds via students to accredited institutions (Gillen et al., 2010; Harcleroad, 1980). Not long after, with the passing of the Higher Education Act (HEA) of 1965, the tie between accreditation and federal funds was made even more complete, as the HEA required institutions receiving Title IV funding to be accredited (Hegji, 2017; Gillen et al., 2010).

The effects for accreditation, particularly institutional accreditation, have been profound. Accreditation had been founded as a voluntary process where member institutions policed themselves, primarily with a focus on quality improvement. And while accreditation still is, by definition, voluntary, with student financial aid being tied to accreditation status the effect has been that almost all institutions of higher education within the United States are now accredited (Flood & Roberts, 2017). In fact, the loss of regional accreditation typically serves as an effective death-knell to an institution (Flood & Roberts, 2017; Gaston, 2014).

Initially, there was trust that accreditation was an effective means of ensuring good stewardship of public funds granted to institutions of higher education through Title IV funding (Eaton, 2012; Hartle, 2012). However, over the decades, as more students took advantage of federal student aid, there were greater public calls for accountability. By the mid-1980s, it was apparent that the role of accrediting bodies deciding which institutions qualify to receive taxpayer dollars was leading to strong calls for accrediting bodies to hold institutions more accountable, a shift in focus from the purported purpose of accreditation being quality improvement (Eaton, 2012; Ewell, 2008; Hartle, 2012). It is also a shift that resulted directly from making accreditors the gatekeepers of federal student funding. And to this day, there remains criticism that accreditation has not lived up to its accountability responsibilities (Eaton, 2012; Ewell, 2009; Gillen et al., 2010; Hartle, 2012; Murray, 2009; Walton, 2022).

Despite these critiques, accrediting bodies have certainly made significant changes to try to address their changing role as engendering both institutional improvement and public accountability (Ewell, 2009). There has been a noticeable shift from focus on inputs (things like requirements for the number of faculty, number of library holdings, dollars committed to

education, etc.) to a focus on outputs (things like rate of degree completion, student learning outcomes, etc.), for example (Brittingham, 2009). Another fundamental shift was the move towards a focus on effectiveness. The Southern Association of Colleges and Schools Commission on Colleges (SACSCOC) first established a standard for institutional effectiveness in 1986 and continues to find it central to the accreditation process (Brittingham, 2009; Ewell, 2001; Jackson et al., 2010; Rodriguez, 2021). It wasn't long before the other regionals came along with their own requirements for institutions to have plans for institutional effectiveness (Ewell, 2001; Ewell, 2009). Today, the seven regional accrediting bodies, and most programmatic accreditors, require some form of ongoing plan for effectiveness that includes defined goals or objectives, the collection of data to determine progress to those goals or objectives, and proof of decisions made to improve institutional effectiveness based on the data collected (Ewell, 2001; Head & Johnson, 2011; Larkan-Skinnner, 2005; Rodriguez, 2021; Welsh & Metcalf, 2003a; Welsh & Metcalf, 2003b). Rodriguez (2021), while reviewing the regional accrediting bodies' mission statements, states that, overall, they show the main purpose of accreditation to be accountability through quality assurance and effectiveness planning processes.

Despite the central role that effectiveness planning has taken in accreditation recently, this ongoing plan for effectiveness has been described as the most problematic part of accreditation. One study found that seventy percent of community colleges undergoing SACSCOC accreditation were found out of compliance with institutional effectiveness standards (Head & Johnson, 2011). Head and Johnson (2011) found that the primary challenge for programs was the ability to look at the standard as a "practice" as opposed to a requirement—

that is, effectiveness planning is ongoing and systematic and not static, and practices need to reflect that. Manning (2011) also notes that the rate of recommendations in effectiveness standards is high and suggests that institutions that look at effectiveness planning as simply an act of compliance, as opposed to a method for transforming institutional culture, do not reap the benefits of the standard. And the problem continues today: In 2019, SACSCOC stated that the standards around effectiveness planning caused the greatest number of compliance issues for institutions and SACSCOC's class of 2021 *Most Frequently Cited* Principles *in Decennial Reaffirmation Reviews* shows 43% of institutions in non-compliance for the Quality Enhancement Plan standard (Southern Association of Colleges and Schools Commission on Colleges, 2021). Rodriguez (2021) suggests that ambiguous definitions of what constitutes appropriate effectiveness planning may be partially to blame for these challenges. Rodriguez (2021) also found that the challenges with effectiveness planning were not significantly different between institutional types, suggesting the challenges are pervasive to all institutions of higher education.

1992 Reauthorization of the Higher Education Act

Some of the critiques of accreditation's ability to hold institutions accountable came to a head with the reauthorization of the Higher Education Act in 1992. Indeed, for a short time it appeared that states might take over the job of the oversight of higher education as a result (Ewell, 2009). This was the culmination of a gradual buildup of frustration in many states at what they saw as the lack of accountability in higher education. This resulted in many states implementing forms of exams or assessment methods for students at their institutions by the end of the 1980s (Ewell, 2001; Ewell, 2011). Coming a bit late to the game, perhaps, in 1989 the U.S.

Department of Education for the first time called on accreditors to include standards on student learning outcomes as part of their recognition process (Ewell, 2001). By the time of the 1992 reauthorization, Congress had decided to establish State Postsecondary Review Entities (SPREs) to review institutions based on performance criteria (Brittingham, 2009; Ewell, 2009), which no doubt would have marked the beginning of the end of accreditation for accountability. However, Congress did not fund the SPREs and they were written out of the 1998 reauthorization of the HEA (Ewell, 2009). At the same time, facing budgetary challenges and recognizing changes being made in accrediting bodies—most dramatically around assessment of student learning outcomes—states mostly backed out of their oversight role of higher education by the early 2000s (Ewell, 2009; Ewell, 2011). This is not to say that concerns over student outcomes and accountability in higher education were completely washed away. Indeed, the Spellings Report of 2006 continued these themes, though the implementation was much more modest than the broad recommendations included in the report.

Accreditation came out of these events in many ways stronger, but also more closely tied to the federal government (Brittingham, 2009; Hartle, 2012). With the states forgoing the mantle of accountability, a precedent was established—that has only been reinforced since—of the federal government taking on the central accountability role over higher education, with accrediting bodies as its mediary (Ewell, 2009). As this historical view moves into the current state of higher education and accreditation we have only seen increases in the federal government's direct influences on accreditation. These include rules around credit transfer, confidentiality, accreditation appeals, textbooks, and distance education (Eaton, 2012; Ewell, 2009). Indeed, the current environment has increased the specialization of evaluation practices

needed at institutions of higher education, often forcing the role of faculty, once central to the process, to be subsidiary to the needs of an increasingly technocratic process, highlighting the increasing need for those trained specifically in the field of evaluation and assessment within institutions of higher education (Ewell, 2011; Rickards & Stitt-Bergh, 2016).

Historical Events in Evaluation and Assessment

While accreditation found its roots in standards creation and the need for the ability to align various levels of education, it eventually grafted with assessment and evaluation movements that had been growing in their own historical setting. This section looks at the important events in the history of evaluation and assessment as they relate to the way they've been incorporated into accreditation practices. These events include the movement away from the classical curriculum in the 19th century, the influence of early research on higher education in the mid 20th century, the domestic programs of Lyndon Johnson labelled as the Great Society, and the rapid growth of assessment as a field in the 1980s, coinciding with a growing skepticism over government spending.

Before we dive into the history, it may be pertinent to discuss the nomenclature around assessment and evaluation. In general parlance, the terms assessment and evaluation are frequently used interchangeably to mean roughly the same thing. In fact, the Oxford American Writer's Thesaurus lists them as synonyms of each other. While the field of assessment and evaluation does make a distinction between the two terms, unfortunately, it's not always followed strictly even by those researching the topics. Technically, assessment is simply some kind of variable—a number, a score, etc.—that is defined on some kind of continuum of possible outcomes, while evaluation is a method for judging quality, value, or importance of something

(Yarbrough, et al., 2011). Assessments are frequently used in evaluation, but—again, technically—the assessment itself carries no quality of judgement of quality. Once we begin judging worth, we have moved into the field of evaluation. For example, a score of an 85% on an exam is an assessment. Once we begin to judge whether that score demonstrates some level of competency in the material assessed, we're using evaluation.

As we'll see, the assessment movement of the late 20th century—which focused primarily on student learning assessment—makes use of evaluation frequently. In fact, assessment without evaluation would be uninstructive as to decisions about what to do next based off the assessment results. Technicalities aside, the term assessment is frequently used to cover both assessment and the evaluation of that assessment, particularly when looking at student outcomes. And student outcomes assessment is a major component of accreditation. Unfortunately, we're left in a state of imprecision at times with the terms simply due to the nature of English and general conventions of use. It is, however, an important finding that has also been echoed by others (Ewell, 2001; Rodriguez, 2021) and I will have more to say about the problem of the definition of terms in Chapter Five.

Movement Away from the Classical Curriculum

The first historical event of importance for the assessment movement in higher education was the slow movement away from the classical curriculum in the 19th century and the eventual foundation in the early 20th century of a new form of curriculum. This shift was long, essentially taking much of the 19th century to fully play out. It was also significant in that it was the first time in U.S. higher education history that we saw institutions, which had been so insular in their early history, largely bowing down to external pressure over internal decisions like curricula and

degree programs offered. By the turn of the century, higher education in the U.S. offered students an ever-increasing number of fields of study which were largely beginning to be overseen by a faculty that was increasingly professionalized and exerting their increased influence on matters of curricular design. This change to faculty as the central driver of curricular establishment and review, with a sometimes tenuous balance of response to outside calls for change, largely follows to the present day.

The colonial colleges relied on the Oxbridge model of education with rote memorization at its core (Gruber, 1975). It was based on a view of human knowledge as essentially static; thus, there was not any particular need for expertise outside of the established curriculum for faculty or, perhaps more historically accurate, tutors; there was certainly not room for deviation in content delivered (Gruber, 1975). It was a model that served not only in the colonial period, but well into the ante-bellum. The Yale Report of 1828, a response to calls for changes to curriculum that reflected a different, more scientific, appraisal of knowledge creation, was mostly a document of resounding support for the classical curriculum, with its defined "furniture of the mind" as the main content of instruction.

However, the fact that the report had to be written at all was evidence of calls for change to the classical curriculum. Indeed, while most of the established institutions were hesitant to make changes to their curricula, many of the new institutions being formed throughout the United States in the 19th century were responding to the demands of the time and establishing curricula in science, technology, teaching, agriculture, the vocational arts, and more (Geiger, 1995). It was the first time in the history of higher education in the U.S. where the curricular decisions were being driven, to a large extent, by calls outside the academy (Geiger, 1975).

While one could argue that these new institutions were largely responding to market demand, at first the academic decisions still lay primarily with the governing boards of the institutions or the legislatures of the states and not the faculty (Harcleroad, 1980).

At the same time as the growth in schools offering diverse courses of study was taking place, the notion of the professor as a profession was beginning to take shape (Cohen & Kisker, 2009). This was partially the result, no doubt, of the need for subject experts in fields that were, themselves, taking on the mantle of professions (Gruber, 1975). In any event, by the turn of the century, the professorship was taking on a form that looks somewhat contemporary, with increasing specialization and professionalization within individual fields of study. By 1915, the American Association of University Professors (AAUP) had been formed and its Committee on Academic Freedom and Academic Tenure produced a Declaration of Principles that was endorsed by the AAUP in 1916 (American Association of University Professors).

One can also see this evolving professionalism with growth in the faculty role. Yale's eventual movement to a more elective style curriculum was led by William Graham Sumner and other faculty on campus between 1883 and 1885—it was both a win for faculty control in review of the curriculum and also a bow to external pressure and competition for students (Harcleroad, 1980). We can also see the change in the University of Wisconsin, whose early academic decisions were under the purview of the Regents and then by the late 1850s, the Board of Visitors. However, by the turn of the century, faculty's role had grown significantly and they were increasingly involved in academic decisions like admissions and curriculum at the University (Harcleroad, 1980). This was a trend that only grew stronger through the 20th century.

The implications for evaluation of these foundational shifts in education were profound. In the classical curriculum, with an established set of unchanging facts that were conveyed, it was much easier to assess the outcomes of students and the curriculum at large. The eventual change to an understanding of knowledge as a dynamic process of creation, requiring an entirely different set of cognitive abilities like critical thinking and analysis, resulted in outcomes and curricula that are much more nebulous. Assessment and evaluation as a field really began to grow out of this environment. And no wonder, as the knowledge and skills required by an early 20th century college graduate looked radically different to those of one at the beginning of the 19th century.

Early Higher Education Research

Early research on higher education, beginning as early as the 1930s but accelerating into the 1960s and 1970s, has also had a profound impact in setting the stage for evaluation in higher education. While not evaluation or assessment, per se, these early forays into research in higher education set a base on which assessment and evaluation in higher education were built (Ewell & Cumming, 2017; Fitzpatrick et al., 2011). Early focus on student learning in college focused not only on cognitive gains, but also on the effects of college on students' attitudes (Ewell & Cumming, 2017). In the 1970s, Astin's research helped to establish the importance of longitudinal studies when examining the effects of college, Bowen's research tied the gains of higher education to public policy goals, and Pace helped point out the importance of environment and its ties to student behavior; Tinto's student integration model also proved useful for future research on student learning—all aspects that have carried over into the assessment movement (Astin, 1977; Bowen, 1977; Ewell, 2002; Ewell & Cumming, 2017; Pace, 1979;

Tinto, 1975). These researchers, and others working at the time, developed research methods and theories that would be adopted by assessment practitioners in higher education; they also established the important tradition of action research that the assessment movement has continued—assessment in higher education is still very much focused on evaluation for improvement (Ewell, 2002; Ewell & Cumming, 2017).

The Great Society

Another important historical event for evaluation in higher education was the legislative work of President Lyndon Johnson. The War on Poverty and the domestic programs labelled the Great Society set the stage for the maturation of program evaluation as a larger field, serving as a critical forebear of the assessment movement in higher education in the United States (Fitzpatrick et al., 2011; Stufflebeam et al., 2000). These federal investments into nearly every aspect of social life were so broad that they required levels of program evaluation that had not existed previously—in fact, finding those with expertise to evaluate such sweeping legislation proved to be a challenge (Fitzpatrick et al., 2011). The Elementary and Secondary Education Act (ESEA) of 1965, the educational aspect of Johnson's Great Society, brought massive increases in government spending in education. During debates on the passing of ESEA, doubts were brought as to evidence of government spending in education leading to positive academic outcomes. The result of these debates ended in the incorporation of the mandate that recipients of ESEA funds report on the results of their use of these federal funds (Fitzpatrick et al., 2011; Stufflebeam et al., 2000; Wholey & White, 1973). One sees in the ESEA debate a foreshadowing of conversations to be held later with regards to higher education that helped spur calls for greater evaluation and assessment.

We also see from this time a rapid expansion of mandates for evaluation actually written within legislation granting federal dollars to public projects (Fitzpatrick et al., 2011). As such, the field from the late 1960s through the 1970s matured dramatically in theoretical and methodological approaches. Such important names in the field as Scriven, Stufflebeam, and Wholey all began to publish explicitly on the topic of evaluation during this time. By the early years of the 1980s professional organizations had begun to form and publish work on standards and ethical guidelines (Fitzpatrick et al., 2011). A vocabulary was being born for those interested in evaluation that assessment in higher education would come to utilize (Ewell & Cumming, 2017; Fitzpatrick et al., 2011). Ironically, while the Reagan administration showed a decrease of federal grants for programs, it also allowed for a diversification within the field (Fitzpatrick et al., 2011; Shadish et al., 1991).

Acceleration of Assessment and Evaluation as a Field

It is this kind of maturation and diversification that would allow Ewell (2008), focusing on student learning assessment, to make the statement that, "the beginning of assessment as an explicit topic of policy discussion in the United States is usually marked at 1985" (p. 7). Indeed, the 1980s showed a combination of an increasing skepticism of government spending and a rapid acceleration of evaluation as a field. This was also the time of the publication of "A Nation at Risk," and growing concerns about the performance of U.S. educational institutions. It also correlated with a shift in higher education accreditation to reflect a greater emphasis on accountability and outcomes, as already described above (Barczykowski, 2018; Brittingham, 2009; Ewell, 2001). The assessment movement, with a particular emphasis on assessing student learning outcomes, had been born.

Perhaps to the chagrin of many in higher education, the assessment movement does not seem to have been a fad and its influence is still heavily entrenched in higher education (Ewell, 2009). Indeed, the wedding of accreditation with requirements for evaluation and assessment that was spurred by greater calls for accountability has only more deeply entrenched evaluation in higher education practices. The current trend is for increased involvement at the federal level, dictating more directly to institutions areas for perceived improvement (Eaton, 2012; Flood & Roberts, 2017; Hartle, 2012). While this may be dismaying for many within the academy, an argument can be made that evaluation mediated by accreditation is still a more cost-effective and less intrusive form of oversight than might otherwise exist (Brittinham, 2009; Lynch, 2018). If true, those wishing for higher education to continue operating with relatively little direct government intervention might consider full-fledged support for evaluation practices as they currently exist to lend credence to its effectiveness.

In this historical overview we've seen how, by historical happenstance, accreditation has come to be the bar by which institution and program quality in higher education is measured. While there have been multiple pushes and opportunities for other entities to take on the cause, accreditation has remained. The relatively recent movement in accreditation from assessing inputs to assessing outputs has come as the result of external pressures for accountability. The movement has kept accreditation as the common currency for accountability when, without the change, it may have been replaced by other external forces. Best practices in accreditation today have absorbed many of the practices of the field of assessment and evaluation to better meet the calls for accountability and quality control.

Assessment, Evaluation, and Effectiveness in Higher Education

When discussing assessment, evaluation, and effectiveness in higher education, it can become difficult to compare literature because definitions can be so divergent (Barczykowski, 2018; Ewell, 2001; Rodriguez, 2021). All seven regional accrediting bodies, and most programmatic accreditors, require some sort of ongoing, systematic evaluation that includes defined goals or objectives, the collection of data to determine progress to those goals or objectives, and proof of decisions made to improve program effectiveness based on the data collected (Ewell, 2001; Head & Johnson, 2011; Larkan-Skinnner, 2005; Welsh & Metcalf, 2003a; Welsh & Metcalf, 2003b). This ongoing plan for effectiveness has been described as the most problematic part of the accreditation process (Ewell, 2011; Head & Johnson, 2011; Manning, 2011; Rodriguez, 2021).

The plan for effectiveness can sometimes take on the name of assessment or evaluation in the literature. And yet, the problem with those terms is that they're often also used in the literature in the very specific sense of the assessment or evaluation of student learning, which is only one part of the totality of the plan for effectiveness. As pointed out by Head and Johnson (2011), there are various aspects of program effectiveness that are not directly related to student learning. Therefore, when discussing the overarching plan that a program or institution puts in place to monitor progress towards goals and make improvements, the term effectiveness is used in this study. When discussing student achievement of learning outcomes, the term assessment is used. Finally, there is much in the research and literature on program assessment and evaluation that has informed movements in accreditation and accountability. When discussing this research

and literature, the term evaluation is used. A note has been added when quoting literature where the above naming conventions may not be followed.

Empirical Research on Accreditation and Program Effectiveness

While there have been a decent number of people looking into assessment and evaluation in higher education from a theoretical standpoint, when looking into empirical studies on how accreditation influences an institution's—or a program's—effectiveness practices, the literature is quite scant. Indeed, as late as 2007, Volkwein et al. stated that, "Systematic studies of the impact of accreditation process on both changes in educational programs and student learning are, to our knowledge, non-existent" (p. 253). While this statement still generally holds true, there are a few different strains of research that have emerged related to accreditation and institutional effectiveness in roughly the last two and a half decades.

One of the strains of research has focused on the change towards outcomes-based evaluation in institutional effectiveness models that has been promoted by accrediting bodies, particularly regarding the perceptions and support of those changes amongst the various constituents in higher education. This movement towards outcomes-based measures of quality requires a greater collaboration between faculty and administration than inputs-based measures (Welsch and Metcalf, 2003b). Welsch and Metcalf (2003a)—following studies by Nichols (1995) and Birnbaum (2000) that suggested faculty resistance to institutional effectiveness measures was the biggest hurdle to their success—studied the influences of faculty support for such effectiveness practices at institutions involved in the SACSCOC self-study process. They found three primary ways that institutions can promote faculty support and participation in institutional effectiveness. The first finding was that framing institutional effectiveness efforts in the context

of its utility in making improvements to the institution (what Ewell (2008) has called an improvement paradigm), as opposed to fulfilling an external mandate like accreditation (what Ewell (2008) refers to as an accountability paradigm), to be the most compelling way of getting faculty support for the practice. Second, they found that engaging faculty in the institutional effectiveness process, giving them ownership of it, was more likely to increase support. Third, they found that faculty who believed in outcomes as the best measure of quality in an institution were the most likely to support efforts at institutional effectiveness; in other words, those who supported outcomes-based definitions as opposed to inputs-based definitions were more likely to support efforts at institutional effectiveness.

In another paper, Welsh and Metcalf (2003b) looked at the differing perceptions of faculty and administration towards institutional effectiveness practices. They found that while both faculty and administrators show overall support for institutional effectiveness, there was a statistically significant difference between the two with administrators showing a greater likelihood to support such practices. While calling for further research to explain the gap between administration and faculty, they suggest that faculty, who often go into their careers for a sense of autonomy, may just be less predisposed to take part in this type of work. As this study looked at effectiveness as it related to institutional efforts, they also suggest that the organized anarchy of higher education may make it less likely for faculty to want to take part in an effort that is not tightly coupled to their academic discipline.

Who buys in to the institutional effectiveness process has been a broader theme of research. When looking at the perceptions of administrators, faculty, staff, and students in the implementation of institutional effectiveness at a small private school, Bernecker (2010) found

no significant difference amongst the groups, in contrast to Metcalf (2003b). This was a very context-specific study, and it may be that the context of a small private school alters the dynamic of administrator and faculty workload and perception. However, Bernecker (2010) did find that administrators had a perception of a greater involvement in the institutional effectiveness practices than faculty, that faculty perceived a greater involvement than staff, and staff perceived a greater involvement than students. LoCascio (2010) in their dissertation found that full-time faculty members showed greater support for institutional effectiveness measures than part-time faculty. They suggest that part-time faculty's lack of training in institutional effectiveness may be the cause. At the same time, LoCascio (2010) found that faculty in the health sciences were more likely than others to place importance on institutional effectiveness, though no reasons for this finding were proposed. However, this supports the idea put forth by others (Ewell, 2009; Welsch & Metcalf, 2003b) that departments with programmatic accreditation and outcomes-based measures in professional practice may show greater support for implementing institutional effectiveness practices. Taken together with the findings of Welsch & Metcalf (2003b), we might deduce that faculty who are in fields that carry programmatic accreditation, or who have licensure and certification requirements tied to outcome measures, are more inclined towards effectiveness practice because they have seen it played out firsthand in their own academic units or work environments. They're both more familiar with it and have likely been better trained.

In their phenomenological study, Bush (2016) looked at the experiences of twelve individuals who bore the main responsibility for preparing their institutions' self-study reports. They found five major themes including: challenges with executive leadership, self-study process management, managing motivation and engagement, education and professional

development, and data management. Bush (2016) was particularly interested in the influence of executive leadership on the process. The importance of academic leadership is also reflected in other research (Kezar, 2013; Larkan-Skinner, 2015; Oden, 2009; Young, 2013). Bush (2016) also emphasized the importance of between-cycle preparation and the long-term cultivation of a culture of assessment, stating, "executive leadership who does not proactively advocate a culture of assessment during cycles when accreditation is not in the forefront, experience challenges in motivating or having people 'buy-in' to the assessment process when accreditation periods begin" (Bush, 2016, pp. 110-111). This statement seems a strong stance towards the improvement paradigm in evaluation and assessment and supports a utilitarian approach towards the overall process—that is, an approach that focuses on using what's found in accreditation self-studies to craft good effectiveness plans that are actually implemented between-cycle.

While these studies were looking at institutional accreditation, there have also been looks at effectiveness practices at the program level. Lewis (2007), in her dissertation focusing on using student learning outcomes as a method of improving dietetics education, found that while changes were instigated by the program director with input from data on student learning outcomes, the main motivators for change, in order of importance, were: (1) the desire to improve student learning outcomes, (2) to improve teaching and learning, (3) to increase students satisfaction, and (4) to meet requirements of their accrediting body. This study again supports the effectiveness of the improvement paradigm when compared to the accountability paradigm.

Overall, the research on attitudes related to program effectiveness changes would seem to suggest that while institution administrators feel a greater level of support and responsibility for institutional effectiveness measures, faculty buy-in is essential for overall institutional culture-

change (Procopio, 2010) and support of effectiveness activities. Faculty are motivated by calls for improvement in outcomes more than arguments for meeting accreditation requirements. Further, evidence suggests that faculty in fields that are accustomed to outcomes-based measures in practice—medicine and K-12 teacher education, for example—or programmatic accreditation are more likely to be engaged and find value in the process. This may be because they're more familiar with the practices of effectiveness planning and outcomes-based assessment and have had more training related to them.

This abuts with another thread of research, which is institutional or programmatic perception of accreditation's worth. Worth can be evaluated in a more subjective way, related to how constituents feel about the process, or whether the process appeared to engender positive emotions, as well as in a more objective way by looking at the actual costs of accreditation. Barczykowski's dissertation looked at programs accredited by the National Association of Schools of Music (NASM) and found that programs with the accreditation found it valuable to raise the credibility of the institution and that the standards positively influenced the quality of education offered (2018). Barczykowski (2018) also found that NASM participation resulted in the creation of comprehensive efforts to evaluate participants' programs in reflection of NASM standards.

Larkan-Skinner (2015) looked at community colleges accredited by SACSCOC and whether the recommendations given by the accrediting body led to institutional changes and improvements. Overall, Larkan-Skinner (2015) found that institutions who had received recommendations in their reaffirmation process showed greater levels of change. They also found that change was greater when an institution was at risk for negative sanction. The study

supported research that suggests institutions make use of their accreditation process to instigate quality improvement measures (Eaton, 2012; Head & Johnson, 2011; Oden, 2009; Young, 2013). Larkan-Skinner (2015) also pointed out barriers to accreditation, reflecting the themes already discussed around leadership and culture discussed by Bush (2016) when they found academic leadership unversed in the process. Larkan-Skinner (2015) also found a lack of assessment knowledge at the institutions, evidence of change, and quality data were all challenges faced in the accreditation process. Finally, this study found that increases in financial investments in the accreditation process led to a greater number of improvements.

On the theme of finances and resource commitment, there have been a couple of studies that directly speak to those themes. Lynch (2018) estimated the average ten-year cost of accreditation at community colleges in North Carolina to be $2.8 million. Academic programs that are programmatically accredited benefit not only from the overall support from the institution that all departments receive, they also, of necessity, have their own internal evaluation and assessment support resources that are devoted to carrying out their own programmatic accreditation requirements. Estimates by Heriot et al. (2009) for the average annual cost for Advance Collegial Schools of Business accreditation at $400,000. Those numbers represent faculty release time, faculty service time, and administrative FTE built into the unit's budget by the university that other programs on campus do not receive. It's no wonder that workload is brought up in discussions around faculty participation in assessment (Walton, 2022).

Theoretical Basis

Institutional theory helps to explain the environment within which program effectiveness and programmatic accreditation take place within institutions of higher education. Institutional

theory began in the mid-twentieth century. The basic premise of the theory—that institutionalized organizations (that is, organizations that have developed within fields to a certain level of maturity) were social organisms influenced not only by concerns around efficiency and effectiveness but also by the socialized construct of the institution itself (Selznick, 1957)—was elaborated by Meyer and Rowan in the late 1970s with an increased emphasis on the effects of environments on institutions (Cai & Mehari, 2015; Meyer, 1977; Meyer & Rowan, 1977).

Meyer and Rowan (1977) proposed that as institutions mature, their organizing structures and practices tend to conform towards conventions agreed-upon by their given field, and society at large, in a process labeled isomorphism. This process would explain why institutional organizations tend to develop similar organizing units, such as human resources, research and development, marketing, etc. In higher education, Morphew (2009) has shown how, since the 1970s, institutions of higher education have become more similar rather than keeping their distinctive characteristics, a good example of isomorphism at play. Isomorphic forms of organization and practice are often further ingrained in institutions by norms established around certification, licensing, or law that perpetuate them (Meyer & Rowan, 1977).

Accreditation fits into institutional theory in that the theory proposes a central tension between institutionalized organizations' tendency to rely on adherence to socially established convention for legitimacy and calls for meaningful evaluation of their performance (Meyer & Rowan, 1977). One of the ways that institutions overcome this tension is through ceremonial inspection and evaluation (Meyer & Rowan, 1977). Accreditation is often seen by administrators within schools as hoop-jumping (Ewell, 2008; Gillen et al, 2010; Barczykowski, 2018), for

example, a view which captures both the minimization and ceremonial tendencies of evaluation in institutionalized organizations (Rickards & Stitt-Bergh, 2016). Accreditation, as an externally validated method of evaluation, is also an organizational structure that is highly institutionalized itself, as institutional theory would predict (Meyer & Rowan, 1973).

As already shown, there are many accrediting bodies working in higher education that are approved by either the Council for Higher Education Accreditation (CHEA) or the U.S. Department of Education. Institutional theory helps to explain the isomorphic processes that have led them to be surprisingly similar despite the variety of programs that they oversee. This similarity has also been noted by Brittingham (2009). This isomorphism makes a study like the one currently undertaken possible—the similarity that exists allows there to be an examination of best practices independent of academic area of study. The structural components of programmatic accreditation are largely content agnostic.

In such an institutionalized setting, a belief in the foundational myths of the organization is important for its ongoing validity (Meyer & Rowan, 1973). This theory may help to explain the importance of faculty buy-in for the success of institutional effectiveness practices within programs and institutions. Patton first outlined the notion of the importance of the evaluator being "close" to the data, as important for some types of research paradigms (1975). This importance of personal connection to the evaluation is what he would later call the "personal factor": "the presence of an identifiable individual or group of people who personally care about the evaluation and the findings it generates" (Patton, 2008). This notion became a significant component to his utilization-focused evaluation (U-FE) approach. The most central tenant of U-FE, however, is the importance of the use of the evaluation. As such, it overlaps well with

accreditation, one of the main purposes of which is to develop a plan for institutional or program effectiveness that the institution or program uses as a process of constant quality improvement.

Thus, differences in success in accreditation, including success in approaches to institutional and program effectiveness, may lie more in the "personal factor" of the participants in the process than in any particulars related to the specific accrediting agency or the way in which the accreditation or effectiveness program is implemented locally. Indeed, Ewell (2008) warns against being wed to any particular assessment approach. Care in identifying and working with the people involved in the process, together with a thoughtful process, are likely the most important factors for success.

Programmatic Accreditation Examples

The Commission on Collegiate Nursing Education (CCNE) is one of two major accrediting bodies for nursing programs in the United States, the other being the Accreditation Commission for Education in Nursing (ACEN). While these are the two major accrediting bodies, there are also some specialty accreditations in nursing, like the Accreditation Commission for Midwifery Education (ACME) and the Council on Accreditation of Nurse Anesthesia Educational Programs (COA). It would be unlikely that nursing program would have degrees accredited by both CCNE and ACEN; however, it is common to have a nursing program have degrees accredited by either CCNE or ACEN and specialized degrees accredited by specialized accrediting bodies. The two nursing programs in this study had most of their degrees accredited by CCNE, so that accrediting body will be the focus of this section with regards to nursing.

CCNE states that their accreditation serves at least five general purposes:

1. To hold nursing programs accountable to the community of interest -- the nursing profession, consumers, employers, institutions of higher education, nursing students, nurse residents and fellows -- and to one another by providing that these programs have mission statements, goals, and outcomes that are appropriate to prepare individuals to fulfill their expected roles.

2. To evaluate the success of a nursing program in achieving its mission, goals, and outcomes.

3. To assess the extent to which a nursing program meets accreditation standards.

4. To inform the public of the purposes and values of accreditation and to identify nursing programs that meet accreditation standards.

5. To foster continuing improvement in nursing programs, and, thereby, in professional practice. (American Association of Colleges of Nursing, n.d.)

CCNE's 2018 Standards for Accreditation of Baccalaureate and Graduate Nursing Programs include four broad standards, with each standard including several "Key Elements" that programs must meet. The four broad standards are Program Quality: Mission and Governance, Program Quality: Institutional Commitment and Resources, Program Quality: Curriculum and Teaching-Learning Practices, and Program Effectiveness: Assessment and Achievement of Program Outcomes.

The structure of accreditation for engineering is different than that for most fields, due to its particular history. Established in 1932 as the Engineer's Council for Professional Development (ECPD), it was a conglomerate of seven separate engineering societies. Eventually, ECPD changed its name to the Accreditation Board for Engineering and Technology, Inc., before changing it once again to, simply, ABET. And while the engineering societies have grown to include 35 different members, ABET has four distinct accreditation commissions that accredit programs: Applied and Natural Science Accreditation Commission, Computing Accreditation

Commission, Engineering Accreditation Commission, and Engineering Technology Accreditation Commission (ABET, n.d.).

Each of the four commissions include the following eight standards: Students, Program Educational Objectives, Student Outcomes, Continuous Improvement, Curriculum, Faculty, Facilities, and Institutional Support. They also include Program Criteria for the accreditation of individual programs within the general commission. For example, The Computing Accreditation Commission standards includes Program Criteria for cybersecurity, computer science, data science and analytics, information systems, and information technology.

Attempts at Best Practice

In his qualitative, grounded theory study Bresciani (2009) examined the practices of 13 institutions around their process of program review of outcomes-based student learning. The institutions were chosen because they demonstrated efficient, effective, and sustainable outcomes-based assessment of student learning practices. The findings showed nine themes within these institutions: clear understanding of goals and expectations for program review, collaboration, use of results, awards and recognition for positive outcomes driven by faculty and staff, resources to support program review, coordination of the process, flexibility, addressing barriers, and evaluation of the program review process.

These findings focused exclusively on student learning outcomes. One of the limitations of the study was that it did not speak to efficiency of the practices or how sustainable they might be. With accreditation's effectiveness plan's focus on ongoing evaluation, these are important considerations when thinking about evaluation in that context. Also, extending the scope beyond student learning outcomes leads to the evaluation of areas like mission, facilities, and fiscal

support, that encompass a larger group of stakeholders and different evaluative focus. What constitutes best practice for program effectiveness in accredited programs is still largely unstudied.

Chapter Summary

This chapter outlined the important historical precedents that have led to the current state of accreditation and program effectiveness planning in U.S. higher education. It also reviewed empirical literature around accreditation and program effectiveness. Finally, it gave a history of the two accrediting bodies that were a major focus of the current research project. The next chapter will discuss the methodology employed for the project.

CHAPTER THREE

METHODOLOGY

There is very little in the way of empirical data that describes the experiences and best practices of practitioners of effectiveness plans within accredited programs. Acknowledging this lack of research, it's implied that practitioners' main source of knowledge around effectiveness planning is either handed down programmatically or derived experientially—or achieved through some combination of the two. As program evaluation and assessment continue their maturation into professions unto themselves, the body of research around the field will continue to grow. As it's presumed that much of the knowledge specific to program effectiveness within programmatically accredited academic units most likely lies with the practitioners, a study of those people seems like a good place to contribute to that body of research.

All research activity is framed within metaphysical constructs that Guba and Lincoln (1984) have called paradigms. Qualitative research tends towards an approach that, axiologically, is explicit in expressing the researcher's position and relationship towards a topic of study. Ontologically, qualitative research tends to be constructivist in its approach to reality and, epistemologically, constructivist about how we know about that reality. That is, qualitative research tends to approach research with a lens focusing on the way humans understand and create meaning in the world (i.e., "construct" reality and their knowledge of it) than to approach the topic of study as a wholly objective, knowable reality independent of those who observe and participate in it. The latter, positivist approach to knowledge, is more common to quantitative research. Finally, qualitative research, methodologically, is inductive and emergent, and heavily shaped by the experience of the researcher in the process of research (Creswell & Poth, 2018).

Understanding the paradigm of qualitative research and its philosophical assumptions, we can say that qualitative research is a methodology that attempts to address "the meaning individuals or groups ascribe to a social or human problem" (Creswell, 2013, p. 44). Interpreting the meanings people make in their real-world settings allows the researcher to make sense of the topic of study through the lens of human experience—both the experience of people related to the topic studied, but also of the researcher (Denzin & Lincoln, 2011). Indeed, most qualitative approaches to research shine a light on the subjective stance of the researcher through careful attention to reflexivity and positionality (Creswell & Poth, 2018). Further, the nature of practices around program effectiveness and accreditation are quite contextual and process-dominant: qualitative approaches to understanding can help to explore deeper how individuals experience these processes within their specific contexts.

In this study, the aim was to understand how practitioners engaged in program effectiveness practices within programmatically accredited academic units make sense of the accreditation process and program effectiveness practices. The study used a grounded theory approach with multiple bounded case studies. Participants in the four case studies were from the disciplines of nursing and engineering in two separate institutions of higher education. The guiding research questions for the current study are:

RQ 1) How do practitioners at two western, doctoral-granting R-1 institutions of higher education understand and experience their programs' formation and implementation of their program effectiveness plans?

RQ 2) How has the context of institution and discipline influenced practitioners' program effectiveness practices?

RQ 3) How do these practitioners' experiences and knowledge contribute to an understanding of generalizable (Ragin & Becker, 1992; Stake, 2006; Yin, 2003) best practice in program effectiveness?

Significance of Study

It is hoped that this study will illuminate best practices related to programmatic accreditation, specifically as it relates to the requirement for ongoing program evaluation as seen in effectiveness plans. As a resource-intensive undertaking, programs would benefit from research to offer confirmation of current practices or serve as a guidepost as they make changes. This study will be of particular importance for programs that are already, or are seeking to become, programmatically accredited. However, due to many similarities in structure and process amongst the various types of accreditation and accrediting bodies, it may be of interest to those working in institutional accreditation as well.

Research Design and Approach

This study, following the methodology of Bresciani (2009), employed grounded theory with multiple bounded cases of practitioners' experiences of program effectiveness practices within their individual academic units. Multiple bounded cases allowed for deep dives into each practitioner's experience with program effectiveness and accreditation practice, while the grounded theory approach allowed for theorizing outside of the contextual constraints of each individual case. Glaser and Strauss (1967) first pioneered the grounded theory method as an inductive way of generating theories that surfaced through the careful collection and analysis of qualitative data. This approach departed from the predominant practice of their time wherein

studies were often designed deductively to test hypothesis derived from existing theories. The original approach derived by Glaser and Strauss has proven to be quite adaptable since its inception and is now used in ways more flexibly than they had initially proposed (Charmaz, 2006; Stake, 2006; Strauss, 1987). This current study's philosophical attitude towards the use of a grounded theory was influenced by Charmaz (2006), whose constructivist approach holds that "[r]esearchers can use grounded theory strategies with a variety of data collection methods" (p. 10). And, while a more traditional approach to grounded theory might call for a larger number of participants than the multiple case study approach employed here (see Creswell & Poth, 2018, p. 150), even Glaser and Strauss in their seminal work on grounded theory state that, "the kind of evidence, as well as the number of cases, is also not so crucial" (1967, p. 30). Grounded theory is the underlayment of the methodological approach to understanding and thinking about the data in the study.

The phenomenon of the study that bound the different cases, what Stake (2006) would call the quintain, was the experience of program effectiveness practices as experienced by practitioners in the field. Yin (2003) suggests that case studies begin with theory and propositions. This study began with institutional theory and the belief that the field of evaluation and assessment in higher education, generally, and the field of programmatic accreditation, specifically, has developed along institutional lines to be isomorphic and ceremonial in practice (Meyer & Rowan, 1977). If the theory holds true, the resulting proposition is that program effectiveness expectations amongst varying accrediting bodies and fields of study should be convergent enough that they can be examined on their own as isomorphic systems, independent of the specific context of academic discipline and institution. Indeed, a study of accrediting

bodies shows that this appears to be true (Brittingham, 2009). The development of program evaluation and assessment as fields of study and professions unto themselves (Ewell, 2011; Rickards & Stitt-Bergh, 2016) supports the isomorphic tendency propounded by institutional theory. That is, if we are developing assessment and evaluation practitioners who are experts in that field, and not necessarily the one in that they are evaluating, there is something unique to assessment and evaluation practices that surely transcends the content of what is being evaluated. Thus, holding the proposition of the isomorphic quality of program effectiveness to be true, this study employed what Yin (2003) has called "replication logic" in its choice of cases. Replication logic, according to Yin, is not to be confused with "sampling logic," where the goal is to find a representative sample. Instead, replication logic seeks to choose cases such that the context and number of cases chosen fulfills the goal of replication reasonably so that a researcher can establish "analytic generalizability" (Yin, 2003).

The use of multiple bounded cases provided other benefits as well. Noting the lack of empirical foundations in extant literature, another proposition of the study was that most of the knowledge of effectiveness practice resides with the practitioners and the systems they have established or inherited. Another benefit is that by focusing on two institutional contexts and two academic disciplines, this study was also able to provide thick description (Creswell & Poth, 2018) of the practitioners within the contexts where it is proposed that practical knowledge on the topic resides. The multiple cases also allowed for data triangulation amongst practitioners in different contexts.

Positionality

My first exposure to evaluation in an educational context was working as the manager at a branch of a franchised private English language school in Beijing. This early exposure to private industry, and private education, was foundational to what became an ongoing professional focus on data driven decision making, evaluation, and process optimization. This early experience in educational administration was very focused on the continuous pursuit of measurable outcomes, with monthly, quarterly, and yearly targets being of central importance. Returning to Bozeman I found work with MSU's intensive English program. In that position I was a part of the program's initial accreditation approval with the Commission on English Language Program Accreditation (CEA), helping to gather data and write portions of the report related to student progression and success. This was my first experience with accreditation and measuring student success, but that experience has proven to be a central part of my professional life ever since.

After the writing of the initial accreditation report, I played a central role in establishing the program's internal, ongoing evaluative practices that were developed and implemented over the next several years. In the role of director of the program three years later, I was responsible for all aspects our three-year continuing accreditation report with CEA. This process of going from establishing practices for an initial accreditation to implementing them over time and writing a continuing accreditation report drove home the importance of the effectiveness planning process that we had devised. The difference between the challenges of the initial-accreditation report and site visit verses the continuing-accreditation report and site visit were striking. The well-designed plan for program effectiveness that we had implemented made the

latter much easier than the former. It also seemed to me that it helped our program to both focus our efforts and enabled us to make data-informed decisions.

At the same time, another realization that was garnered from the process was that we, as a program, were constantly questioning whether the systems that we were developing were the "right" way to do things. Our guide was the accreditation standards that we were trying to meet, but accreditation standards are often intentionally non-prescriptive in how you meet them. While this allows for a great deal of flexibility to account for the variety of programmatic and institutional contexts we see in U.S. higher education, it can also leave those who are working on a program's assessment and evaluative practices in the dark as far as acceptable practice in the field. This is compounded when those practitioners come from a discipline-specific background and not an assessment and evaluation background.

After a successful continuing accreditation cycle with CEA in the intensive English program where I worked, I took a position with the College of Nursing with program evaluation—couched in the context of the College's programmatic accreditation—serving as a central part of my role. Soon after my hire, I led an overhaul of the College's program effectiveness plan and, together with the Associate Dean for Academic Affairs, am largely responsible for the plan's ongoing implementation and tracking. I've also been involved in the development of a revamped Teaching/Learning Assessment Plan within the College. My contribution to these efforts has been largely influenced by my earlier professional experiences, but the plan for program effectiveness at the level of an academic College is significantly more complex, particularly with regard to the increased number and types of constituents involved in the process. Relational and political considerations are much more important and thoughtful

communication seems to be key in garnering buy-in. It has been interesting to see that the importance of relationship building and garnering the buy-in of constituents has shown up consistently in the literature around program evaluation and assessment (Miller, 2013; Patton & Campbell-Patton, 2022; Rickards & Stitt-Bergh, 2016; Walton, 2022; Welsh & Metcalf, 2003b).

Far from changing my general attitudes towards effectiveness planning, this experience has only solidified my belief that good planning not only helps accreditation efforts but leads to better program-level decision making and improvement. In fact, I now see the benefits of effectiveness planning to meet accreditation requirements—what Ewell (2008) refers to as the "accountability paradigm" of assessment—to be ancillary, while I see the benefits for program improvement and evaluation—what Ewel (2008) calls the "improvement paradigm"—as primary. Further, I have grown increasingly aware of the importance of building utility into the establishment of the plan for program effectiveness—whatever approach to planning that is adopted, if the plan is not structured in a way that lends itself to ongoing use there is little hope of effecting ongoing, positive systemic change across all levels of the program, an attitude reflected in Patton's utilization-focused evaluation method (Patton & Campbell-Patton, 2022).

Bringing the experiential knowledge that I have with the practices of program effectiveness and accreditation, a social constructionism lens will be used to co-construct and make meaning of the participants' responses. My own positionality is affected by my professional role where presenting data with impartiality before making inferences as to meaning is central to the process I use to gather trust; this influence shows in my overall approach to this study, where my findings in Chapter Four are presented mostly "as-is," and my co-construction of meaning is developed primarily in Chapter Five. With deep experience in the topic of study,

there comes the threat of too-readily aligning participants' experiences with my own, which is another reason why I've presented the findings in a way to try to let the data stand for themselves. At the same time, as Dahlberg (2006) points out, bracketing is not really possible—even more particularly so with my insider positionality. And so, all attempts were made to bridle my understanding of the phenomenon to better understand participants' experiences and not to, "understand too quick, too careless, or slovenly" (Dahlberg, 2006, p. 16).

Context of the Study

Tracing its roots back to the late 19th century, the field of accreditation has proven to be quite adept at adapting to changing public and academic pressures to keep—and at times even to elevate—its central role as the main public accountability system in higher education in the United States (Eaton, 2012; Ewell, 2008; Ewell, 2009; Hartle, 2012). Over the last 30 years, one of the main ways in which accreditation has adapted to keep that role has been with a shift towards emphasizing the central importance of effectiveness and assessment in accreditation standards and moving away from more prescriptive, quantitative measures of quality (Brittingham, 2009). The shifting accreditation focus has, at times, been accompanied—or even prompted—by accountability regulations promoted by state governments (Ewell, 2009). And indeed, there has been a push and pull between state government as the primary drivers of accountability and accreditation as filling that role (Ewell, 2009; Ewell, 2011). Presently, however, accreditation has largely subsumed state accountability efforts and still holds as the primary accountability mechanism in higher education (Brittingham, 2009; Ewell, 2009; Hartle, 2012).

This is not to say that accreditation lives in an accountability paradigm outside the realm of governmental influence. As already alluded to, much of the more recent shift towards effectiveness planning and program assessment has been a reaction to state governments' attempts to hold institutions of higher education accountable (Ewell, 2009). And, while state governments have no direct say on how accrediting bodies police the institutions they accredit, the federal government is quite involved in the process by the direct sway it holds over the U.S. Department of Education (USDE). For institutions of higher education to qualify for Title IV funding, something virtually indispensable to any institution (Brittingham, 2009; Hegji, 2017; Gillen et al., 2010), they must be accredited by a USDE-approved accrediting body. This mechanism has allowed the federal government to channel the direction and focus of accreditation and accountability in higher education (Ewell, 2009).

Added to the largely institutional accreditation context just cited, several fields of study within higher education take part in programmatic accreditation. While some do this for reputational strength, many are essentially required to partake in programmatic accreditation due to licensure or legal requirements for program graduates to practice in the field (Overbay & Aaltonen, 2001). While it is hoped that all academic programs take part in some sort of systematic evaluation of the work that they do—even if only driven to do so by institutional requirements—those programs that are accredited are required to do so to maintain their accreditation. Thus, programmatically accredited units are at an interesting intersection of evaluation in higher education where not only is evaluation expected, it is required; and that evaluation itself is subject to external evaluation periodically through the re-accreditation process. At the same time, these programs' practices need to fit within institutional evaluation

practices that set standards for the institution at large and fit within the institutional accreditation context. Thus, while all academic units at institutionally accredited institutions no doubt take part, to some degree, in the program effectiveness practices required by the institution, it stands to reason that plans for program effectiveness developed by programmatically accredited academic units, considering their additional accountability burdens, would be more purposeful and stand to greater scrutiny than those academic units that do not face the same external oversight. As such, programmatically accredited academic units, and the people within those units involved in program effectiveness planning, become solid units of analysis for best practices in program effectiveness in higher education. Most likely, they are the ones who serve as pioneers at their institutions at large for practices that units not facing the same scrutiny might seek to emulate.

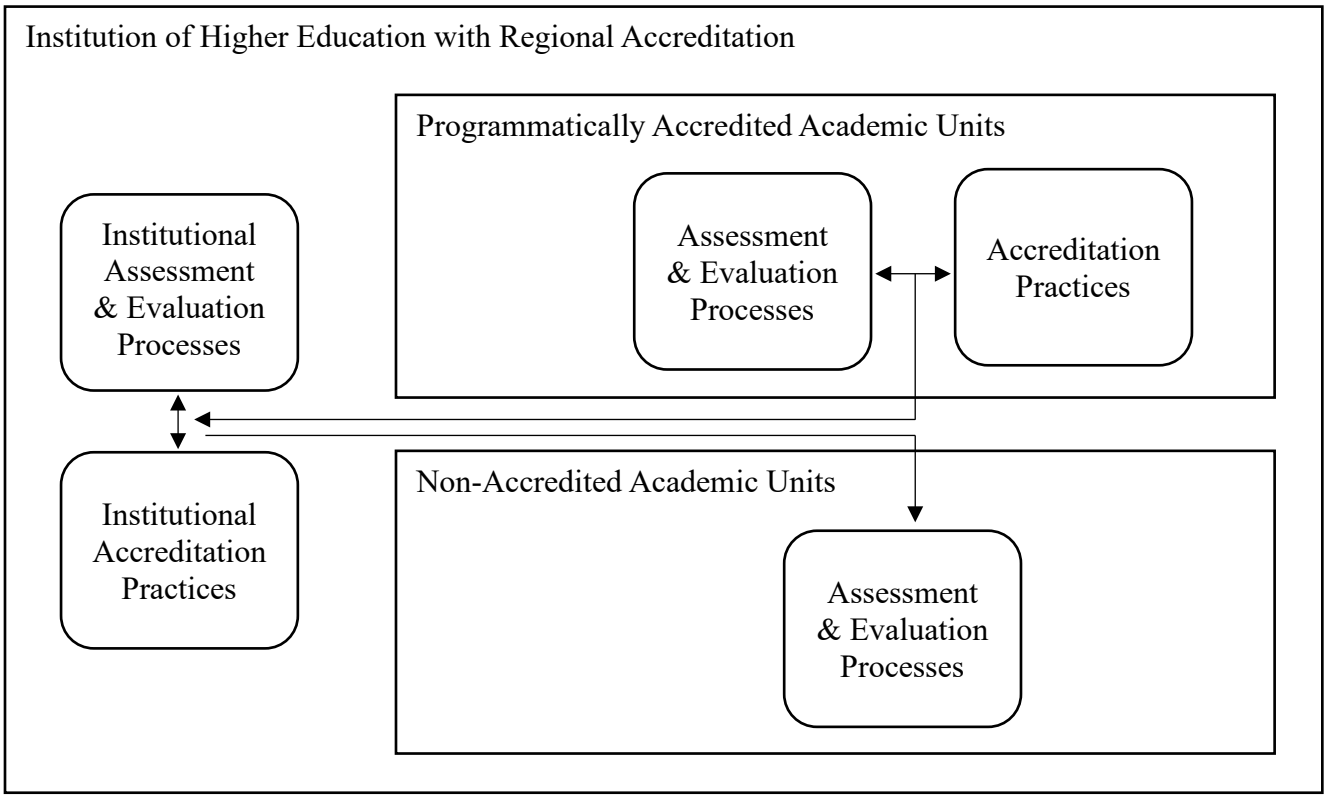

Figure 3. Ecosystem of Accreditation in Higher Education

Constructs

The constructs for this study employed those outlined in the initial conceptual framework in Chapter One. How the notions of stakeholder buy-in, communication, plan structure, and utility interplay within accredited programs was interrogated in the interview protocol, as can be seen below. These broad constructs also served as a starting point for the analysis strategies employed in the coding process.

Population, Sample, and Sampling Strategy

Programmatic Fields of Study

For this study I was interested in gathering the experiences of practitioners of program effectiveness to gain experiential knowledge about best practices. My own experiences suggested that program effectiveness can lead to a variety of positive changes in outcomes for programs, but the knowledge I had gained around the structure and implementation of program effectiveness planning had largely been through personal trial and error. Further, the lack of research in this area, generally, suggested that those creating and implementing plans for program effectiveness have most likely also relied on some combination of experiential learning and inherited practices to arrive at their knowledge as practitioners in the field. The thought was that by engaging at a deep level with common practitioners about their effectiveness practices, themes may emerge that could be of help to current and future practitioners of program effectiveness. And so, this study engaged practitioners in two separate fields—engineering and nursing—in two separate institutions, to better understand their experiences with programmatic accreditation and effectiveness practices.

To gain multiple perspectives from practitioners in the field of program effectiveness, this study chose to look at two separate areas of academic inquiry that operate under programmatic accreditation—nursing and engineering. Both fields have a long history of participating in accreditation. Nursing accreditation can trace its roots back to 1893 with attempts to establish standards for nurse training and, even more formally, to 1917 when the National League of Nursing Education published its *Standards of Curriculum for Schools of Nursing* (American Association of Colleges of Nursing). ABET, the accrediting body for engineering, was founded

by seven existing engineering societies as the Engineers' Council for Professional Development in 1932 (ABET, n.d.). Both fields have decades of experience working with accreditation and the accrediting bodies have had significant time to mature.

These fields share some commonalities in the way that their accreditation plays a role in the professional practice and graduate education of their graduates as well. Graduates in both nursing and engineering require or encourage licensure to practice; also, pursuit of advanced degrees in their fields typically would require their baccalaureate work to have been completed within an accredited program. As part of their accreditation reporting, they both also report on the results of summative, high stakes testing that their students take part in. These fields' length of engagement with accreditation, as well as the fact that their accreditation status is generally used as a minimal bar to enter their professional fields, suggests that these disciplines are more likely to have developed mature effectiveness practices than those practicing in disciplines where accreditation is truly more voluntary or whose programmatic accrediting bodies have not been in place as long.

The sampling strategy for this study was primarily criterion-based: all participants were practitioners of program effectiveness either in nursing or in engineering. Having defined the two disciplines to study, I chose the two institutions to study, starting at my own institution. This choice was made in the hope that proximal factors unique to one institution—things like organizational structure, institutional assessment requirements, the localized culture of assessment—might be exposed when contrasted with the other. There are ties between the nursing program at Institution 1 and Institution 2, so I was able to reach out via a colleague to Jennifer with the nursing program at Institution 2. Jennifer also happens to have deep

experiences working with CCNE at very high levels—she was able to present a perspective that drew from considerably more experience than any of the other participants. After meeting with Jennifer, she was able to connect me with Casey in nursing at Institution 2, and also the Associate Dean in engineering, Charlie. After our interview, Charlie then suggested I work with Mark in the mechanical engineering department.

Coming back to my home institution, there was an opportunistic opportunity of sampling that I took advantage of after having a conversation with a colleague in electrical and computer engineering about my research, Bill. It just so happened that he was the main person responsible for program effectiveness within his department. That conversation snowballed and I ended up not only interviewing Bill, but also the Department Head for his unit. After the interviews with electrical and computer engineering at Institution 1 I reached out to Justin in mechanical and industrial engineering to give a parallel to my work with that department at Institution 2. Unfortunately, I wasn't able to have completely perfect parallelism between positions interviewed within the two institutions as I would have like. Specifically, I was unable to secure participation from two desired participants: the individual responsible for program effectiveness in electrical engineering at Institution 2 and the Associate Dean for Engineering at Institution 1.

Institution 1

Institution 1 is a mid-sized, western, doctoral-granting R1 institution that is regionally accredited by Northwest Commission on Colleges and Universities (NWCCU). The flagship, land-grant institution of the state, the College of Engineering is one of the largest on campus and is the College for which the institution is best known. It houses several engineering disciplines as well as computer science within its academic ranks, offering bachelor's, master's, and PhD

degrees. The undergraduate degree programs in engineering are accredited by varying commissions of ABET, an accrediting body for engineering and computing programs.

The College of Nursing at Institution 1 is the oldest and largest in the state. It educates students in the last two years of the bachelor's degree both on the main campus and at satellite campuses throughout the state to capitalize on the clinical capacity at some of the state's larger cities. It also offers master's, and Doctor of Nursing Practice (DNP) degrees via distance-delivery. Its degree programs at both the undergraduate and graduate levels are accredited by the Commission on Collegiate Nursing Education (CCNE).

Institution 2

Institution 2 is a large, western, highly ranked, doctoral-granting R1 institution accredited by NWCCU. The College of Engineering offers bachelor's, master's, and PhD degrees in variety of engineering and computing disciplines. Its degree programs are accredited by varying commissions of ABET at the undergraduate level. The School of Nursing offers bachelor's, master's, PhD, and DNP degrees at three separate campus locations. Most of its degree programs at both the undergraduate and graduate levels are accredited by CCNE with its Nurse-Midwifery DNP and Graduate Certificate programs being accredited by the Accreditation Commission for Midwifery Education (ACME).

	Size	Carnegie Classification	Governance Structure	Regional Accreditor
Institution 1	17,000	R1	Public	NWCCU
Institution 2	33,000	R1	Public	NWCCU

Table 1. Institution Characteristics

	Founded	No. of Faculty in Program	No. of Students in Program	No. of Administrators in Program
Institution 1 - Nursing	1937	130	1,200 1,000 UG	• 2 Associate Deans • 3 Program Leads • 4 Academic Programs Staff • 2 Advisors
Institution 1 – Mechanical Engineering	1893	35	1,200 UG	• Department Head • 2 Academic Advisors • 1 Office Manager • 1 Receptionist • Shop members
Institution 1- Electrical & Computer Engineering	1893	17	350 UG	• Department Head • 1 Administrator • Lab manager
Institution 2 – Nursing	1920	130	1,000 500-600 UG	• 3 Associate Deans • 3 Program Directors • 20 Staff in Student Academic Services
Institution 2 – Mechanical Engineering	1903	65	500 UG	• 1 Department Chair • 2 Associate Chairs • 1 Accreditation Coordinator • 3 UG advisors • 10 Fiscal staff • 3 Lab support • 3 Lab techs

Table 2. Program Characteristics

Interviewees

Pseudonyms were used for all participants for confidentiality as per this study's IRB protocol.

	Title	Role in Process	Years in Program	Experience
Institution 1				
College of Engineering				
Mechanical & Industrial Engineering				
Justin	Professor, Undergraduate Curriculum Coordinator	Department champion for ABET accreditation	28 years	• MEd coursework in program assessment and evaluation • Four ABET cycles • ABET evaluator
Electrical & Computer Engineering				
Kyle	Professor, Department Head	Assigns faculty, responsible for non-curricular aspects of self-study report	6 years in current role, 23 years at institution	• Submits ABET report • Oversees accreditation process
Bill	Professor	Department champion for ABET accreditation	22 years as faculty, 7 years as lead for assessment	• 7 years as assessment lead • ABET workshop attendance
College of Nursing				
Kathy	Associate Dean for Academic Affairs	College champion for CCNE accreditation	5 years in current role, 18 years in leadership roles with College prior	• 1 CCNE accreditation cycle • Working on CCNE CIPR report • 1 ACME initial accreditation cycle
Institution 2				
College of Engineering				
Charlie	Associate Dean, Office of Academic Affairs	Organize and support ABET accreditation at the college level	3 years	• 3 rounds of ABET accreditation with home department • 1 round of curriculum coordinator in home department

				• Not yet taken full program through ABET accreditation
Mechanical Engineering				
Mark	Associate Chair for Academics	Departmental champion for ABET accreditation	32 years as faculty, 23 years with accreditation	• First experience with ABET in 2001
School of Nursing				
Jennifer	Former Associate Dean for Academic Affairs	School champion for CCNE accreditation	6 years	• Experience with accreditation since 1980 • Associate Dean at four different institutions • Involved in eight accreditation cycles with various institutions • Site visitor for CCNE • Has Team Led 15 site visits • Two terms on CCNE accreditation review committee • Co-chair of CCNE accreditation review committee for five years • Review of roughly 1200 self-studies with CCNE • Co-chair of CCNE accreditation standards revision committee
Casey	Director of Student and Academic Services	Support of CCNE accreditation processes	6 years in current role, 17 years total in School of Nursing	• Several accreditation cycles and CIPR reports • MEd coursework in program assessment and evaluation

Table 3. Participant Characteristics

Data Collection

Yin (2003) discusses six different types of what he calls evidence that may be collected in case studies. For this study, the primary sources of data included two in his list: documents and interviews. For each of the fields studied, a close analysis of the accreditation bodies' websites and the accreditation requirements was undertaken. I primarily interrogated the accrediting bodies' websites for information about the history and development of the accreditation requirements. I also looked at them with the questions of, "what is this accrediting body's overall position on the role of accreditation?" and, when comparing with their accreditation requirements, "is there a discrepancy between what they say is their stated position and their actual accreditation requirements?" When looking at the accreditation requirement documents, the main question was, "what do these requirements as written actually ask a program to do to meet them?" Another question when looking at the accreditation requirement documents was, "how are the ABET requirements different from the CCNE requirements?" Overall, these documents served as a way to corroborate the data collected via participant interviews and to fit them within the context of their accreditation requirements (Yin, 2003). It also allowed for a detailed analysis of the discrepancies between stated objectives of accreditation and the actual requirements that became a theme when exploring the differences between the two accrediting bodies approaches and the two disciplines' effectiveness practices.

Open-ended participant interviews were the second major source of evidence for this study. These interviews were conducted via Webex and recorded on the Webex platform according to a prescribed interview protocol. The protocol focused on six main areas: participant background, program background, the effectiveness plans' formation and creation, the

effectiveness plans' implementation, institutional context, and best practice. Each question of the protocol was also mapped to both a research question, as well as one or more of the constructs from the initial framework that I was working with at the time. The protocol can be seen in the Appendix.

The process of gathering this data was iterative. I spent time moving from website or accreditation requirement analysis to interview participation, and back to the websites and accreditation requirements. For example, I already had a thorough understanding of the accreditation standards of CCNE before beginning my study, but I spent time reading through the ABET accreditation criteria prior to beginning my first interview. One of my questions was around the distinction that ABET makes between their general criteria and their program criteria. I was curious as to how that actually played out in practice and whether or not the process was centrally managed or managed in a distributed model by the individual programs. I could envision a possible scenario where the general criteria were managed at the college level and the program criteria were managed at the department level—was that how the accreditation process worked in the cases that I was analyzing? Another example was when, after a few interviews in both disciplines, I dove much deeper into the specific wording in each accrediting bodies' effectiveness requirements to try to understand the disciplinary distinctions I was finding in my participants' responses. This was an area where my experiential understanding of how accrediting bodies had dealt with effectiveness requirements had blinded me to a stark difference with the way that ABET had outlined their requirements. I hadn't seen the difference in my initial reading of their requirements, and it was the results I was seeing in my interviews that directed me back to a deeper textual analysis.

Data Analysis Strategies

Grounded theory is appropriate when the research is centered around a process with defined steps (Creswell & Poth, 2018). Plans for programmatic effectiveness are, by their definition, processes that programs define with steps to follow. Following the grounded theory approach outlined in Creswell and Poth (2018), the transcripts of the interviews were initially coded using an open coding process, line by line through the transcribed interviews. While the constructs presented in the initial conceptual framework of stakeholder buy-in, communication, plan structure, and utility presented themselves as obvious segments of data due to the way my protocol was structured, care was taken not to force the data since in open coding, the researcher is looking to find the main categories present in the data to allow the researcher to think deeper about the phenomenon (Strauss & Corbin, 1990). In fact, my analysis resulted in a departure from my initial theoretical framework and led to a more refined framework that I present in Chapter Five.

The second round of coding in grounded theory is axial coding, where connections between the categories defined in the open coding process are examined (Creswell & Poth, 2018). The researcher then returns to the data and re-codes it with categories based off of the core phenomenon, developing an axial coding paradigm (Strauss & Corbin, 1990). In the final step, selective coding, the researcher unifies the coding around a core category, or central phenomenon (Strauss & Corbin, 1990). In the next section I will describe my process of coding and the theme exploration that took place during my research.

Throughout the process of data gathering and analysis, care was taken to remain reflexive about what I was hearing and how I was interpreting what I was hearing. Some of the questions I

managed were: how does what I'm hearing not align with my understanding of effectiveness practice; am I painting a picture that just matches what I want the outcomes to look like; and would someone without experience in the field come to the same conclusions when looking at the data? This reflexivity helped me when I at times found myself trying to force the data and led to what I believe is a much more trustworthy representation of phenomena presented in the data.

Coding and Theme Exploration

The preliminary steps of initial coding for this study took place during the interview process. While following the interview protocol already described, I took notes while listening to the participants, highlighting quotes or thoughts that resonated with my own experience or things that seemed particularly intriguing or worthy of further inquiry. Examples of notes taken during the interviews that were pertinent to the initial coding were: "somebody at the helm," resource commitment, documentation, faculty involvement, "wet blanket," "reasonably user friendly," "we do what they want," fidelity vs feasibility, "we would want to do this anyway," and "ABET is prescriptive." I also used a process of analytic memoing throughout each stage of the interview and coding process to help me sort what I was seeing with the emerging data and to think about how the codes might relate to each other, or how the various contexts that I was witnessing were similar or different. As these ideas were generated through the note taking and memoing process, I was able to home in and focus on them in subsequent interviews, or to take them back to previous interviews to see whether or not they were reflected in those data.

One example of this process was the development of an important theme, which was that accrediting bodies' requirements have a large influence on how the programs in my study envisioned their effectiveness practices and there was a big difference between the ABET

requirements for engineering and the CCNE requirements for nursing. This was something that took time for me to discover. I had already finished my interviews with engineering at University 2 when, in my first interview with a participant in engineering in University 1, I had the realization that there was something fundamentally different in the way engineering viewed the scope of their effectiveness process when compared with nursing. Early in this interview, I wrote down a quote in my notebook that I referenced above: “we do what they want,” that referred to the fact that this participants’ effectiveness practice aligned with what ABET requires of them. Later, looking at the transcript of this interview, I can almost see when I had the full realization that what ABET wants was quite different than my experiences with accreditation, and much different from what I was hearing with the nursing programs. After a lengthy description of their process for assessing student outcomes, it seemed like something significant was missing from their process when compared with my own experience and the experiences I was hearing from nursing. I asked Bill about what I called in the moment the “periphery parts of accreditation.” I went on: “things like adequacy of your facilities, support of the institution, faculty preparedness—things like that. Does that fit into your continuous improvement plan?” When it became apparent that these things did not fit into the plan, it led me to both revisit ABET’s accreditation criteria, but also to revisit the engineering interviews that I had completed at Institution 2 with an eye to noticing the difference that I had become aware of between engineering and nursing in this later interview. This realization also shaped the two future interviews I was to take part in with engineering participants in University 1 and flesh out the theme a little more.

This realization and the revisiting of previous interviews, combined with an iterative memoing process, led to the development of the code “Culture of Assessment” in the open coding stage. In my final analysis, which reflected also drafts of writing and discussions with the chair of my committee, this code eventually split into two distinct aspects: first, when combined in the selective coding stage with some of the other major themes that emerged, it helped to drive the Culture of Evaluation Model in Engineering and Nursing that is found in Chapter Four; second, some of the quotes that I initially coded as “Culture of Assessment” were moved in the axial coding phase to a code dealing with best practice: “Develop a Culture of Evaluation.” This code dealt more with the themes around intentional culture development that I saw, as opposed to the quotes that I coded as showing a distinction between the different types of culture in nursing and engineering that led to the model mentioned above. The change in word choice from assessment to evaluation also reflected a fine tuning of my own use of terms that I developed through the research process and that I speak to at greater length in Chapter Five.

This example shows how my code and theme development started early with my note taking and memoing during the interview process itself and shaped my thinking very early in the research. This more informal process eventually intertwined with the more formal process of initial coding during the transcription stage, but the latter was informed in some significant ways by the former. As for the formal transcription phase, after my interviews were complete, I uploaded the recorded transcripts produced by Webex into NVivo and listened to the interviews. There were many transcription errors that I corrected by watching the recordings and editing the errors in the file in NVivo. I then began the open coding process, which generated 56 unique codes.

During the axial coding stage, I was able to group my initial open coding themes into 31 distinct themes before discarding five that had no relevance to the current study, arriving at 26. I made use of NVivo's hierarchy chart functionality to group my codes and visually explore how they were related and to see what themes were expressed with the greatest frequency. I also continued my process of using analytic memos in this stage, pairing it with a number of conversations with colleagues, family, and my committee chair, helping me to focus on what was important in the data presented by these practitioners and to group the data accordingly. One example of an axial coding grouping that developed was one related to the open codes that I had of "Innovation," "Intuitive Change," and "Alternatives to Accreditation." These codes were all related to the processes that programs used—whether formal or informal—when making change. In this round of axial coding—again, in constant interaction with my memoing, the discussions I was having about my data, and the drafting process as I started writing Chapter Four—I ended up subsuming them all into the code of "Change Process." This code joined with another open coding stage code that I had labelled "Process" in my selective coding stage and operated as a theme that distinguished between approaches to change in the two disciplines studied with "ad hoc approach to change" for engineering, and "systematic evaluative approach to change" for nursing as the final themes.

The selective coding stage developed primarily out of my process of writing as I began drafting Chapter Four. While all stages were influenced by my positionality as an insider with 12 years of experience as a practitioner in the field, this stage was perhaps influenced the most as I began to unify the codes that had emerged into thematic units. As I began writing around the themes that emerged in the axial coding, I focused most heavily on developing a list of best

practices, with the themes of support of organizational structures and resources, plan structure and documentation, feasibility, having champions & taking the time, developing a culture of evaluation, thinking about succession, and the importance of metaevaluation. Influenced by some conversations I continued to have about my research—but most significantly by feedback from, and conversations with, my committee chair—the selective coding solidified in my later drafts around two central phenomena: a model outlining the distinct cultures of evaluation in nursing and engineering as well as generalized best practices for practitioners in program effectiveness. The model explored more deeply the relationships and differences in themes that I had found in nursing and engineering and theorizes influences that contribute to the distinct cultures of evaluation that I found in the two disciplines. The best practices for practitioners in program effectiveness pulled from themes that emerged from the interviews with my participants that also resonated with my own experience to create a sort of roadmap for individuals involved in effectiveness practice.

Trustworthiness & Limitations

There are a number of different approaches to trustworthiness in qualitative research (Creswell & Poth, 2018) although, essentially, it is a similar notion to what is described as validity in quantitative research. The question is how do we know that the research accurately describes what was observed, and to what extent can we generalize any of the findings. Lack of rigor has been used to criticize qualitative research (LeCompte & Goetz, 1982). At the same time, some qualitative researchers have pushed back on that quantitative-paradigm-focused criticism and eschewed the term of validity, instead preferring to define qualitative research in terms of its own use that are not shared with quantitative research (Anzul, et al., 2003). Lincoln

and Guba (1985) list trustworthiness, credibility, authenticity, transferability, dependability, and confirmability as methods used in qualitative research to gain validity. Wolcott (1990, 1994) goes so far as to practically break with the concept altogether. For this study, the term trustworthiness is used. In the socially constructed perspective used in this study, validity—with its claims of true objectivity—seems inappropriate. There was a conscious and reflective process by myself to take part in the construction of the meaning that I found.

Whittemore et al. (2001) list 29 strategies; Creswell and Poth (2018) list nine commonly used strategies and recommend that researchers use at least two approaches. The nine listed by Creswell and Poth (2018) are: triangulation, negative case analysis, reflexivity, member checking, prolonged engagement in the field, collaborating with participants, external audits, thick descriptions, and using a peer review. For this study, trustworthiness was increased by the use of thick descriptions of the context and experience of the participants. By describing their pertinent experience and the setting in which they operate, as I've done above but flesh out more in Chapter Four, it's hoped that others can find themselves in the data and that the findings increase in their transferability. Triangulation was also sought to find multiple points of confirmation of findings presented. For example, in my theme of "Accreditation as a 'Wet Blanket'" in engineering, quotes from four of the five engineering participants were used in support of the theme and to show that it was found across engineering broadly, and not just with one or two participants. Further, most of my final best practice themes were discovered through the confirmation of a majority of the participants. Triangulation is a significant tool in my own practice of effectiveness, where multiple points of corroboration are typically encouraged, and I tried hard not to present a theme that didn't seem grounded in data that was most often presented

by a number of individuals. This ability to triangulate my findings confirmed that I was reaching saturation in my data, with similar themes emerging across the participants.

I also presented a number of negative cases in the data. In fact, the finding of the differences between the approaches of nursing and engineering that I present in my model in Chapter Four show significant negative cases between disciplines. Not only that, they were differences that, going into this project, I had not expected to find and went against my understanding of effectiveness practice. They were differences that required reflexivity on my part to put aside my preconceived notions of how effectiveness happens in practice and to listen carefully to what I was finding in my interviews. Between-discipline differences in evaluative focus, approaches to change, timing of review, number of individuals involved, and paradigmatic approaches to accountability were all types of negative case examples found in the data. There wasn't always such a clean distinction of the negative cases happening only as between-discipline difference, however. For example, mechanical and industrial engineering at Institution 1 showed a more comprehensive process and approach to change than the other engineering programs that I studied, and I point that out in the data. Another example is around the theme of "Organizing the People." While all of the participants spoke to this in a general sense, I present the various ways that this was accomplished in the cases studied to give a broad and deeper understanding of approaches that programs might take, depending on context, to effectively organize people to accomplish their effectiveness tasks. This was another area where I tried not to force my understanding onto the data but to, instead, let the data speak for itself.

There were design limitations. Follow up messages to a couple of the participants gathered no response, so some attempts at member-checking were not able to be carried out.

While I, myself, have experience working in the field that I think helped me to understand with a great deal of clarity what the participants meant in their responses, there still were a couple of areas that I hadn't asked about in earlier interviews that I would have revisited with earlier participants to gain a greater degree of clarity. One of those would have been asking more detailed questions to the engineering participants in Institution 2 about their process for change. Much about their process for change did come up in the interviews, but it was more indirect and in response to other questions not explicitly designed to elicit their approach to change. It would have been nice to confirm with direct questions their approaches. In the end, I still feel like my conclusions around this topic are valid since my participants in engineering at Institution 2 all talked about a more ad hoc process for change, which also aligns with what I saw from engineering participants in Institution 1. It would have been nice to explore this more explicitly, however.

Chapter Summary

In this chapter I outlined my positionality and the methodology I utilized in my study of practitioners in program effectiveness in the fields of nursing and engineering. I employed a grounded theory method to examine multiple case studies to try to find areas of agreement and difference—both in discipline and contextual setting—prior to arriving at my theoretical conclusions. In the next chapter I will present the results and research findings.

CHAPTER FOUR

RESULTS AND RESEARCH FINDINGS

Introduction

The purpose of this grounded theory study was to understand the practices of current practitioners in programmatic accreditation and effectiveness practice. Further, this study also sought to theorize a set of best practices based off of those practices. The research questions that helped to guide the study were:

RQ 1) How do practitioners at two western, doctoral-granting R-1 institutions of higher education understand and experience their programs' formation and implementation of their program effectiveness plans?

RQ 2) How has the context of institution and discipline influenced practitioners' program effectiveness practices?

RQ 3) How do these practitioners' experiences and knowledge contribute to an understanding of generalizable (Ragin & Becker, 1992; Stake, 2006; Yin, 2003) best practice in program effectiveness?

Overarching Thematic Structure

After interviewing practitioners in program effectiveness in nursing and engineering programs, there emerged a pattern of difference in the structure and approaches towards effectiveness between the two disciplines that ultimately presented an overarching culture of evaluation that was consistent within each field, but distinct when comparing across fields. The structure and interrelation of the themes discovered through the interview process is shown in

Figure 4 below. I'll use the structure in Figure 4 to present the themes and the overarching cultures of evaluation that were seen. After my discussion on the themes that emerged in the culture of evaluation model, I will spend some time theorizing best practice in the field of programmatic effectiveness.

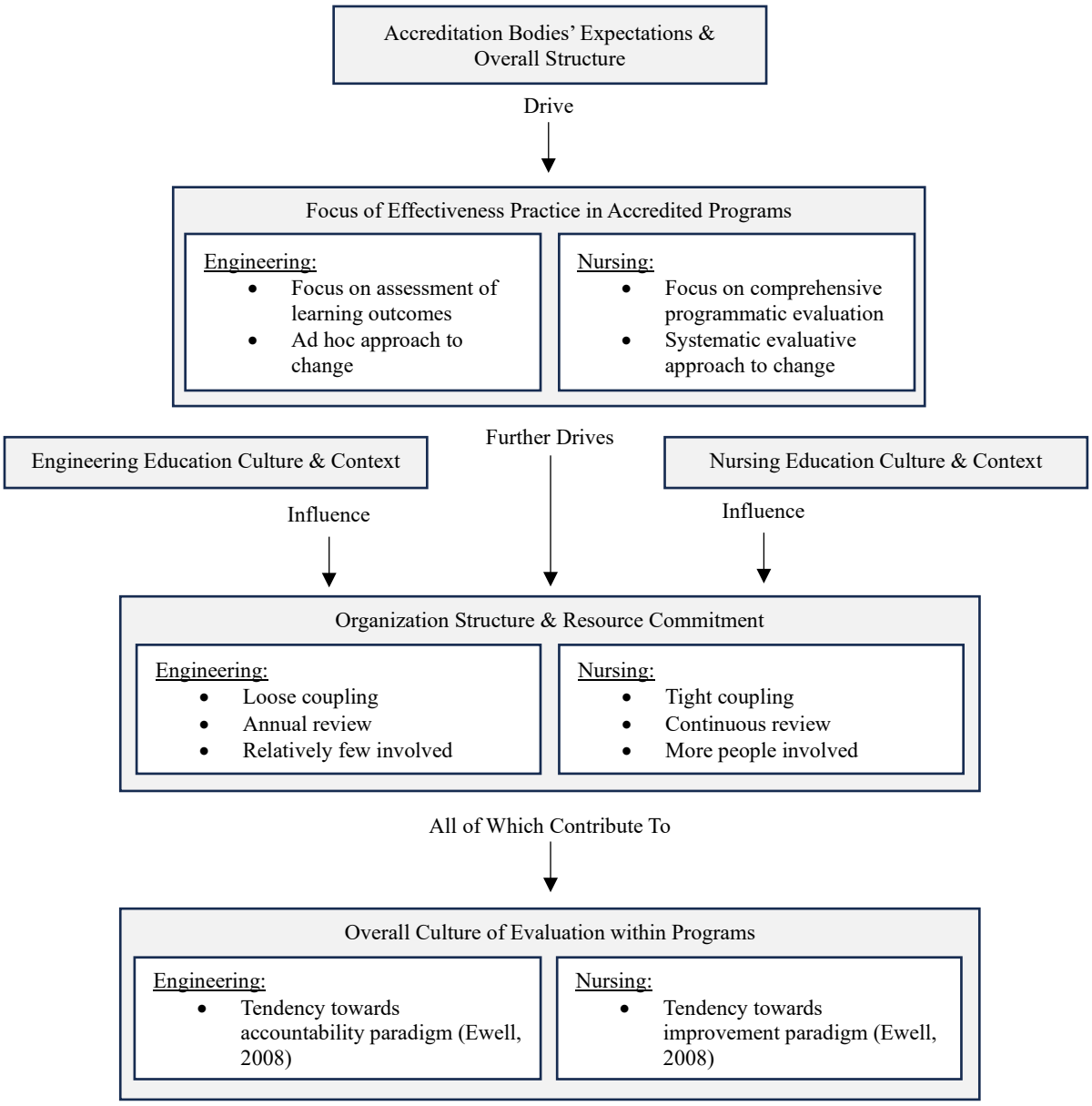

Figure 4. Culture of Evaluation Model in Engineering and Nursing.

Influence of Accrediting Bodies' Expectations and Structure

One of the most striking differences in approaches towards effectiveness between the nursing and engineering practitioners involved in the study was the scope of the focus of their effectiveness efforts. In this section I will explore the themes that emerged around the tendency towards a more narrowly focused effectiveness effort on student learning and outcomes assessment with engineering versus the more comprehensive programmatic evaluation focus in nursing. I'll then look at the structure of the accreditation expectations between the two fields' accrediting bodies—ABET and the Commission on Collegiate Nursing Education (CCNE)—to posit an association between accreditors expectations and the difference in focus between the two disciplines. Themes that emerged in this section included accreditation as having value, limiting constraints in engineering, and creating a comprehensive evaluation in nursing.

Accreditation has Value

In my interviews with practitioners in programmatic accreditation, it became clear that there was not really a feeling that accreditation was a worthless endeavor. All the participants mentioned at some point—though sometimes reluctantly—the value that the effectiveness practices forced upon programs by accreditation brings to the programs themselves. Charlie spoke to this, pointing out how his view towards accreditation has changed over time:

> Even as someone that may not have loved it in the first place, I genuinely recognize that it's useful to have standards and to make sure that you're [thinking about] those standards when you're trying to produce graduates that are going to go on and make things that hopefully make the world and society a better place.

Some of the benefit appears to be just in forcing programs to think about the quality of their programs. Kyle spoke to this when he said of ABET's program effectiveness requirements:

> It's just like anything I think: it's having that deadline that forces you to think about it and do it. It's just like when you're submitting a paper: unless there's a deadline, you just don't get it done because there's other things that are pulling your time away. So, I think it forces us to go through the process.

As Charlie succinctly put it, "ABET provides the stick that makes continuous improvement real."

Accreditation as a "Wet Blanket"

In this section I look at how the engineering participants made frequent reference to ABET as a constraining factor to their programs. Within this theme we see sub-themes of the tension between innovative thinking and what one participant called ABET's "thinking little about things" as well as a theme of change-making occurring outside the framework of ABET processes. Programmatic change in engineering was observed to occur in a more ad hoc manner than in nursing. The constraint was not, however, always seen as a negative. There were some participants who spoke to the constraint of ABET's requirements being helpful in stopping more far-fetched ideas from monopolizing conversation.

Despite the recognition of the value accreditation brings to their programs, the engineering participants made frequent reference to the constraints that the ABET requirements brought to their programs. Mark critiques ABET's effectiveness requirements, saying:

> I think that to me is one of the common problems of this, is that the system they have in place doesn't direct it towards thinking big about things, it directs it towards thinking little about things, "Oh, the writing wasn't as good, maybe we need to somehow come up with some tweak to how we teach the writing," and "they're missing a piece of the design thing." "Okay, well gee, you better go in and teach an ethics module in there to make sure that happens." It's not really designed for something like this: to say, "Okay, let's stand back with a blank piece of paper and really get a big committee together and spend a year debating this to come up with a new approach to doing our design teaching."

The quote was within a larger conversation about how his program had gone through a process of curricular change utilizing more of the "stand back with a blank piece of paper" approach that he felt occurred largely outside of the process for change that ABET would have dictated. He said, as a way of self-preservation, they would find a way to reverse-engineer the decision-making process that had led to the changes to make it better aligned with ABET's expectations when it came time for them to report on the change, though: "We will somehow find a way to attribute this thing to an ABET process, even though that isn't really fundamentally what happened."

Change happening outside of the prescribed process of ABET was a theme that was also reflected by Bill with Electrical and Computer Engineering at Institution 1 when he said:

> The main changes that we made to the curriculum, I cannot say that they've been based on assessment. But always, when we're making changes, we have to keep the ABET requirements in mind. I mean, we're going through the process now. We're starting where we're looking at hopefully making the program a bit more flexible for students so that they can get through as quickly as reasonable. But any change that we make, we have to make it with the understanding that there are certain requirements that we have to abide by for accreditation. But the results of assessment have never driven any significant particular change in my opinion.

This theme of higher-order change operating at a level outside of the prescribed ABET process was consistent in engineering, where ABET seemed more to act like guardrails that would prevent programs from unconstrained change. Charlie, Associate Dean of Engineering at Institution 2, described it this way: "Ninety-nine percent of the conversation going on is much more organic, and ABET, I would say, is prescriptive." The important thing to him was: "How do these two feed each other. How does that organic conversation…ABET is the wet blanket. Let me be a little more precise with that language. ABET is a constraint." He went on to state that the constraint of ABET:

> keeps you from just going crazy with new ideas. So, in a meeting when people start going off and saying, "Wouldn't it be amazing if we did the following?" And you generally have to say, "Well, how would we then assess that?" Or, "How would that be seen by ABET?" Or, "Maybe we need to reach out to ABET and talk to them before we were to pursue this" […] ABET is a constraint that makes sure that when you're having these conversations you're recognizing that if you were to make changes to your programs you'd have to make changes in light of your accreditation requirements. And so, it's a constraint that, honestly, keeps you from chasing the next shiny object blindly.

In this way, Charlie seemed to see the constraint of ABET as positive at times to prevent the conversation from going south when entertaining a decision-making process that is, overall, quite organic.

In contrast, Mark, with Mechanical Engineering at Institution 2, spoke to the process that they engage in when speaking of a curricular change they had made outside of the prescribed ABET process. In doing so, he critiqued the constraints that ABET presented when thinking about such changes, noting that: "There's a certain amount of intuition that goes into teaching that's hard to capture." To Mark, their decision making was more ad hoc, and ABET's constraints didn't allow for that process to be captured due to its narrow focus on assessment of student outcomes. Mark goes on to say:

> When I think of continuous improvement, I tend to think a little bit bigger picture, strategic-type things. To me, that kind of is more entrepreneurial, that's more opportunity-driven. […] To be honest, like I said, before I don't believe that comes out with this kind of evaluation of outcomes.

While the engineering participants consistently shared that most of the changes to their programs happened outside of the ABET process, there was no alternative system described by most participants that would have provided a platform for carrying out "opportunity-driven" decisions, nor a process for evaluating whether such changes were effective. In fact, there was only one program in engineering that described such a systematic process for this type of

decision-making, and it explicitly operated *outside* of the ABET process. Justin, with Mechanical and Industrial Engineering at Institution 1, discusses how they split their continuous improvement process into two paths when he states:

> We have the ABET part. I really don't want our programs to be driven by ABET. I don't think that's the right thing. I mean, they have good ideas. They want us to make sure we have a quality program, and we want to be ABET accredited, but that's not the only thing we want. So, we also have a continuous improvement path where we do other things—like curricular reviews; faculty meetings where we talk about the program and make changes—that aren't really driven by ABET. They're driven more by our pulse on the industry, for example. We're working with our industrial advisory board or, whatever, just more general continuous improvement.

In a follow-up to his response, it was noted that this process was reported in their ABET reports: "It ends up being part of the document we have that we submit because I think it makes it more of a full, better story of what we do. Yeah, and it just helps. It's not really necessary, but it does help." And so Mechanical and Industrial Engineering at Institution 1 was an outlier in the fact that they did actually create a process for continuous improvement that contained elements outside of assessment of student outcomes. They were the outlier, however, and had created the process as one distinct from their ABET requirements.

Innovation Situated within Comprehensive Evaluation in Nursing

The practitioners in the nursing programs did not express the same distinction between innovative thinking, change-making, and their accreditation processes. It was clear that the approach nursing took was one where "thinking big about things" was wrapped within their systematic effectiveness approaches. Jennifer, with nursing at Institution 2, showed this difference in approach when she stated, "We value innovation and we encourage that, but you've still got to be looking at the data. So, I guess the compliance of looking at the data would be the

most important thing. And then the reason you're looking at the data is for quality improvement."

And whereas the engineering participants discussed ABET almost exclusively within the context of the assessment of student learning outcomes, the nursing participants viewed their effectiveness processes tied to their CCNE accreditation much more comprehensively. Kathy, with nursing at Institution 1, spoke to this when she said, "I see assessment of teaching and learning as a subset of program effectiveness. It wasn't the primary effort in the Master Evaluation Plan to begin with. I think we were trying to provide a structure for program's evaluation, first, and then build those [evaluative] processes." These differences in the scope of the approach toward effectiveness between the disciplines are a significant finding, and one that can be explained, I believe, by looking at the differences in the wording and structure of the two fields' accrediting bodies' expectations.

Accrediting Bodies' Expectations Matter

It was surprising to me to see such a stark difference between engineering and nursing with regard to the fundamental scope of the processes they'd developed around meeting their accreditation's requirements for effectiveness—even more so because my experiences with programmatic accreditation in an intensive English program with the Commission on English Language Program Accreditation (CEA) prior to working with CCNE in nursing was so much more in line with what I was hearing from the nursing participants. In fact, the main reason that I interviewed both Bill and Kyle within Institution 1's Electrical and Computer Engineering Department was because I had been pushing Bill about where the rest of their evaluation efforts fit in. It struck me as odd that the person who was responsible for their accreditation was focused

so steadfastly on the assessment of student learning outcomes. When I asked him about how they evaluated faculty outcomes or the adequacy of their facilities, he suggested I speak to the department head who was more involved in writing that part of the accreditation self-study that is produced every six years. However, I found that Kyle's view of the scope of their effectiveness practices was very much in line with Bill's. When I asked Kyle about whether their process of continuous improvement (the term used in engineering for what I would call effectiveness) led to changes in or reviews of something like facilities, he said, "I always hope that it would, but it never seems to," before going on to talk about how engineering is constantly changing and they constantly update their curriculum and equipment to reflect that change. There was obviously something fundamentally different in the way that engineering approached the scope of their expectations for effectiveness and so I did a deeper analysis of the accreditation requirements between the two bodies I was studying.

While both CCNE and ABET place their general plans for evaluating program outcomes in a standard or criterion that is focused primarily on the educational outcomes of students, the majority of CCNE's sub-standards (called key elements) outside of the student outcomes standard mention that the key element must be "reviewed periodically" or "evaluated on a regular basis" when discussing the requirements for meeting the standard. Further, CCNE doesn't only mention that these things must be reviewed or elaborated regularly, the elaborations on these key elements frequently point to the need for, "A defined processes [to be] used" to meet the standard requirements (Commission on Collegiate Nursing Education, 2018). That is, there is a requirement to have a comprehensive effectiveness plan that evaluates, on a regular basis, all the aspects of the program that are contained within the CCNE Standards.

In contrast, while ABET says of their accreditation criteria: “these criteria are intended to foster the systematic pursuit of improvement in the quality of engineering education that satisfies the needs of its constituencies in a dynamic and competitive environment” (ABET, 2024b), their actual criteria are not written in a way that would require programs to develop an effectiveness plan that encompasses areas outside of the evaluation of student outcomes. There is no requirement for a systematic approach to the evaluation of mission, goals, faculty, facilities, institutional support, or student support services, for example. The only explicit reference to effectiveness practices, what ABET calls Continuous Improvement, is in Criterion 4. However, this criterion is directed only towards evaluating student learning outcomes and does not require that the continuous improvement process extend to other areas of the program. This approach contrasts with CCNE, which requires that there exist, in writing, a systematic and ongoing plan for evaluating each aspect of the program that is reviewed during accreditation.

A good example of the difference can be seen by comparing ABET’s Criterion 7 with CCNE’s Standard II-B. ABET’s Criterion reads:

> Classrooms, offices, laboratories, and associated equipment must be adequate to support attainment of the student outcomes and to provide an atmosphere conducive to learning. Modern tools, equipment, computing resources, and laboratories appropriate to the program must be available, accessible, and systematically maintained and upgraded to enable students to attain the student outcomes and to support program needs. Students must be provided appropriate guidance regarding the use of the tools, equipment, computing resources, and laboratories available to the program.
>
> The library services and the computing information infrastructure must be adequate to support the scholarly and professional activities of the students and faculty. (2024b)

Compare this wording with that of CCNE’s around their similar requirement:

> **Physical resources and clinical sites enable the program to fulfill its mission, goals, and expected outcomes. Adequacy of physical resources and clinical sites is reviewed periodically, and resources are modified as needed.**
>
> *Elaboration: Physical space and facilities (e.g., faculty and staff work space, classrooms, meeting areas) are sufficient and configured in ways that enable the program to achieve its mission, goals, and expected outcomes. Equipment and supplies (e.g., computing, laboratory, and teaching-learning materials) are sufficient to achieve the program's mission, goals, and expected outcomes. The program is responsible for ensuring adequate physical resources and clinical sites. Clinical sites are sufficient, appropriate, and available to achieve the program's mission, goals, and expected outcomes.*
>
> *A defined process is used to determine currency, availability, accessibility, and adequacy of resources (e.g., clinical simulation, laboratory, computing, supplies, and clinical sites), and modifications are made as appropriate.* (Underlines mine; Commission on Collegiate Nursing Education, 2018)

We see in these quotes that the criterion for engineering is one that is only required to be addressed during each reaffirmation cycle—every six years for ABET—and there is no requirement to be thinking about facilities in between cycles. In contrast, CCNE expects that programs have a defined process to review facilities that presumably is taking place in between cycles on shorter timelines according to the process defined by the program. This is just one example that could be mirrored in most of the requirements of these two accrediting bodies.

Some of this difference may have to do with the timing of some of the significant structural changes to the accreditation requirements with the accrediting bodies. Engineering was quite early to the program effectiveness train with their Engineering Criteria 2000 edits to their standards in 1997 (ABET, n.d.; Pavlakis & Kelley, 2016). These changes were a direct reaction to the pushback against institutions of higher education and accreditation that was seen at the end of the 20th century. At this time, calls for true assessment of educational outcomes, as opposed to a focus on the educational inputs, led to a number of changes to how institutions of higher education were expected to be held accountable to the public in their education of students

(Eaton, 2012; Ewell, 2008; Hartle, 2012). Accreditation followed suit with a push towards greater focus on outcomes as opposed to a focus on inputs (Ewell, 2009). Since then, however, the overall structure of ABET's standards has changed little, with changes mostly focused on content requirements within separate fields. As Mark at Institution 2 stated, "this current format for ABET is twenty-three years old." He goes on to say with a bit of humor, "It makes me worried a little bit that they're gonna look at it and say it's become stale and we want to redo it now. And I certainly hope that if they decide to do that it will be after I've retired because I don't want to go through a whole different cycle again of this whole thing."

A stated above, there were many references from the engineering participants that implied that ABET did not meet the needs of what some referred to as innovation or strategic thinking. Mark at Institution 2 said, "The ABET things are, basically, they correct things that are in the fundamentals that are bad. The actual answer is I feel like we're doing okay, there, you know? When I see the ABET process, the fundamental process doesn't really aim at that [i.e., strategic thinking], you know?"

While at the time of Engineering Criteria 2000 perhaps ABET's move to looking at outputs seemed groundbreaking, listening to the engineering participants speak of their requirements today they are strikingly inputs-driven. Justin at Institution 1 states, "There are some things that you have to just comply with. You can only have two-thirds of your credits be technical and one-third has to be other, and you have to have these topics." Jennifer, when speaking of her experience as a site-evaluator for CCNE, contrasts this with an approach much more focused on outcomes in nursing:

> The data is important. You can have a kind of a mess of a school that seems to be doing something incredibly out of the norm, and yet their outcomes indicate that

> their students are employed, the students are happy, the students are passing the NCLEX or certification [exams], and that surrounding community of the interest is pleased with the graduates. Then, why not?

By examining the differences in the requirements of these two accrediting bodies, it's not hard to see why the programs in this study have such vastly different approaches to the scope of their effectiveness practices. ABET's criteria are written in a way that focuses continuous improvement on the evaluation of student learning outcomes and that's what the programs in this study did. Bill, with electrical and computer engineering at Institution 1, was as explicit to state: "We do what, essentially, they [ABET] want." CCNE's standards are written in a way that require comprehensive evaluation of the entirety of the program and, similarly, that's what the programs in this study did. And so, the important finding is that the way accreditation standards and criterion are written have a substantial influence on how accredited programs go about envisioning their effectiveness processes. In nursing, both programs studied took an all-encompassing view of the process, where their effectiveness plans attempt to address every aspect of the evaluation of their programs. In engineering, their effectiveness plans were primarily focused on evaluating student learning outcomes.

Organizational Structure and Resource Commitment

The implications of the difference in accrediting bodies' expectations outlined above carry into the way that the organizations are structured and how resources are committed to the effectiveness processes in nursing and engineering. There also appear to be disciplinary tendencies and attitudes at play. In this section I'll explore these broad influences and also look at the sub-themes of how tightly or loosely coupled the structures are in the disciplines, the timing of their effectiveness practices, and the number of people involved in the process.

Disciplinary Influence

Another surprising finding in the research was the difference in the structures of accreditation between the bodies of nursing and engineering. The structure of CCNE's accreditation allows for one college or school of nursing to seek accreditation for all of its programs within one umbrella of accreditation standards. CCNE has one set of standards for graduate and baccalaureate programs in nursing. Sub-disciplinary differences—which in nursing is more common at the graduate level—are all subsumed within one set of standards. And so, if a program prepares some graduate students to be Psychiatric Mental Health Nurse Practitioners, and some to be Family Nurse Practitioners, as both universities in this study did, they are still working under one set of accreditation standards. And while the individual degree-granting programs are separately accredited, the process is such that the accreditation has the same general expectations, and the programs can all be accredited at the same time with one report. That is, one self-study can be produced that incorporates all the programs and the site visit can encompass a review of all of the accredited programs.

ABET is structured differently. While they have General Criteria for Baccalaureate Level Programs that everyone must meet, they then have separate Program Criteria for the sub-disciplines of engineering. To date, there are 29 sub-disciplines with distinct Program Criterion listed on ABET's webpage (2024b). These 29 sub-disciplines have additional criteria for curriculum and often for faculty that are overseen by the various member societies of ABET, which number 35. And while this may not be the case with all engineering programs, for the ones in this study each department wrote their own individual accreditation reports that were inclusive of both the general and program criteria. That is, each department created their own unique report separate from each other engineering department.

This has significant ramifications for the level at which the responsibility for accreditation lies within the programs studied. For the nursing programs, the accreditation responsibility lay at the college or school level. For the engineering programs, the responsibility for accreditation lay at the departmental level. Weick (1976) first proposed the idea of the difference between loose coupling and tight coupling in organizational systems, with tightly coupled systems showing greater levels of interdependence amongst their constituent parts than loosely coupled systems. Following Weick, the nursing programs in the study showed a tightly coupled organizational structure with a large interdependence amongst the parts of the organization. This tightly coupled structure allowed for centralized resource deployment in accreditation at the college or school level. Conversely, the engineering programs studied showed a loosely coupled structure with a low level of interdependence amongst the various departments within the college or school of engineering. The loosely coupled organization led to de-centralized resource deployment for the accreditation process; that is, the resources were departmental-level, and not pooled at the college or school level like I saw in the nursing programs. There are real ramifications for efficiency in the difference between these two models as will be discussed below.

Number of People Involved in the Process

Looking at both the differences in the scope of the accreditation expectations and the disciplinary influences already discussed, it's no surprise that the two disciplines showed differences in the number of individuals involved in the effectiveness process. The nursing programs in the study showed a consistently higher number of individuals involved in the process and both programs had non-nurses in positions that play a significant role in the

effectiveness process. Jennifer spoke to how the team involved in the effectiveness processes in her role at Institution 2 was expansive, with student services, advising, recruitment, admissions, curriculum, and scheduling staff involved in their Quality Improvement Team. Because of the expansiveness of the groups involved, she says, "we didn't embed it in the faculty organization structure because it wasn't just faculty. So, 'Mary', the curriculum person, 'Cindy', the student affairs person, a budget person, are all part of that Quality Improvement Team, which meets regularly."

At Institution 1, the nursing college's governance system is one that is set up to support the work of accreditation that lies primarily within the purview of the college's Academic Programs Team. Being embedded in this way helps to steer the committees and keep them on track with regard to effectiveness practices:

> So, the College of Nursing has a very structured faculty governance system, made up of a Faculty President, a number of committees that are staffed by faculty from across the College, and across the state, two of which are our curriculum committees—academic curriculum committees at the undergraduate level and the graduate level. Academic Programs sits as ex officio on those committees and helps to keep the faculty moving forward on a timely schedule to really review the curriculum itself on a regular basis […] and keep the work of the of the College and the curriculum going. Another committee […] is an Executive Committee that is advisory to the Dean [and included membership of Academic Programs] and allows kind of those higher-level conversations to evolve for the college itself—kind of drive initiatives that need some work and, of course, accreditation is kind of a piece of that ongoing work of the College.

While there did appear to be some attempts within engineering to have some centralized support for the processes, the overall feeling from those involved was that there were very few human resources committed to doing the work. Charlie discussed the Accreditation and Continuous Improvement Committee, that has both departmental and college-level branches:

> The ACI Committee is a standing body at the departmental, and at the college level, where they're responsible for really doing the work between review cycles, which

> is making sure that you're actually paying attention to…are you delivering on what you were accredited to do and are you reviewing to see what's working and not working.

Some of the engineering programs split those functions in two. As Charlie said, "Some departments have two individuals: one doing accreditation and one doing curriculum."

The Mechanical Engineering department at Institution 2 is one that tends to have two people, but it doesn't always work out that cleanly, as Mark says when he was discussing that he currently is serving in both roles: "Typically we have a separate person who acts as the accreditation coordinator. So that would be outside of my job. I support that, but there has been a couple of times I've worn both hats at the same time."

Mark went on discussing whether the effectiveness processes were defined top-down from the college-level, or more at the departmental level. From his perspective:

> That's pretty much your little department. It's kind of a handful of people that are interested in it that kind of serve on an ABET steering committee and talk this stuff over. Just a couple people that sort of sign up for interest. The kind of stuff you don't put somebody on that doesn't express an interest in it because you'll got nothing other than some nominal service for that.

This was representative of a theme that emerged in engineering—finding people to be involved in the effectiveness process seemed to be a consistent challenge. Justin with Mechanical and Industrial Engineering at Institution 1 stated, "it's just it's another thing we have to do, that all the faculty have to do. They already have too much to do and so it's hard to keep that as a priority." With Electrical and Computer Engineering at Institution 1, they pay faculty for work in summer to compensate for this. Kyle states, "What we do in this department is we give one month of summer support to our ABET accreditor. […] So we purposely take some of those program fees that we're getting from the students and dedicate it to having one person that's

going to focus on doing our process for our accreditation. So, we dedicate some faculty resources."

Bill, who receives that summer support at Institution 1, however, still echoed Mark's "handful of people" statement with his comment regarding who is involved in the effectiveness plan in his department: "I think it rests in few hands."

Another factor that no doubt influences the number of people involved is the differences in the number of faculty to call upon. Some of this is a result of the tight versus loose coupling already discussed, where nursing is able to pull from the entire nursing faculty, while engineering is pulling from within the smaller departmental units that create their individual accreditation reports. Faculty composition and institutional type probably also play a role. It may also be an area where discipline norms come into play. Both nursing programs have substantial faculty numbers, with a high percentage of adjunct faculty, whereas the engineering departments involved have much smaller faculty numbers, with the majority being tenure track. They both exist within R1 institutions and there are obviously differing incentives for faculty depending on track and rank. Charlie stated:

> The one nuanced piece then is that you're often really focused on finding faculty that have the bandwidth to do this work. And in a system like Institution 2, where you have different tracks of faculty—we have research faculty who are really shut off from all of this. They're just on soft money and their grants, doing research. You have tenure track faculty that have to have a balance of service, teaching, and research. And you have your teaching faculty that have a balance between teaching and service. In some limited cases, you may find more of this effort, then, pushed onto the teaching faculty, which is an extra burden. [...] There are equity issues around something as time and effort intensive as an accreditation like ABET, and that is real.

The large number of adjunct faculty in the nursing programs interviewed no doubt lend themselves a little better to participation in the process as they don't have the same pulls on their

time that tenure track faculty have. However, the nursing programs were not bereft of tenure track faculty, and they also took part in the process. It's an area where it seems like there is most likely some type of influence, but it might be that the rest of the culture around effectiveness practices in nursing is such that it is more likely to bring tenure track faculty into those practices, despite the conflicting pulls on their time.

Timing of the Effectiveness Processes

Perhaps it's not a surprise, reflecting on the differences in resource allocation between the two disciplines, that there is also a difference in the timing of their effectiveness processes. The engineering participants in the study described their effectiveness practices as occurring less frequently than the nursing participants. Bill, with Electrical and Computer Engineering at Institution 1, stated that for his program:

> It's more of an annual basis and, to be honest, we wouldn't even have to do it annually based on how we've argued things out in our reports, but naturally we do. Occasionally it will come up other times in the year. But really you can count on every faculty retreat in August I have an hour or so where I discuss the current status of things. We [also] have our industrial advisory board come in every year and I present results to them and ask for their opinions because again, they're one of the stakeholders that we need to consult with.

When describing how they keep track of the timing of their effectiveness practices at Institution 2, Mark states:

> It's kind of a master timetable for when things happen over the six years. So they're sitting there saying, okay, here's each of the six years going in here, so what's supposed to happen in each year. So it's kind of the years and across the top would be the evaluation things and there'll be an X on what's going to happen that year, and then the head ABET person's in charge of making sure that gets addressed.

The nursing programs plans for effectiveness were, essentially, continuous. Both programs in the study had a named plan that they utilized: for Institution 1, it was called a Master

Evaluation Plan and for Institution 2 it was labelled the Systematic Evaluation Plan. While they'll be covered more in-depth in the section on best practices below, it was clear that the plans that they had developed for their effectiveness practice were things that they looked at continuously. The overall effectiveness plans in nursing were structured in ways that dictated that they were being constantly updated and worked on.

I've shared how I believe the structure of the accrediting bodies in nursing and engineering contribute to differences in not only the scope of the evaluative efforts, but also the human scope of who is involved and how much time is committed to the process. Nursing's accrediting body dictates a process of comprehensive, ongoing evaluation that the programs in the study embraced. Engineering's accrediting body focused on the assessment of student learning outcomes, leaving most of the programs in the study to approach higher-order change in a more ad hoc way. These changes all have ramifications for the overall cultures of evaluation that were seen in nursing and engineering. I will discuss those ramifications in the next section.

Overall Culture of Evaluation

When summing the preceding differences that exist in effectiveness practices between engineering and nursing, the result is a significant difference in the cultures of evaluation that existed in the programs studied: nursing tended more towards what Ewell labelled the improvement paradigm, where engineering tended more toward what Ewell called the accountability paradigm (2008). That is to say that the effectiveness practices within the nursing programs in the study were undertaken with the spirit of improving the programs, whereas the engineering programs, despite not dismissing the process as having no value, seemed to be more

involved in the process to align with accountability requirements. This can be seen clearly when looking at the way the two different fields' participants discussed buy-in with the overall process.

Buy-In

One area in which the literature on accreditation had been conflicting was around buy-in to the accreditation process. Most of the research around accreditation is focused on institutional accreditation, however, which encompasses a much larger group of constituents than does programmatic accreditation. This research suggests that garnering leadership, administration, faculty, and staff buy-in can be a significant challenge (Bush, 2016; Procopio, 2010). However, there is also some research that suggests that buy-in may be less of an issue in programs with programmatic accreditation (Ewell, 2009; Welsch & Metcalf, 2003b). The latter was certainly found to be supported in the interviews conducted where there were frequent references to support at all levels and the requisite buy-in from the various parties involved. Mark with engineering at Institution 2 states, "we don't really get pushed back on it. You know? Nobody says, I think this is a stupid idea, I'm not gonna deal with being involved in it. That doesn't happen."

But this isn't to say that accreditation and program effectiveness were at the forefront for all parties. Charlie talked a bit about the timing of the focus, as well as the difference of engagement, when he said, "I would say [for] most faculty this is something that is at back of their mind and then for a significant chunk of people involved in the renewal process, it becomes front of the mind for about a year and a half." However, there were shades of difference in the buy-in for many of the programs—most strikingly with the engineering participants where the

feeling of the buy-in being more for compliance with accreditation shone through. Charlie goes on:

> No one likes having to do anything that they're being made to do. So, speaking as an instructor, there's nothing about this that's fun. But, in the same way that most of us don't actually enjoy exercising, but we like the results. I would put it in that category of you do it begrudgingly. In part because you have no choice, but in part because you actually benefit from having done it.

Mark, also at Institution 2, discusses the importance of having the accreditation for their students' employment: "the federal government will not hire an engineer out of a non- accredited program. And there's a handful of employers—like Boeing—that won't hire out of a non-accredited program. So, it's something you don't want to mess with." What we don't see in the interviews in engineering, is any real excitement about the ways in which their effectiveness processes improve their programs and the buy-in that comes from everyone seeing that improvement.

In nursing, while there were still some mentions of challenges with buy-in, there seemed to be a much greater acceptance of the process as being important because it adds value to the program. Kathy, who has been in her current role with nursing at Institution 1 for five years now, stated, "There's always some pushback from people who don't really see the value, but I feel like that's less and less so. I feel like our faculty like having data and that we're becoming more reliant upon having data with which to make decisions, which is a good thing." Kathy also showed a different approach than that of the engineering programs with their "wet blanket" attitude towards ABET as a grounding constraint: "I try not to use the CCNE card too much because I really do want them to think about how we can improve things." Casey, with nursing at Institution 2, stated of their ongoing evaluative process and whether it was driven by their accreditation requirements: "We would want to be doing this anyway in the interest of program

quality." So, where we saw that engineering participants often referred to the culture of their accreditation process as being one that was used to stifle changes people on their faculty might want to make because it would put them out of compliance (accreditation as a "wet blanket"), the nursing participants focused on putting the notion of putting improvement at the forefront: because they viewed improvement as the main goal of their effectiveness work.

Summary of the Model

In the preceding sections, I've taken an overarching approach to exploring the themes from my interviews with practitioners in program effectiveness within engineering and nursing as presented in Figure 4. Those themes show a significant difference in overall approach between the two disciplines starting from their accreditors' requirements through to their overall cultures of evaluation. In the next section, I'll shift to theorizing best practices within program effectiveness for units that are programmatically accredited. As stated in Chapter Three, I used institutional theory to propose that the field of accreditation presented enough isomorphic tendencies that there could be more generalizable conclusions drawn around best practice in effectiveness practice. This is an area where the literature is relatively silent; however, since the practitioners of effectiveness practice in this study had all originally come to the field as novices, having some guidelines for newcomers would be helpful in practice for new practitioners.

Best Practices

After analyzing the interviews, themes for best practices in programmatic accreditation and program effectiveness did emerge. The major themes that were discovered when interviewing these practitioners were the support of organizational structures and resources,

effectiveness plan structure and documentation, feasibility, development of a culture of evaluation, consideration of succession, and metaevaluation. I'll explore these broader practices, as well as some of the subcomponents of those practices.

Support of Organizational Structures and Resources

While in the themes explored above, we also discussed organization structure and resource commitment from a more macro-level, in this section we'll explore the importance of structuring organizational structures and resources in ways that allow programs to meet their evaluative needs, including efficient collection, analysis, and distribution of key findings, while also allowing for decision-making to come from those findings. Some of the subcategories that emerged when looking at organizational structure and support were those of having an accreditation champion, ensuring the organization of people is conducive to the process, and organizing the work to be done.

Have a Champion. Virtually all of the participants discussed the importance of having one, or a few, individuals who were going to be the ones to take on the mantel of championing their effectiveness practices. Most of the participants in this project were, themselves, those champions. Part of the importance of having the champion is to have a point person, but it also became apparent that having the champion also meant that there was someone who would be taking the time to make sure the work got done. Justin with Mechanical and Industrial Engineering at Institution 1 spoke to both the time required and the importance of having someone to lead the effort:

> It's easy to get the data. It's not so easy to summarize the data, and then it's even harder to talk about it and figure out what to do about it and then implement that and make sure that happens and keep track of it. There's a lot of administrative things that seem pretty simple that aren't because it just takes a lot of time and effort

> and energy to make it happen. And so, the challenge is having a champion for your continuous improvement program. And you got to have one and it's hard to find one.

Assigning the role of "champion" for the effectiveness process was a best practice espoused by all of the participants. Casey with nursing at Institution 2 noted how officially assigning it as a job role, as opposed to just adding to a faculty member's workload, was important: "We've got a faculty member who's specifically charged with this kind of work instead of trying to find a tenure-line faculty member who can do it as part of their service commitment, or something like that. You just don't get as consistent of follow-through as you do when somebody who, now it's officially part of your job." As Jennifer, also with nursing at Institution 2, states, it's about "prioritizing these efforts as [being] as important as teaching." She goes on, saying that, "There has to be somebody at the helm and somebody pulling it all together. Somebody has to be the champion: speak the language and do the campaigning for how important this is, to get people involved and to coordinate all the efforts."

Organizing the People. Important to the practice of effectiveness is ensuring that the organizational structures of people within the unit is such that they are conducive to carrying out the work. Kathy, with nursing at Institution 1, talked about how the last reaffirmation cycle of their CCNE accreditation led to a restructuring of existing positions, as well as the addition of a new position:

> Reflecting upon the site visit in 2018 and the new role that Dean Shannon set up with a combined Associate Dean for Academics, both at the undergraduate and graduate level, and fully overseeing program accreditation, assessment of teaching and learning, [and] the curricular issues for the programs, [we added the positions of] the Director of Programs and Evaluation and myself.

The change in structure has allowed for the nursing program to have the true champions of the effectiveness process that were discussed above. With the structural change came the people and time to do the work.

Jennifer, with nursing at Institution 2, talked about how lack of this organizational support structure led to a "rescue mission" when she was working at a previous institution. When she was hired, "The report was due in eight months, or something ridiculous, and I had no idea about the program. So again, none—or very little—staff support and neophyte faculty had been assigned, people who never thought about or looked at accreditation processes. So that was a rescue mission."

We already saw how in Jennifer's current role, her Quality Improvement Team is organized in an expansive way that incorporates individuals from student services, advising, recruitment, admissions, curriculum, and scheduling staff, and not just faculty. However, it is not that faculty groups are not involved in the process. They can even lead to structural change with their participation in that process. Casey with nursing at Institution 2 discussed how the faculty committee groups involved in the curricula were instrumental in instigating an increase in support with the addition of a Program Director position for each program. This had the added benefit of leading to more unity in the school:

> The faculty on those committees were advocating for something like a Program Director to be put in place because the program director now ties all of the committees together in an overall school effort instead of having it be a little bit siloed of—this group looks at undergrad stuff, and this group looks at the DNP stuff—but they don't really talk to each other very much.

These curricular committees are also centrally involved in the data-driven decision making that happens in their effectiveness planning. Casey goes on:

> We have coordinating committees for each of our degree programs: that's a group of faculty for each of our departments, a couple of admin staff, and students have representation on there as well. But that's in direct response to our internal faculty governance structure where our faculty have ultimate deciding voice on matters of curriculum and program quality. So those committees, when I talk about doing the assessment, those committees are often the first to see the information and the first to take a pass at what needs to be done about it.

So, best practice would dictate that programs have a champion for program effectiveness. Further, the people within the units—inclusive of, but not exclusive to, the champions—need to be organized in a thoughtful way to address the needs of the effectiveness work. Next, we'll look at the importance of organizing the work itself as critical in establishing effective practices.

Organizing the Work. As already seen, there can be a large amount of work involved in program effectiveness. Units that are successful in their practice all showed intention around how they organized and scheduled the work that needed to be done. While we already saw earlier in this chapter that engineering's support structures for effectiveness are less comprehensive than nursing's, there is still recognition that during a reaffirmation cycle the units need additional resources to get the work done:

> During the height of preparation for an upcoming renewal, you will typically have a professional staff member, or even a group of staff, that are staffing this work because the department is turning in a phone book sized report that's being submitted for each unit right? […] And that just is a tremendous volume of material that needs to be collected. And so there really is a lot of time, effort, and energy needed.

The need to ramp up support by redistributing resources within the organization is one way to deal with additional workload around accreditation reporting periods. In contrast to that approach is the one taken by nursing at Institution 1, which has structured its plan and documentation process in a way that allows for greater ease of continuous reporting, thus lessening the need for such resource distribution:

> The plan itself, it helps us to be successful in being on top of the CCNE standards and being able to say [that] we should be able to do our CIPR report here this next year—our Continuous Improvement Progress Report for CCNE—we should really be able to pull from the Master Evaluation Plan in articulating how we're meeting these individual standards pretty easily.

The actual work of program effectiveness can be distributed into more bite-sized, ongoing pieces, like it was with nursing, or in more larger pieces focused around accreditation reporting cycles. Either way, there needs to be a recognition that there is a significant amount of work to be done and resource allocation needs to rise to the level of that need.

Effectiveness Plan Structure and Documentation

A well-designed effectiveness approach needs to include a purposeful approach to structuring effectiveness practices and documentation of the work done. As already stated, both nursing programs showed detailed attention to the structure of the plan and the documentation process accompanying it. Not only that, but the plans themselves in nursing were more encompassing than engineering. And while ABET doesn't require programs to have such a comprehensive effectiveness plan, the experience of the nursing programs suggest that having a comprehensive plan is probably best practice.

Both nursing programs had an actual name for their effectiveness plan. At Institution 2, they called it the Systematic Evaluation Plan. At Institution 1 it was dubbed the Master Evaluation Plan. This approach was not echoed in the engineering participants, who mostly discussed ABET's criterion of Continuous Improvement and assessment of student outcomes when discussing their effectiveness practices. So, when discussing plan structure, for the engineering programs there often wasn't a set structure in the same way that there was with nursing, though they certainly had an established process. The process itself for the engineering

participants seemed to mostly be focused around some way of keeping track of when certain aspects of their evaluation were supposed to take place and then a process for documenting their results.

Justin's department's structure of documentation for effectiveness is a file-structure system with assessment rubrics, examples of student work, and aggregate results of student assessment. They keep track of what needs to be done through the use of a calendar. For structure at Institution 2, Mark talks about having an electronic post-it note type system for site evaluators to be able to click on by content categories: "if you're interested in ethics here's three places in the curriculum where ethics was. You know, something the students did that involved an ethical exercise in here, or something like that."

The nursing programs have a significantly more nuanced approach to the organization and structure of their effectiveness plans. At Institution 1, the structure of the plan is a Microsoft Excel document with a tab for each standard. Under each standard are listed the sub-standards (which CCNE labels Key Elements) which then have columns for timelines, parties involved, areas where evidence are kept, the outcomes seen, and the response to those outcomes. Kathy discusses it this way,

> So, the platform itself allows us to be able to track each standard individually with a plan in place of: how often; who's responsible for it; how often are we going to look at that particular problem—is it something that needs to be done annually or not so frequently; if it's being met; what the data source is; what the outcome is; and what our strategy is for improvement if that needs to be [improved].

At Institution 2, their plan was also overseen by a state regulatory body in an interesting intersection of another spoke of the "triad" of oversight in higher education. Jennifer spoke of the body's influence over their plan, leading it to become what she thought was an "unhelpful" 40

pages. However, we see a very similar, overarching and systematic approach to their evaluation of the program within a structured plan:

> Having that Quality Improvement Team kind of monitoring all of that and using the Systematic Evaluation Plan as a guide to make sure we were kind of marching through it. Forty pages of things like "oh, look, we need to be looking at our mission values and outcomes every 5 years or every 3 years, we better sit down and do that." That's great, but then you also have to document it. What meeting did that happen in? What other committees within that faculty structure have to know that?

Jennifer also talks about the importance of the documentation of the process,

> The Quality Improvement Team minutes are a great place for any on-site evaluators to look and say, oh, look, they did something. They used this data and made a quality improvement. So, it's all packaged right there for them. So that's the biggest thing that I did is kind of formalize that quality improvement effort that is just crucial to accreditation. [You have to] have that documented and have faculty involved in it.

The results around the area of the overall structure and comprehensiveness of the plans for program effectiveness again show that accrediting bodies' standards have a real influence on how their accredited programs go about evaluating themselves. The suggestion for best practice at the program level, regardless of accreditation requirements, would be to take the time to form and implement a comprehensive effectiveness plan. The nursing programs in the study used their accrediting body's standards as the general structure for organizing their plans, which might serve as a template for other programmatically accredited units looking to develop a true comprehensive evaluative structure for their programs.

Feasibility

Whether comprehensive or not, all participants spoke to the need for making their effectiveness plans feasible in their implementation. The pulls on faculty time have already been discussed. Developing a plan that hits all the requirements exceptionally but that does so in a

way that overburdens those involved in the process is counterproductive. Mark with Mechanical Engineering at Institution 2 stated, “People recognize that accreditation is important as long as you make the stuff that you give them reasonably user friendly. So, things like, okay, I’m gonna ask you to fill out a report on your course. They’ll get a form with specific questions on it, and, you know and [they] address those.” While from the discussion directly preceding this one might think that the comprehensiveness of nursing’s approach would produce a heavier burden on their faculty, nursing also discussed the importance of making their plans feasible. Kathy at Institution 1 discusses how one approach to getting this done would be to break up the things that need to be evaluated into bite-sized chunks: “I think it could be pretty overwhelming to bring all the data at once and have them somehow processes it appropriately. I think what we’ve tried to do is break out some tasks […] in certain parts of the academic years so that they’re not completely overloaded.”

Breaking up evaluative requests throughout the academic year for Institution 1’s nursing program is important not only for keeping the volume manageable, but also to ensure that the decisions that are made align with university timelines: “we try to time those with other university processes so that we’re efficient in getting things done and not having to wade through this slog of academia.” An example of this purposeful timing would be to review items that result in curricular changes in the Fall semester so that changes can be made in time for the following academic year’s catalog. So, part of the feasibility requires thinking through the best time to get these things done to fit within the larger structure of the university.

For the engineering programs, this timing aligned with their annual faculty retreats in the Fall where they could discuss curricular decisions as a group. Another approach mentioned in

getting to a feasible process was the need to make reductive decisions about what to look at. Justin with Mechanical and Industrial Engineering at Institution 1, said, “We have a pretty good assessment process now and we’ve minimized it to the point where—we used to take all kinds of data in all kinds of classes—now we’ve got it narrowed down to where we really want it. Then we just summarize that data and discuss it.” This has come from developing a better understanding of the process and what’s important for them to reach their evaluative goals. At another point in the interview, Justin discussed the importance of understanding the difference between assessing and evaluating and the process of discovery in finding their current solution: “We found out we were collecting a ton of data, but we weren’t really doing anything with it. We were really assessing but not evaluating, and we were assessing a lot of things that were just more subjective, rather than real data. We didn’t have real data.”

Best laid plans are useless if they are overwhelming for the people who need to carry them out. By making sure that the effectiveness processes are tailor-fit to the unit to be feasible for those who have to carry them out, programs are able to derive the greatest benefit from their practices. Even in nursing, with its more comprehensive approaches to evaluation, there was a recognition of the need to keep it simple and manageable.

Develop a Culture of Evaluation

All of the best practices leading up to this point are important because they help to develop a culture of evaluation within the programs. Carefully considering the organization of people and distribution of the work, structuring a plan that is effective and feasible for those involved, all contribute to the ways that the parties involved view the process. But it is also

important that everyone within the organization is involved and knows of the importance of the process. Kathy at Institution 1 talks about the importance of this when she states:

> We certainly try to keep it in the forefront. And I think the way it's designed does help us with that. We can constantly be saying, here's some survey data, here's the latest NCLEX data, here's some information on retention and enrollment from the university. So, we do kind of try with all the committees to give them some information throughout the academic year, so that they can keep this in their forefront. I think that helps a lot.

The notion of putting the information in front of the people who have a stake in the process influencing the overall culture of assessment of the unit is echoed in Bill's statement, "I've given them [faculty] some models of what a rubric could be like, and they've run with it and developed their own. And to be honest, some people have been very enthusiastic about it and tried new things." That enthusiasm gets at an attitude of approaching effectiveness practices from a perspective that is truly focused on improvement, as opposed to one that is focused on accountability (Ewell, 2008). That is, it gets at the development of a culture of evaluation within the unit.

Think About Succession

Perhaps cultivating an organizational culture of evaluation would help to ease the challenges of transitions that many of the participants discussed. Engineering at Institution 2 is just starting to ramp up for a reaffirmation cycle with ABET. Charlie said of this cycle that, "I can probably say most of the chairs for this cycle were not chair the last time ABET accreditation happened." That presents an awful lot of learning for someone if they were not aware of the evaluative processes that go into their accreditation. So, a culture of evaluation may be one way to help with that continuity. But there are challenges with handing off the work. Mark, who is closer to the end of his career than the beginning, succinctly put his desire to transition away

from his role with accreditation in his unit this way: “I want to give it up to a new guy who’s going to be around here longer than I am.” But, as Justin stated, there’s “not so many volunteers to be in charge of assessment around here, I can tell you that.”

When I spoke with her, Jennifer had just recently stepped back into a faculty role at 0.5 FTE as she transitions into retirement. She discussed how those succeeding her have been steeped in the evaluative culture, which should help in the organization’s transition with her stepping back:

> I’ve been available to the person who stepped into my role. She was put in as an interim. And she had previously done the Undergraduate Program Director, and then the Graduate Program Director. So, within my six-year span, then, she started out undergraduate and moved to graduate. So, she had that full span. Yeah, she was wonderful. Our partnership was great. Still is. And I’m still available to her. But—she is top notch—just a young faculty coming into it, so she has a lot to learn, but she is doing a stellar job. Before I stepped out of the role, I asked two other, again, stellar individuals to assume the Graduate Program Coordinator and the Undergraduate Program Coordinator positions. They’re terrific.

She also, however, talks about real challenges to continuity and transition that are familiar to those within academia. So, while the “core is solid”:

> In the meantime, the Dean is leaving. The associate Dean for DEI has accepted the Dean position at [another institution]. So, he stepped down. Another department chair, who was the department chair for 22 years, has retired and is moving into a faculty role, and the other department chair for the Child, Family, and Community Health nursing has taken a job as a nurse researcher at [a clinical agency]. So all of that top leadership: gone.

No unit seemed to have a silver-bullet solution to the problem of succession. And, as can be seen from their descriptions of change in their units, it does present a challenge to keeping up effectiveness practice through the vicissitudes that time brings in academia. My sense is that Jennifer’s description above comes the closest: by cultivating a culture of evaluation, and embedding evaluative processes deeply within the unit, academic programs are given the best

chance at continuing to effectively evaluate their programs despite the inevitable changes that come in their units.

Metaevaluation

Another area of best practice that was supported by the research (Brescianni, 2009) was that of incorporating metaevaluation—an evaluation of your evaluation—into your effectiveness practice. So, not only should programs have a defined process for evaluating their programs, but they should take the time to step back and think about their evaluative processes themselves—are we evaluating the things that we want to be evaluating? Are our evaluations actually giving us the data that we think they are? Is our evaluative process fair and trustworthy? These are just some of the types of questions that might be asked when thinking about a metaevaluative framework. When asked, the engineering participants did not say that they included metaevaluation in their process. And so, this was another area where the disciplines of nursing and engineering diverged. This may be, again, a result of the differences in the scope of the effectiveness practices required by the two accrediting bodies. However, there is enough support in the scant literature that exists for the process that combined with what I saw in nursing to reaffirm metaevaluation as an important best practice.

Jennifer, who had gone through a complete revamping of their Systematic Evaluation Plan at Institution 2, hadn't had the time to get the metaevaluation stage, but acknowledged its presence in their plan: "That's in the Systematic Evaluation Plan. We didn't get that implementation […] in early enough, but we're due." Kathy, with nursing at Institution 1, spoke directly to the presence of metaevaluation before I was able to even ask the question, "We've kind of built into it an evaluation process. So, how well are we evaluating the process of

evaluating of our work?" She goes on later in greater length to discuss how their teaching and learning assessment process that evaluates students' achievement of program objectives, also utilizes a metaevaluative framework:

> It gives some ability to judge not only these scores that you get when you assess how well students might learn a particular program learning outcome; it allows us to kind of assess and place some quality to it as well. […] We tied in these concepts of, is it accurate? Is it useful? Is it measuring what it's supposed to measure? And is it fair? You know? For students, is it fair? Is it a fair system? I'm hoping that it kind of helps us sidestep faculty feeling like it's an attack on them. I don't want them to feel like that. I want them to recognize that we're trying to look at this really comprehensively. And, yeah, maybe students are scoring well on this particular exemplar, but does it really have value? Is it meaningful? Does it really tell us that they're going to be successful nurses? And if not, then, why are we using it?

In short, no system is perfect as designed. Programs might employ one evaluative strategy that they think will meet their needs only to find that it's not a good fit. Without a process built in to force a reflection on established practices, it might be easy for ineffective practices to continue. Incorporating metaevaluation into the effectiveness plan helps programs to deliberately think about their processes and to make changes to those areas that are not working well.

Chapter Summary

This chapter presented the findings from an analysis of the practices of practitioners in programmatic evaluation and accreditation in the fields of nursing and engineering at two large, R-1 institutions of higher education, focusing on best practices that might be used as a guide for others within the field of evaluation in higher education. One important finding was that the language and structure of accrediting bodies has a significant influence on the scope of the effectiveness practices that units engage in.

Regardless of scope, however, other best practices were discovered. The importance of the support of organizational structures and resources was important in realizing all participants' effectiveness practice. Those support structures and resources were foundational in establishing effectiveness plan structures and documentation practices that allowed the programs to stay on top of their work and document what they'd done. Another aspect all participants spoke to was the importance of making the plans feasible and not overburdensome. Having a champion of the process who is able to take the time to make the process meaningful helps to engage other parties within the unit and to develop a culture of evaluation to help when there are faculty and staff changes to those involved in the process. Finally, having a structure to evaluate the effectiveness of your evaluation is recommended to ensure the quality of the process. In the next chapter I will discuss the conclusions and implications of the findings.

CHAPTER FIVE

CONCLUSIONS AND FINDINGS

Introduction

The purpose of this qualitative grounded theory study was to understand the experiences of practitioners in program effectiveness in accredited programs of nursing and engineering at two large, very high research, institutions. Program effectiveness can be defined as a comprehensive, ongoing practice of systematic evaluation of all aspects of an academic program and is often incorporated into accreditation requirements (Ewell, 2001; Head & Johnson, 2011; Larkan-Skinnner, 2005; Welsh & Metcalf, 2003a; Welsh & Metcalf, 2003b). Practitioners from each of the four bounded cases in engineering and nursing at two institutions were interviewed about their experiences carrying out their program effectiveness practices within their contexts. These practitioners' experiences were then analyzed and used as a catalyst for theorizing best practices in program effectiveness in higher education units that are programmatically accredited. The study sought to address the lack of empirical research around program effectiveness practices in higher education and to give practitioners real-world advice on how to better carry out their requirements for holistic program evaluation.

The study utilized institutional theory and the idea that accreditation, as a field, has become isomorphic, with structures and processes that propagate across disciplinary and accrediting boundaries. The multiple bounded cases utilized different fields, as well as different institutional contexts, with the expectation that the differing accrediting bodies were similar enough that findings could be drawn for programmatic effectiveness generally, and that these

findings did not have to be bounded to any individual field's practice. Exploring the different academic disciplines within two distinct institutional contexts also allowed for areas of similarity and difference depending on that context to emerge. And so, one aspect of the study was also to ascertain to what degree are programmatic accrediting bodies similarly structured and to examine what we can learn from best practices in the areas of discrepancy. Further, what influence does institutional context have on program effectiveness practices and are there any best practices to be ascertained from those similarities or differences.

My interest in the study comes from twelve years of personal experience working in program effectiveness within two separate disciplinary contexts and with three different accrediting bodies. That experience has led me to believe that well-executed program effectiveness practices have wide-reaching, positive outcomes for the programs that carry them out. At the same time, working within the context of an institution of higher education where accreditation—whether institutional or programmatic—is a not infrequently discussed topic, I've been surprised by the general attitude of irritation that most of my colleagues express towards accreditation. Accreditation is often dismissed as a nuisance that is done because it's required, as opposed to an opportunity to spend the time to truly evaluate, from top to bottom, the effectiveness of a program or institution. Perhaps through outlining best practices and describing an overall attitude and approach that might be different from what most people think of when they think of accreditation, more colleagues in the academy—both those within programmatically accredited units and those in units without programmatic accreditation—will begin to embrace practices that allow data-driven approaches to continuous quality improvement within their units. I believe that the vast majority of those working in academia would embrace a

process that truly sought to improve the effectiveness of the work that they do if it were done in a way that was feasible and led to demonstrable positive change. This belief is not hypothetical; I've seen shifts in the attitudes of administrators, faculty, and staff towards effectiveness practices as a result of the implementation of well-designed approaches towards program evaluation. In the current study, I used a social constructionism lens to engage with the material I gathered from my interviews with practitioners within the field of program effectiveness and accreditation. The following conclusions are constructions based on my own personal experiences within the field of program effectiveness and the experiences of the participants in this study. And while one might get the sense from Chapter Four, and the conclusions in this chapter, that my background in nursing is biasing me towards favoring the approach of the field, I think the results and conclusions I present support that the type of approach undertaken by nursing leads to better outcomes.

Conclusions

In this section I will revisit the guiding research questions for the study and discuss the conclusions garnered through the process of data collection and analysis:

Research Question 1

How do practitioners at two western, doctoral-granting R-1 institutions of higher education understand and experience their programs' formation and implementation of their program effectiveness plans?

Two themes emerged in response to this research question. The first theme was the difference between effectiveness practices and effectiveness plans. This theme addresses how

accrediting bodies' effectiveness requirements reflect the effectiveness practices that those accredited by them employ. The second theme was the history of the formation of the programs' effectiveness plans. This theme addresses inheriting practices and the building of effectiveness practice knowledge through experience. Both themes are described in greater detail in the following sections.

Difference Between Effectiveness Practices and Effectiveness Plans. The comprehensiveness of the effectiveness practices for engineering and nursing were driven by the requirements of the accrediting bodies that they were working with. Engineering employed an approach that I will call an implementation of effectiveness practices. These practices were focused almost solely on the assessment of student learning outcomes. Nursing utilized effectiveness plans that outlined processes for comprehensive evaluation of the program in toto. The difference between practices and plans had implications for the focus of effectiveness, the time and resources employed by the programs in the study, and the nature of the thinking surrounding change-making in the programs.

The focus of the effectiveness practices within the engineering programs in this study was almost wholly on the assessment of student learning outcomes. The programs described a documented process of continuous evaluation of their students' attainment of course and program outcomes; however, even when pushed, they mostly had no real process in place for the evaluation of other aspects of their programs; things like evaluation of mission and program goals, fiscal resources, faculty outcomes, adequacy of facilities, and support services did not fit within the scope of their ongoing evaluative processes. This is why I have decided to term what I saw in the engineering programs in this study as "practice" versus the "plan" approach that I saw

in nursing. The nursing programs in the study saw the assessment of student learning outcomes as being embedded within a much broader system of program evaluation. That broader system incorporated the intentional evaluation of those things listed as absent in engineering's approach above. The comprehensive system that the nursing programs used to evaluate their programs were documented and continuous. They were comprehensive evaluation plans that systematically looked at every portion of their accreditation requirements in an ongoing way. This disciplinary difference in approach shows some support for the research of LoCasio (2010), who found that health science faculty placed a greater importance on accreditation standards for effectiveness than those in other disciplines. This study also shows at least partial support for literature that suggests professional programs have a greater propensity towards engaging in effectiveness practice (Ewell, 2009; Welsch & Metcalf, 2003b). Taken together, I suggest that there potentially is a hierarchy of support for effectiveness in accredited programs, with health professions showing more comprehensive support for effectiveness than other professional programs like engineering, and these other professional programs then showing greater support than non-professional programs.

The structure of the program, availability of resources, and degree of centralization of the process also contributed to the robustness of the disciplines' accreditation practices. The loosely coupled structure of the engineering programs in the study, combined with the more bounded effectiveness method that they employed, correlated with an approach that was more limited in time and resource allocation. The smaller numbers of faculty within an engineering department, as well as the various pulls on that faculty's time, seem to explain to some extent the fact that the engineering programs' engagement with their effectiveness practices was episodic—often

annual—and that very few people were involved in the process of the implementation of the effectiveness practices. Conversely, the nursing programs in the study were more tightly coupled, allowing for more efficient centralization of resource deployment. The nursing programs engaged more people in the process and were involved with their effectiveness practices continually.

The extent to which the disciplines in the study incorporated their change process within their program evaluation efforts also varied. Engineering's approach to broader-level innovation and change within their programs was distinct from the approach of the nursing programs in the study. Only one of the engineering participants described a defined approach to change and innovation. The descriptions given for the processes involved in change-making in engineering were generally ad hoc and driven by intuition. Generally, ABET was seen as a constraining factor—both stifling innovation, but also suppressing half-baked ideas. Unsurprisingly, there did not appear to be a defined method of evaluating the effectiveness of the changes undertaken. By contrast, the nursing programs in the study described defined, systematic approaches to change and embedded innovation within their effectiveness plans. Built into that planning were processes for evaluation of the effectiveness of the changes made.

As discussed in Chapter Four, these distinctions are most likely a reflection of the wording and structure of the accreditation expectations within the two main accrediting bodies that the programs engaged with. The only engineering participant that described a more systematic approach to change explicitly stated that it existed outside of the scope of their ABET requirements. The fact that most of the engineering participants did not have this type of process, however, speaks to the importance of accrediting bodies' expectations. Many of the participants

discussed how one of the main values of accreditation lies in the fact that it forces these units to do certain things that they might not otherwise do because of the constraints on time and resources that everyone in academia faces. And so, this study supports existing literature that suggests that accreditation is one of the main drivers of quality improvement practices (Eaton, 2012; Head & Johnson, 2011; Larkan-Skinner, 2015; Oden, 2009; Young, 2013). Not explicitly requiring a more comprehensive approach to overall program evaluation means it's more likely that that type of approach is not undertaken due to the perception that they would require resources and time that the programs simply do not have.

We can see that the formation and implementation of program effectiveness practices are highly dependent on the approach of the accrediting body. The engineering participants who had been around long enough mentioned how changes to ABET's expectations with the Engineering Criteria 2000 had led to significant changes in their practices and a shift towards focusing on an evaluation of program learning outcomes. This change was a reaction to increasing calls for accountability in higher education at the end of the 20th century (Barczykowski, 2018; Brittingham, 2009; Ewell, 2001) and was a change that ABET led at the time. However, at this point nearly a quarter of the way through the 21st century, many accrediting bodies have gone beyond the requirements that ABET outlined at the turn of the century. They have explicit requirements for comprehensive, ongoing evaluation of all aspects of the programs that they accredit.

Not only that, but ABET's criterion are looking a bit outdated when compared to the Commission on Collegiate Nursing Education (CCNE) with regards to its focus on inputs, even though initially ABET's Engineering Criteria 2000 were hailed as a move towards an outputs-

based evaluative framework. For example, Criterion 5 for the Criteria for Accrediting Engineering Programs, 2024-2025, focuses on curriculum. Part of the Criterion requires, "a minimum of 30 semester credit hours (or equivalent) of a combination of college-level mathematics and basic sciences with experimental experience appropriate to the program." Another section requires, "a minimum of 45 semester credit hours (or equivalent) of engineering topics appropriate to the program, consisting of engineering and computer sciences and engineering design, and utilizing modern engineering tools." In contrast, CCNE's similar Standard III, Key Element B, reads:

> Baccalaureate curricula are developed, implemented, and revised to reflect relevant professional nursing standards and guidelines, which are clearly evident within the curriculum and within the expected student outcomes (individual and aggregate). Baccalaureate program curricula incorporate *The Essentials of Baccalaureate Education for Professional Nursing Practice.* (p. 13)

In the elaboration to this Key Element with CCNE we find: "The baccalaureate degree program incorporates professional nursing standards and guidelines relevant to that program and each track offered. The program clearly demonstrates where and how content, knowledge, and skills required by identified sets of standards are incorporated into the curriculum" (p. 13). When comparing the two standards, we see how prescriptive ABET is with its requirements, with specified inputs of minimum credit hours, as opposed to an approach like CCNE's, which focuses on the outcomes that the students gain through the program design and being less prescriptive with the inputs. The differences seen in the two fields' attitudes towards evaluation along Ewell's accountability versus improvement spectrum (2008) supports the literature that suggests that movements towards outputs-based evaluation help in promoting the improvement paradigm within academic units (Welsch & Metcalf, 2003a). In sum, accrediting bodies like CCNE that explicitly require programs seeking accreditation to comprehensively evaluate all

aspects of their program and who have a fairly strict focus on outputs, force units seeking accreditation to grapple with a level of effectiveness practice that rise to a systems-level approach—the nursing programs in this study showed true effectiveness plans, not just effectiveness practices, and they seemed to be operating very much in the realm of the improvement paradigm.

I think this is an important area for programmatic accrediting bodies to consider in the formation of their accreditation requirements: what type of approach do they want their accredited programs to take with their continuous improvement processes? This study suggests that the type of approach taken by CCNE leads to a more rigorous approach to nursing programs' evaluative frameworks. This study also supports literature around the importance of developing a culture of evaluation that is focused on improvement as opposed to accountability (Bush, 2016; Larkan-Skinner, 2015; Lewis, 2007). This might also explain why, while the engineering programs didn't explicitly express a lack of faculty buy-in, the difference in the levels of engagement between nursing and engineering was quite clear, with nursing engaging more people in the process. This seems to support literature that suggests faculty buy-in is best garnered when it's focused on improvement (Welsch & Metcalf, 2003a) and that that buy-in is essential in promoting changes to organizational culture around evaluation (Bush, 2016; Larkan-Skinner, 2015; Procopio, 2010). Also, I would say that it most likely leads to better approaches to the overall evaluation of their programs and the comprehensiveness of their quality improvement efforts. Instead of hoping that programs would engage in these practices because they lead to better outcomes, it might behoove accrediting bodies to require it.

Development of Expertise and Formation of Effectiveness Practices. The second theme that emerged with regard to the first research question was one around the formation and implementation of the effectiveness practices of the programs in the study. In essence—how did the effectiveness practices come into existence and who carried them out. Through the interviews with the participants, it became apparent that in engineering inheritance was the main method that the programs' practices came into existence; further, the engineers in the study primarily came to their effectiveness roles as novices and gained experience and mastery on the job. Both of these points serve as contributions to the literature, as the formation of effectiveness practices within academic units and the professional preparation of practitioners of effectiveness are largely undocumented.

All of the engineering participants discussed the formation of the effectiveness practices at the units where they had worked as being one of inheritance and iteration. That is, when they came into their positions they inherited the existing practices of the academic unit and then, over time in their roles, iterated as they gained in experience and knowledge. A couple of the participants expressed the importance of learning from their roles as site evaluators, or through workshops and conferences with different bodies affiliated with accreditation, but almost all of them expressed the beginning state of their effectiveness practice as inheriting the one that had previously existed and making changes as they learned more, or as they evaluated the data collected from their programs. We can take from this that the practitioners in program effectiveness in engineering that were involved in this study came to it not necessarily because they were experts in the field. Conversely, they become experts in the field through the experiences that they gained.

In contrast, both nursing programs described large structural changes to the evaluative practices within their programs when the current champions moved into their roles. I think we see from this that the people involved in the effectiveness practices in the nursing programs in this study had a clearer understanding of what they wanted out of their practices at the outset than the engineering participants did. This is most likely explained by the fact that both of the nursing programs in the study employed people in roles related to their effectiveness practices who had been hired, at least in part, due to their existing professional experience in effectiveness planning. Both my position, and Casey's at Institution 2, are very much the type of roles that exist because of expertise in effectiveness practice. In short, the nursing programs in the study showed an acknowledgement of program evaluation as a profession in and of itself and that they sought professionals when thinking about their hiring decisions.

This supports Rickards and Stitt-Bergh (2016) and Ewell (2011) who suggested that as the expectations for accountability and evaluation continue to increase, the field of evaluation in higher education will come to be led increasingly by professionals in evaluation as opposed to disciplinary experts. I expect the influence of organizations like the American Evaluation Association (AEA), which formed in 1986, a year after the 1985 mark that Ewell (2008) posits as "the beginning of assessment as an explicit topic of policy discussion in the United States," to increase (p. 7). This association is composed explicitly as one "devoted to the application and exploration of evaluation as a profession" (American Evaluation Association, n.d.) and in 2018 published a list of five evaluator competencies for practicing professionals (American Evaluation Association, 2018). And while AEA does not explicitly refer to accreditation, it certainly includes "evaluation in education" as one of their main types of evaluation. At the very least, AEA reflects

the budding profession of evaluation as one that requires specific competencies that are distinct from individual disciplines and contexts.

Research Question 2

How has the context of institution and discipline influenced practitioners' program effectiveness practices?

In this section I will explore the themes that emerged around the influence of each program's institutional and disciplinary context. With institutional influence, there were subthemes of the ability to promulgate cultures of evaluation at the institutional level as well as the ways in which the programs' effectiveness practices fit within the larger context of the institutions' effectiveness practices. I explore these themes in greater detail below.

Institutional Influence. When looking at the literature around programmatically accredited units in higher education, there's some suggestion that these programs help the overall institution to develop a greater culture of evaluation (Brescianni, 2009). This finding would be important because the research also suggests that lack of expertise in evaluation at the institutional level is a challenge for accreditation (Larkan-Skinner, 2015). So, if programmatically accredited units help to promote the institution's culture of evaluation it could influence an institution's decision to engage in the expensive process of programmatic accreditation. The findings in this study with regard to the influence of programmatically accredited units on the institution's culture of evaluation were mixed, however, with the amount of influence varying by institution and departments.

The Assessment and Outcomes Committee at Institution 1 is an institutional-level committee with representatives from every college. Kathy with the nursing program at Institution

1 was the representative for the College of Nursing and there did seem to be an endorsement that her contributions to the committee helped in sharing some of the nursing programs experiences to other units with less experience in evaluation. However, the representative from the College of Engineering on this committee was the Dean of the college, and the individuals who I spoke with in engineering were basically unfamiliar with the work of the committee. There is no doubt an opportunity lost with the degrees of separation between where the work is being done and where the representation on this committee sits. This is an instance where the flatter organizational structure and tight coupling of the nursing program at Institution 1—where they don't have individual departments—probably led to a more direct connection between those involved in the actual work of effectiveness practices and institutional committee representation. It also has allowed for the nursing program at Institution 1 to contribute to the university discussion at the practitioner level and to be able to speak from experience in how to engage in effective program evaluation and develop a culture of evaluation.

The same question around the ability of practitioners in program effectiveness to positively influence the institution's culture of effectiveness was met with reservation by the participants at Institution 2. None of the participants could think of ways that the work they did in their units might influence the overall institution's culture of effectiveness. Nearly all of them made reference to how siloed the institution was. If there was a committee similar to the Assessment and Outcomes Committee at Institution 1, none of the participants at Institution 2 spoke to it. This may be institution-specific; it could also be related to size—while both institutions are large, Institution 2 is significantly larger than Institution 1. In sum, it's unclear how much having accredited programs might influence overall institutional culture with regard

to effectiveness practice. What little support I found would suggest that intentionality in institutional committee structure would be important to garner the greatest effect in promoting an institutional culture of evaluation drawing from programmatically accredited units' experience and could be an area of further study.

One area that was consistent with the literature on best practices for institutional influence on programmatically accredited units was the fact that the institutions were not forcing their institutional effectiveness structure or practices onto these programmatically accredited units (Brescianni, 2009). None of the participants expressed feeling burdened by the institution's effectiveness efforts. In fact, most of the participants expressed only a vague understanding of how their units' work was incorporated into the institution's accreditation. On the one hand, this is a positive because these units already do a great deal of work—no doubt more than non-accredited units—to evaluate their programs. Forcing these units into practices that don't align with their own accreditation could produce an increased and unnecessary burden. On the other hand, not connecting these experts of evaluation with other units in the institution to share lessons learned is another area of lost opportunity. And so once again, institutional leaders should be thinking about committee structure and organization if they want to see the work done by practitioners of program effectiveness within individual units spread to the institution more broadly. There is an efficiency lost by not capitalizing on these people's experience and potential influence on their peers.

Disciplinary Influence. Much time has already been taken discussing how the disciplinary differences in engineering and nursing education have contributed to differential levels of coupling and the focus of evaluative efforts between the disciplines. This is indeed a major

finding that penetrates most of the sections of this conclusion. Instead of rehashing some of those findings that are found in other sections here, instead I will focus on some of the areas of similarity that were found between the two disciplines and the implications that those similarities have when thinking about more dissimilar units on a college campus and the generalizability of the findings of this study.

There were, in fact, many ways that these two disciplines were quite similar. The have both been in existence for a very long time and both have been working with accreditation for periods of time than span decades—even close to the century mark. Both fields graduate students in preparation for professions that are well defined. Unlike a liberal studies major, for example, someone who graduates in nursing or engineering has a clearly defined career path, should they choose to walk it. Both fields also have professional, summative exams that can be used as a measure of program effectiveness, with the slight distinction that the passing of the NCLEX is required for licensure and practice as a Registered Nurse, while the Fundamentals of Engineering exam is not a requirement for most graduates to move into professional roles in engineering.

These similarities are important when discussing the findings of this research and thinking about their generalizability. Indeed, much of the research suggests that academic units who are used to outcomes-based assessments in practice are more likely to support and engage in effectiveness practice (Ewell, 2009; LoCasio, 2010; Welsch & Metcalf, 2003b). How do program effectiveness practices change within the context of a major like English or Art, where there aren't necessarily standardized outcome expectations to measure for students, nor a standardized test to benchmark against national averages? Nursing and engineering participants both spoke to the importance of gathering input from organizations that employ their graduates to make sure

their students were meeting the needs of their future employers. Again, how would one gather that type of feedback for majors where the possible employers are so varied? Ultimately, it seems easier for professional programs like engineering and nursing to triangulate data points to evaluate their outcomes than it would be for many of the other academic units that comprise the university.

Research Question 3

How do these practitioners' experiences and knowledge contribute to an understanding of generalizable (Ragin & Becker, 1992; Stake, 2006; Yin, 2003) best practice in program effectiveness?

Despite the institutional and disciplinary distinctions discussed above, there were still enough areas of commonality between the cases within the study to be able to draw conclusions as to best practices in program effectiveness. In the following section I will outline both the best practices discovered, revisit my framework from Chapter One, and discuss implications for the findings within the field of programmatic accreditation and effectiveness practice in higher education.

Best Practices in Program Effectiveness. This study supported the inference that I had coming into this study that most effectiveness practitioners in programmatically accredited academic units gain their knowledge through inherited practice and on the job experience. In nursing, where the hiring focused on experts, it still was found that those experts had gained their knowledge previously through on the job experience. This finding is not only significant because the preparation of effectiveness practitioners has not been studied, but also because it has ramifications for suggestions for best practice. If many of the practitioners of effectiveness are

coming to it not as experts, we need to think about how we prepare them to do their job well. The published standards or criterion for accreditation are typically written to be somewhat open-ended, but that lack of specificity can be a challenge for those new to effectiveness practices. Taking the time to discuss practices with practitioners who have been doing this work for many years allowed me to take a deep look at best practice, but also at areas of deficiency. Looking at these cases, the themes of support of organizational structures and resources, intentional effectiveness plan structure and documentation, feasibility, the development of a culture of evaluation, consideration of succession, and metaevaluation came through as themes for best practice. This list contrasts with Brescianni (2009), who developed a list of best practices for assessing student outcomes, where only the themes of resources and metaevaluation were shared by both this study and that of Brescianni. This may be partially due to the difference in scope of analysis, as this study focused on effectiveness versus just outcomes assessment.

I began my study with a Framework for Best Practices in Program Effectiveness Planning based off of the scant literature on program effectiveness:

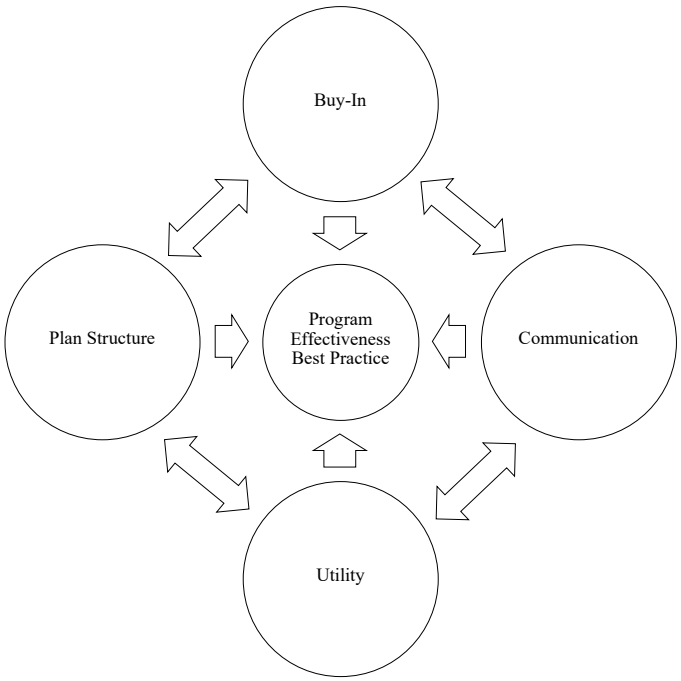

Figure 5. Initial Framework for Best Practices in Program Effectiveness Planning

Reflecting on this framework, I think it reflects the paucity of research at the level of programmatic accreditation where the issues of buy-in and communication are probably less pronounced than they are within the context of institutional accreditation, where the literature is somewhat more robust. This is significant because the extant literature on institutional effectiveness is heavily focused on buy-in (Bernecker, 2010; Birnbaum, 2000; LoCascio, 2010; Nichols, 1995; Procopio, 2010; Welsch & Metcalf, 2003a; Welsch & Metcalf, 2003b). As already discussed, I found that the issue of buy-in wasn't so much one of having to garner it; instead, it was one of how deeply engrained it was. My findings would suggest that buy-in, as a construct, is probably more dimensional than binary. In this study, the degree of buy-in also appeared to be related to a perception of time availability. Ultimately, buy-in for the programs I studied differed by how deeply the evaluative culture had penetrated the organization and how much time people

perceived they had to engage in the process. The question of whether or not general buy-in existed was moot—overwhelmingly the participants said that there weren't problems with buy-in in their programs.

Similarly, the importance of communication that showed some support in the literature (Bush, 2016; Ewell, 2009; Kezar, 2013; Oden, 2009; Young, 2013) was not a theme that I found supported in my work with the participants in this study. The research around this point was often focused on the attitudes of institutional leadership and the communication of the importance of effectiveness practices. Again, the difference between the institutional-level of accreditation and programmatic-level of accreditation, perhaps combined with the importance of the champion found at the program-level, seems to explain the difference.

Finally, most participants acknowledged that the stakeholders in the effectiveness process found utility in it—another area where this study's findings seem to vary with the literature on institutional accreditation (Bresciani, 2009; Larkan-Skinner, 2015, Lewis, 2007). This may be because focusing on utilization has become a foundational aspect of effectiveness practice, as Ewell (2009) suggested it might. It may also be related to the proximal locus of the effectiveness practices in programmatic accreditation versus those in institutional accreditation, which may seem more removed from everyday practice.

So, rethinking my conceptual framework at the tail end of the research that I conducted, I think the following model better describes what I found with my participants:

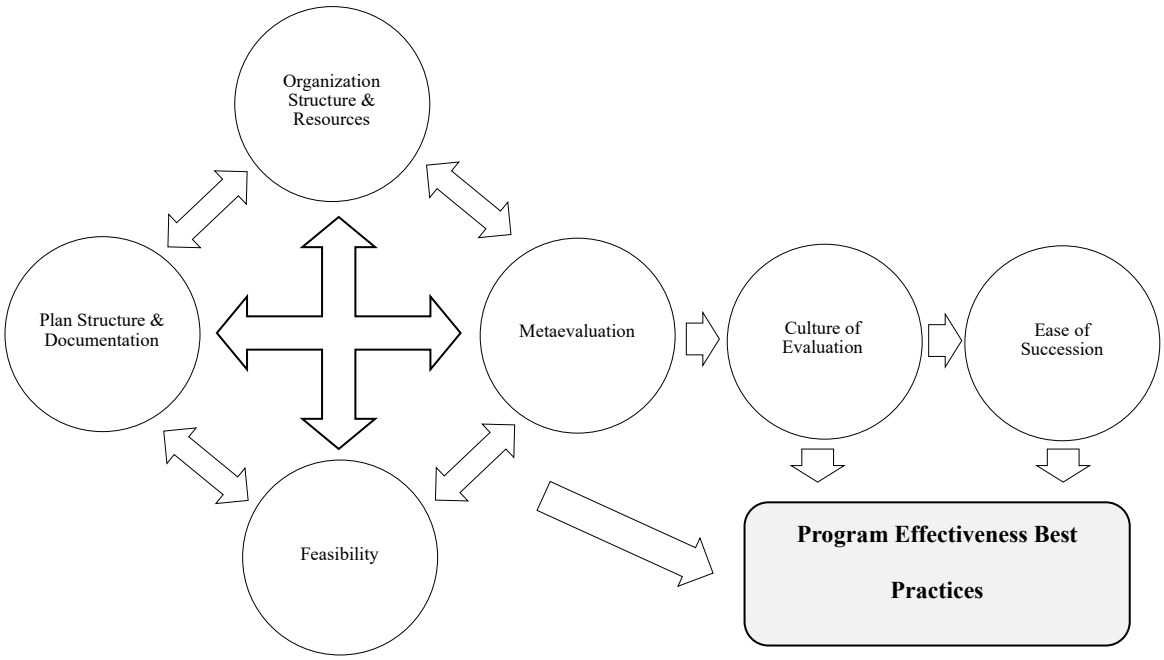

Figure 6. Framework for Best Practices in Program Effectiveness Planning

Like the initial framework, the practices of organization structure and resources, plan structure and documentation, feasibility, and metaevaluation all actively contribute to the others' success. For example, good organization structure and resource deployment enhance the ability to create an effective plan structure and documentation process; well-implemented plan structure and documentation processes can influence organization structure and resource allocation. Those four practices, themselves, produce an organization's culture of evaluation, which contributes to the ease of succession with changes in effectiveness practitioners. All aspects align to lead to best practices in program effectiveness.

These best practices are ones that any current or future practitioner of program effectiveness within a programmatically accredited unit can refer to as they're reflecting on their practice. It was clear that the individual contexts of each program influenced exactly how they addressed these practices, but I saw careful consideration of these points for each unit. In my

own practice, these themes also resonate as ones that have helped in the formation of an effectiveness practice that leads to positive programmatic changes.

The one best practice that the engineering programs, generally, did not seem to undertake in a way that leads to the best results was to have a written, systematic plan for ongoing comprehensive evaluation of their program's effectiveness. I am mentioning this frequently because in my experience the careful design of this effectiveness plan is the secret sauce that makes accreditation a powerful tool for far-reaching program improvement. As already stated, I think this is a problem with the way the ABET criterion are worded, and I'll have recommendations in the Implications for Practice section below.

Limitations of the Study

This study did not look at many of the other programmatic accrediting bodies and disciplines that operate in the context of higher education. While institutional theory would suggest that the structures and processes of these various accrediting bodies would be similar, not exploring more of them was a limitation. There is, obviously, a feasibility aspect to this as there are 96 U.S.-recognized institutional and programmatic accrediting bodies listed by CHEA. Looking at all of them in the way I did would be impossible for any one study. However, there were some differences in the accrediting bodies studied here that were significant to the findings and it may be that there are other important differences with some of these bodies that would shed more light on the landscape of programmatic accreditation.

Another nuanced aspect of the context of accreditation within higher education is that there are also accrediting bodies that institutions might deal with that are not directly involved with academics. For example, Institution 1's student wellness center is accredited by the

Accreditation Association for Ambulatory Health Care. Both institutions have an affiliated museum accredited by the American Alliance of Museums. The National Collegiate Athletic Association also has a structure for eligibility that's not dissimilar from the programmatic accreditation discussed here. How these accrediting bodies might be similar or dissimilar from the ones in the current study is unknown.

While my sample size snowballed during my study at Institution 1 to include Electrical and Computer Engineering, attempts at reaching the corresponding individual in the same field at Institution 2 were fruitless. I also was unable to gather participation from the Associate Dean at Institution 1. The sample, therefore, lost some symmetry of comparison that I had hoped to have in my initial design. It's a limitation that there was not another departmental voice in a separate field of engineering at Institution 2 nor the college-level perspective in engineering at Institution 1.

Recommendations for Future Research

The two institutions involved in this study are, in many ways, quite similar. The limited support seen in this study for the influence of accredited programs on the institution's overall culture of evaluation may be a result of the size and focus of the institutions. Perhaps smaller schools, or schools with a less research-focused mission, would be in a better position to allow effectiveness practitioners to mingle more frequently and meaningfully with people outside their academic areas and thus exert a greater influence on the institution's overall culture of evaluation.

As alluded to in the Limitations of the Study section, the entire landscape of accreditation within institutions of higher education can be quite vast. It might be interesting to study a couple

of institutions as individual cases but to look at the entirety of the accreditation processes that take place within them. It seems to me that there might be a great deal of inefficiencies that happen within institutions with different units all recreating the wheel, so to speak, as they implement their processes for program evaluation.

Implications for the Fields of Accreditation and Evaluation

Interestingly, the areas of deficiency identified have mostly been at levels higher than the units where my participants worked. While, generally, there was support for the isomorphic trend in accreditation that institutional theory would suggest, the areas of misalignment point out areas of thought for those involved in accreditation and also suggest that institutional theory was only partially supported in my findings. One of the first areas that the field of accreditation might consider is the consistency of terms. When reading about accreditation and program evaluation, or when discussing these topics with practitioners, you have to be constantly gauging the context and intent for the use of terms like assessment and evaluation, or effectiveness, continuous improvement, and continuous quality improvement. There are times when speakers are using these terms to mean the same thing, though there are also some distinct times when they are purposefully being used with differentiated meaning. Unfortunately, that kind of specificity is not common, even amongst the people I interviewed who are professionals in the area.

Ewell (2001), one of the earliest and most prolific researchers in assessment and outcomes in higher education, points to the challenges with terminology:

> *Assessment*, finally, refers primarily to the methods that an institution or program employs to gather evidence of student learning. But historically, the term has been applied in several ways. For accreditation purposes, the common meaning refers to the collection and use of *aggregated* data about student attainment to examine the degree to which program or institution-level learning goals are achieved. But the

> term assessment is also commonly used to describe the processes used to certify *individual* students or even, in some cases, to award grades. *Evaluation* also commonly refers to evidence-gathering processes that are designed to examine program or institution-level effectiveness. But the object of evaluation usually extends beyond learning outcomes to examine a much wider domain of institutional performance. (p. 7)

Ewell's suggestion in 2001 that the terms used around accreditation "remain relatively underdeveloped" (p. 5), unfortunately, remains true today. Indeed, Rodriguez (2021) was still making reference to these challenges just recently. In practice, Ewell has consistently used the term assessment when measuring student outcomes—even though it is, by definition, an evaluative process—and separates other aspects of effectiveness evaluation into "administrative processes" (2009, p. 17). I think this sectioning off of duties between assessment and "administrative processes" that Ewell points to may partially be the source of the challenge in understanding the comprehensiveness of effectiveness planning for both faculty and administration in institutions—the comprehensive evaluation of effectiveness planning requires the bridging of the pedagogical and administrative aspects of the academy and an understanding of how to measure success in each, both separately and together.

However, even if these two aspects are integrated successfully, the lexical challenge of inherited convention remains. Nursing at Institution 1, where they use the term Teaching and Learning Assessment to talk about the evaluation of course and program outcomes, is a good example. This nomenclature was adopted partly to distinguish this process from the overarching program evaluation that the program takes part in, but it also aligns with the tendency to use the term assessment when talking about evaluation of outcomes. In fact, the university itself uses the term assessment when discussing this evaluative process, so nursing's adoption of the term also aligns with the institutional use. So, we see again in practice that while both outcomes

assessment and program evaluation are evaluative exercises, the terminology is confused. Indeed, it may be that some of the misunderstanding around program evaluation for people not well-versed with the practice is due to the confusion of these terms—not only in everyday parlance, but within the field itself.

It's not within the scope of our influence for people who practice program effectiveness to change the conventions of English usage. However, it might behoove us to define terms consistently within the discipline so that even we, as experts, are aware of what's being discussed. I would propose a two-tiered approach to the defining process—one centered around teaching, learning, and curricula, which adopts the common terms already at use in the field, and one centered around the assessment and evaluation involved in the other aspects of the program. While distinct for clarity in terminology, the two tiers would mirror each other in their essential structure and processes. See Figure 7 below.

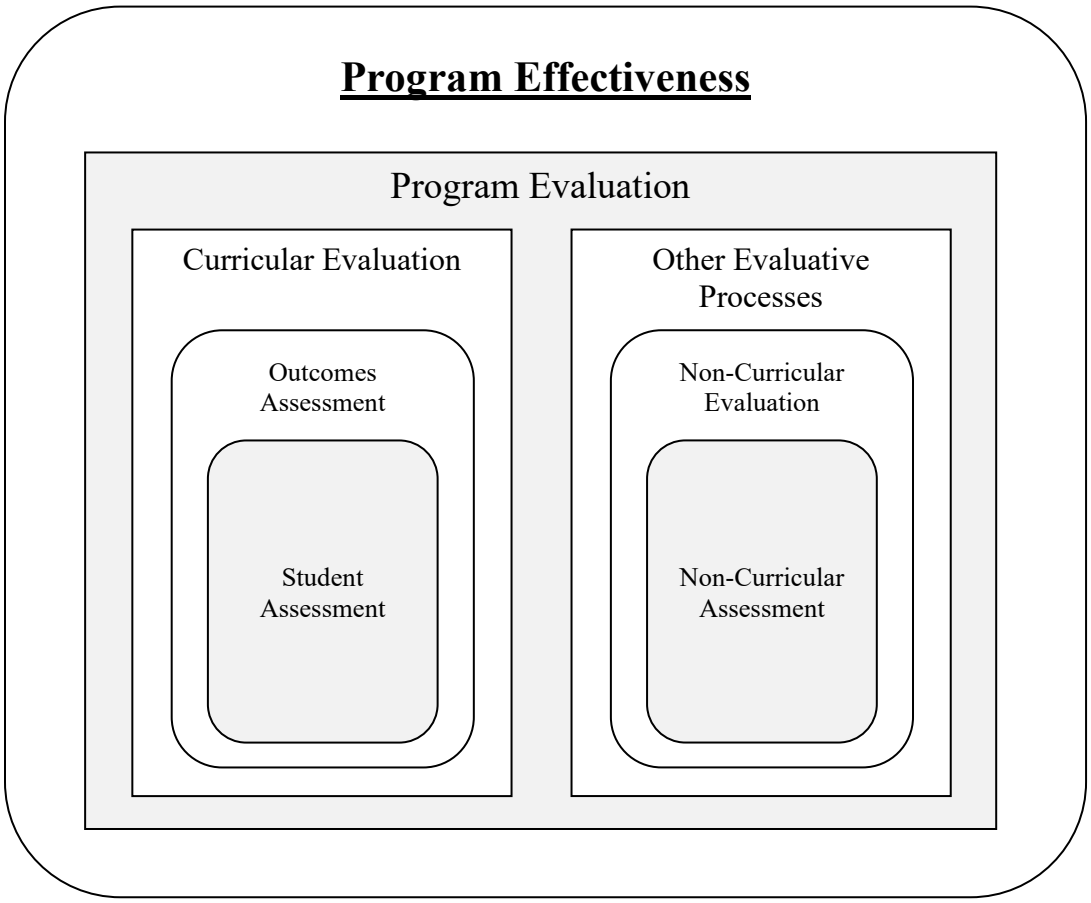

Figure 7. Proposed Nomenclature for Program Effectiveness

In this model, the teaching and learning side would use the term assessment when referring to artifacts that attempt to measure individual student competence (i.e., an exam given by a faculty member). Outcomes assessment would be used when referring to the more aggregated evaluation of a groups of students' achievement of course or program outcomes (i.e., a curricular committee using aggregate measures of an exam against a set benchmark aligned with a program outcome). While this second step is truly an evaluative exercise, common parlance has pegged it as outcomes assessment. That terminology is probably too established to change at this point.

The benefit of this approach is that we can keep the terms commonly used in outcomes assessment, but it allows us to distinguish between the evaluative processes related to teaching and learning and those related to other aspects of the program. So, we see that when evaluating non-curricular aspects of a program in this model we might have an assessment (i.e., a student survey sent to all graduating students), but then we would use the term evaluation to talk more generally about the significance of that non-curricular assessment (i.e., aggregated results from the student survey when compared against benchmarks). Both the curricular and non-curricular tiers would be sub-components of program evaluation—that is, the evaluation of the various aspects of the program.

Finally, effectiveness can be used to talk about the systems-level approach to evaluating a program's success in total. This definition of effectiveness would be similar to the Joint Committee on Standards for Educational Evaluation's definition of "design," but tailored to the context of the ongoing accreditation requirements in higher education (Yarbrough, 2010, p. 286), and would align well with the Southern Association of Colleges and Schools Commission on Colleges (SACSCOC) definition. It is important to note that it does not, however, align with how the term effectiveness is used with all accrediting bodies, which is part of the problem.

In this definition, effectiveness contains information regarding the program's assessment, evaluation, and continuous improvement or constant quality improvement practices, but would be reserved to refer to the overarching plan and approach to collecting data and measuring success in total. In this usage, the term of effectiveness would be in reference to a systems-level approach to answering the question of "is this program successful?" The nuanced differences in these definitions would allow people to neatly understand the processes involved at each level

and their roles within those processes. If this approach, or something similar, were adopted more broadly by accrediting bodies it would allow practitioners to speak more easily with each other and to know clearly what the other was talking about.

Another less-critical lexical example is the difference in the terms used for accreditation expectations by the various accrediting bodies. CCNE uses the term standard to refer to their accreditation expectations. This is a commonly used term. However, ABET and ACME (an accrediting body that both nursing programs in this study work with in addition to CCNE) use the term criterion. This also appears to be a commonly used term. Both terms, however, refer to the exact same thing. Again, the difference in terms makes talking across disciplines or accrediting bodies more challenging for those who have only been exposed to their own context's usage. Perhaps a body like the Council for Higher Education Accreditation (CHEA) could take the charge in suggesting consistent terminology, particularly as it appears that there is an increasing professional class of evaluators and effectiveness practitioners whose experience in those areas allow them to cross disciplines.

The final implication for practice revolves around the difference in expectations for effectiveness practices already discussed at various points in this and the previous chapter. To be clear, this isn't a discussion around the definition of the term. However, the definition I just gave for effectiveness presages what is to come. This discussion is really around the ultimate goals and scope of accreditation. It seems limiting to reduce the requirement for ongoing evaluation of the program to only the assessment of student learning outcomes. Demonstrating that a program is designed such that students attain a level of competency required for the field of study is incredibly important for a program's accountability. However, the scope of the work done within

academic units at institutions of higher education, and the scope of what academic programs should be evaluating to ensure they're doing their best in all of the areas that they engage in, is much broader than just the design of their curricula and the outcomes of their students, as important as those are. Having something like accreditation demand that programs evaluate their success *in its totality* is a way of ensuring that these units are continuously making data-informed decisions at all levels and evaluating the results of the changes that they make.

Programmatically accredited units already report comprehensively on these aspects when they complete their self-studies during their reaffirmation process, and so the only true addition would be extending the ongoing evaluative requirement that exists with ABET for curricula and student outcomes to all the other areas evaluated by the accrediting body during their cyclical reaffirmation, which for engineering is every six years. This would mean the addition of, at a minimum, an ongoing evaluation requirement for faculty preparedness and outcomes, adequacy of facilities, and institutional support in the case of ABET, though many accreditors require an evaluation of mission and goals as well. While there would no doubt be complaints from programs like engineering whose accrediting bodies do not currently require such a comprehensive ongoing approach, the programs of nursing in this study have demonstrated the feasibility of implementing such an approach. I would say from my own experience in two separate fields that required a more comprehensive approach, programs would come to find greater value in the process if it were more comprehensive, as long as it was executed well. The results of the process lead to questions and solutions that are more inclusive and that cross more boundaries than a singular focus on student outcomes could ever produce.

It may be that ABET is an outlier in their approach. When looking at AACSB, an accrediting body for business in higher education, the very first standard of the 2020 Guiding Principles and Standards for Business Accreditation describes the requirement for a comprehensive effectiveness planning process, though again we see the problem of inconsistent terminology as they refer to it as a Strategic Plan (AACSB, 2023). The National Architectural Accrediting Board (NAAB), in their 2020 Edition of their Conditions for Accreditation, lists the requirement for a comprehensive effectiveness plan in Condition 5.2, (National Architectural Accrediting Board, 2020). Notice again the inconsistent use of terminology, as NAAB refers to their requirements as conditions. Indeed, all of the regional accrediting bodies also incorporate language around the need for comprehensive and systematic evaluation that, at the very least, is not limited to just student outcomes, although we again find lack of consistent vocabulary (Accrediting Commission for Community and Junior Colleges Western Association of Schools and Colleges, 2024; Higher Learning Commission, 2020; Middle States Commission on Higher Education, 2023; New England Commission of Higher Education, 2021; Northwest Commission on Colleges and Universities, 2020; Southern Association of Colleges and Schools Commission on Colleges, 2024; WASC Senior College and University Commission, 2023). So, it does appear that it is at least not uncommon for accrediting bodies to include a comprehensive effectiveness planning requirement. In fact, these examples would suggest that ABET may be relatively unique in its limited scope.

So, it seems odd that ABET's criterion do not explicitly do this. CHEA, who recognizes accrediting bodies in a process similar itself to accreditation, in its 2021 CHEA Standards and Procedures for Recognition, does include in Standard 1, Academic Quality and Student

Achievement, a requirement for accrediting bodies to enforce standards that "advance academic quality using quantitative and/or qualitative measures" and "support implementation of innovative practices" (Council for Higher Education Accreditation, 2021, p. 9). I think they're missing an opportunity in this wording to explicitly require accrediting bodies do what a large number already appear to be doing: include comprehensive effectiveness planning as a required part of their accreditation requirements. A more explicit requirement might also help to resolve some of the ongoing problems that programs and institutions have related to their program effectiveness requirements, as the research shows that this requirement tends to be the most common area of compliance problems in accreditation (Head & Johnson, 2011; Manning, 2011; Southern Association of Colleges and Schools Commission on Colleges, 2021; Rodriguez, 2021). Rodriguez (2021), in a similar vein to my argument here, posits lack of clarity around the definition of effectiveness planning as a contributing factor in institutions' and programs' ability to consistently meet the effectiveness requirement in accreditation. That research and this study combine to suggest that this is an area where the field of accreditation could improve.

Chapter Summary

This chapter reviewed the major conclusions derived from the study. I set out in this study to try to understand the practices of people involved in program effectiveness in units that were programmatically accredited. The hope was to gain a deeper understanding of their work and, through that understanding, discover best practices for implementing program effectiveness strategies in programmatically accredited academic units. Through this research and a review of the literature it quickly became apparent that accreditation is integrated into institutions of higher education in more levels than I had understood. As a time and resource-intensive process aimed

at program improvement, it's important for institutional decision-makers understand the process, its benefits, and its challenges.

REFERENCES CITED

AACSB (2023). 2020 Guiding Principles and Standards for Business Accreditation (Updated). https://www.aacsb.edu/-/media/documents/accreditation/2020-aacsb-business-accreditation-standards-june-2023.pdf?rev=d31cfbe864e54792816ff426fe913e65&hash=33A159779F107443A64BDACBBB7000C5.

ABET (n.d.). History. https://www.abet.org/about-abet/history/.

ABET (2024a). Accreditation Policy and Procedure Manual (APPM), 2024-2025. https://www.abet.org/accreditation/accreditation-criteria/accreditation-policy-and-procedure-manual-appm-2024-2025/.

ABET (2024b). Criteria for Accrediting Engineering Programs, 2024 – 2025. https://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2024-2025/.

Accrediting Commission for Community and Junior Colleges Western Association of Schools and Colleges (2024). Accreditation Handbook (January 2024 Edition). https://accjc.org/wp-content/uploads/Accreditation-Handbook.pdf.

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, *50*(2), 179-211.

Alstete, J. W. (2007). A brief history of college accreditation. In *College Accreditation: Managing Internal Revitalization and Public Respect* (pp. 11-21). Palgrave Macmillan, New York.

American Association of Colleges of Nursing. History of Accreditation. https://www.acenursing.org/about/acen-history-of-accreditation/.

American Association of Colleges of Nursing (n.d.). Who we are. https://www.aacnnursing.org/CCNE-Accreditation/Who-We-Are.

American Association of University Professors. 1940 Statement of Principles on Academic Freedom and Tenure. Retrieved August 1, 2022, from: https://www.aaup.org/report/1940-statement-principles-academic-freedom-and-tenure.

American Evaluation Association (2018). *The 2018 AEA Evaluator Competencies*. https://www.eval.org/Portals/0/Docs/AEA%20Evaluator%20Competencies.pdf.

American Evaluation Association (n.d.). Values of the American Evaluation Association (AEA). https://www.eval.org/About/About-AEA.

American Medical Association. AMA History. Retrieved August 1, 2022, from https://www.ama-assn.org/about/ama-history/ama-history.

Anzul, M., Ely, M., Freidman, T., Garner, D., & McCormack-Steinmetz, A. (2003). *Doing qualitative research: Circles within circles*. Routledge.

Astin, A. W. (1977). Four Critical Years. Effects of College on Beliefs, Attitudes, and Knowledge.

Baker, K. A. (2007). Organizational communication. *Management Benchmark Study*, *1*(1), 1-3.

Banta, T. W. (Ed.). (2002). *Building a scholarship of assessment*. John Wiley & Sons.

Barczykowski, J. L. (2018). *The Goals of Specialized Accreditation: A Study of Perceived Benefits and Costs of Membership with the National Association of Schools of Music.* (Doctoral dissertation, University of Southern California).

Bardo, J. W. (2009). The impact of the changing climate for accreditation on the individual college or university: Five trends and their implications. *New Directions for Higher Education*, *145*, 47-58.

Beno, B. A. (2004). The role of student learning outcomes in accreditation quality review. *New directions for community colleges*, *2004*(126), 65-72.

Bernecker, K. J. (2010). *A Comparison of the Perceptions of Four Groups of University Constituents Concerning Institutional Effectiveness at a Private University*. ProQuest LLC.

Birnbaum, R. (2000). The life cycle of academic management fads. *The Journal of Higher Education*, *71*(1), 1-16.

Bowen, H. R. (1977). Investment in Learning: The Individual and Social Value of American. *Higher Education*, 142-57.

Boyd, R. G. (1973). The development of accreditation and its influence upon curriculum development in higher education. *Journal of Thought*, 188-196.

Bresciani, M. J. (2006). *Outcomes-based academic and co-curricular program review: A compilation of institutional good practices*. Stylus Publishing.

Bresciani, M. (2009). Evaluating the efficiency, effectiveness and sustainability of outcomes-based program review. *The Journal of Faculty Development*, *23*(1), 30-39.

Brittingham, B. (2008). An uneasy partnership: Accreditation and the federal government. *Change: The Magazine of Higher Learning, 40*(5), pp. 32-38.

Brittingham, B. (2009). Accreditation in the United States: How did we get to where we are? *New Directions for Higher Education*, *2009*(145), 7-27.

Brubacher, J.S., Rudy, W. (2017). *Higher Education in Transition: A history of colleges and universities* (4th ed.). Routledge.

Bruhn, J. G. (1993). The nemesis of allied health accreditation: the emperor needs new clothes. *Journal of Allied Health*, *22*(4), 335-351.

Bush, K. C. (2016). *Preparing self-assessments for regional accreditation reaffirmation: A phenomenological study of higher education leaders* (Doctoral dissertation, Colorado Technical University).

Cai, Y., & Mehari, Y. (2015). The use of institutional theory in higher education research. In *Theory and method in higher education research*. Emerald Group Publishing Limited.

Challa, C. D., Kasper, G. M., & Redmond, R. (2005). The accreditation process for IS programs in business schools. *Journal of Information Systems Education*, *16*(2), 207.

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Sage Publications.

Cohen, A. M., & Kisker, C. B. (2009). *The shaping of American higher education: Emergence and growth of the contemporary system*. John Wiley & Sons.

Commission on Collegiate Nursing Education. (2018). *Standards for accreditation of baccalaureate and graduate nursing education programs*. Commission on Collegiate Nursing Education.

Congressional Research Service (2017). *An Overview of Accreditation of Higher Education in the United States.*

Cooke, M., Irby, D. M., & O'Brien, B. C. (2010). *Educating physicians: a call for reform of medical school and residency* (Vol. 16). John Wiley & Sons.

Council for Higher Education Accreditation (n.d.). *Browse Database and Directories*. https://www.chea.org/directories.

Council for Higher Education Accrediation (2021). CHEA Standards and Procedures for Recognition.

Creswell, J. (2013). Qualitative inquiry and research design: Five different approaches.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches,* 4th ed. Sage publications.

Crow, S. (2009). Musings on the Future of Accreditation. *New Directions for Higher Education*, *145*, 87-97.

Dahlberg, K. (2006). The essence of essences–the search for meaning structures in phenomenological analysis of lifeworld phenomena. *International journal of qualitative studies on health and well-being*, *1*(1), 11-19.

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2011). *The Sage handbook of qualitative research*. Sage Publications.

Dickeson, R. C. (2009). *Recalibrating the accreditation-federal relationship*. Council for Higher Education Accreditation. Washington, DC.

Eaton, J. S. (2007). Institutions, Accreditors, and the Federal Government: Redefining their "appropriate relationship". *Change: The Magazine of Higher Learning, 39 (*5), 16-23.

Eaton, J. S. (2012). The future of accreditation. *Planning for higher education*, *40*(3), 8.

Eaton, J. S. (2015). An Overview of US Accreditation. Revised November 2015. *Council for Higher Education Accreditation*.

Ewell, P. T. (2001). Accreditation and Student Learning Outcomes: A Proposed Point of Departure. CHEA Occasional Paper.

Ewell, P. T. (2002). An emerging scholarship: A brief history of assessment. *Building a scholarship of assessment*, 3-25.

Ewell, P. T. (2008). Assessment and accountability in America today: Background and context. *New Directions for Institutional Research*, *2008*(S1), 7-17.

Ewell, P. T. (2009). Assessment, accountability, and improvement. *NILOA Occasional Paper*, *1*

Ewell, P. T. (2011). Assessing student learning outcomes in college: The role of the states. *The states and public higher education policy: Affordability, access, and accountability*, 151-172.

Ewell, P. & Cumming, T. (2017). Introduction: History and conceptual basis of assessment in higher education. In T. Cumming & M. D. Miller (Eds), *Enhancing Assessment in Higher Education: Putting Psychometrics to work.* Stylus.

Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2011). *Program evaluation: Alternative approaches and practical guidelines*. Pearson Education Inc.

Flexner, A. (1910). The Flexner Report. *Medica/Education in the United States and Canada: The Carnegie Foundation for the Advancement of Teaching, Science and Health Publications Inc*.

Flood, J. T., & Roberts, J. (2017). The Evolving Nature of Higher Education Accreditation: Legal Considerations for Institutional Research Leaders. *New Directions for Institutional Research*, *172*, 73-84.

Friedlander, J., & MacDougall, P. R. (1990). Responding to mandates for institutional effectiveness. *New Directions for Community Colleges*, *72*(4), 93-100.

Gaston, P. L. (2014). *Higher education accreditation: How it's changing, why it must.* Stylus.

Geiger, R. L. (1995). The era of multipurpose colleges in American higher education, 1850-1890. *History of Higher Education Annual*, *15*, 51-92.

Gillen, A., Bennett, D. L., & Vedder, R. (2010). The Inmates Running the Asylum? An Analysis of Higher Education Accreditation. *Center for College Affordability and Productivity (NJ1)*.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. AldineTransaction.

Gruber, C. S. (1975). *Mars and Minerva: World War I and the Uses of the Higher Learning in America.* Louisiana State University Press.

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. *Handbook of qualitative research*, *2*(163-194), 105.

Hanna, K., Duvall, J., Turpin, R. L., Pendleton-Romig, K, Parker, S. (2016). Mastering the journey to continuous compliance: Process improvements for excelling at CCNE Accreditation. *SAGE Open, 6*(2). https://doi.org/10.1177/2158244016656231.

Harcleroad, F. F. (1980). Accreditation: History, Process, and Problems. AAHE-ERIC/Higher Education Research Report No. 6, 1980.

Hartle, T. W. (2012). Accreditation and the Public Interest: can accreditors continue to play a central role in public policy? *Planning for Higher Education*, *40*(3), 16.

Hawkins, H. (2007). Toward system. *The History of Higher Education* (3rd ed.), pp. 332-345. Pearson Custom Publishing.

Head, R. B., & Johnson, M. S. (2011). Accreditation and Its Influence on Institutional Effectiveness. *New Directions for Community Colleges*, *153*, 37-52.

Hegji, A. (2017). An Overview of Accreditation of Higher Education in the United States. CRS Report R43826, Version 7. Updated. *Congressional Research Service*.

Heriot, K. C., Franklin, G., & Austin, W. W. (2009). Applying for initial AACSB accreditation: An exploratory study to identify costs. *Journal of Education for Business*, *84*(5), 283-289.

Higher Learning Commission (2020). Criteria for Accreditation. https://www.hlcommission.org/Policies/criteria-and-core-components.html.

Ingersoll, G. L., Sauter, M. (1998). Integrating accreditation criteria into educational program evaluation. *Nursing and Health Care Perspectives, 19*(5).

Jackson, R. S., Davis, J. H., & Jackson, F. R. (2010). Redesigning Regional Accreditation: the impact on institutional planning. *Planning for Higher Education*, *38*(4), 9.

Johnson, E. L. (1981). Misconceptions about the early land-grant colleges. *The Journal of Higher Education*, *52*(4), 333-351.

Kelderman, E. (2009). Struggling colleges question the cost–and worth–of specialized accreditation. *The Chronicle of Higher Education*, *56*(8), 1-3.

Kezar, A. (2013). Examining non-tenure track faculty perceptions of how departmental policies and practices shape their performance and ability to create student learning at four-year institutions. *Research in Higher Education*, *54*, 571-598.

Kuh, G. D., & Ewell, P. T. (2010). The State of Learning Outcomes Assessment in the United States. *Higher Education Management and Policy*, *22*(1), 1-20.

Kuh, G. D., Kinzie, J., Schuh, J. H., & Whitt, E. J. (2005). Never let it rest lessons about student success from high-performing colleges and universities. *Change: The Magazine of Higher Learning*, *37*(4), 44-51.

Larkan-Skinner, K. D. (2015). *The role of SACSCOC recommendations in changing community college practices for institutional effectiveness: A quantitative analysis* (Doctoral dissertation, Texas Tech University).

LeCompte, M. D., & Goetz, J. P. (1982). Problems of reliability and validity in ethnographic research. *Review of educational research*, *52*(1), 31-60.

Lewis, L. C. (2007). *Closing the Loop Or Tightening the Noose?: An Examination of the Use of Student Learning Outcomes to Improve Dietetic Education*. University of Minnesota.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.

LoCascio, S. H. (2010). *Faculty Perceptions of Institutional Effectiveness*. (Doctoral dissertation, Capella University).

Longenecker, P. A. (2012). *Institutional assessment in colleges with multiple programmatic accreditations: A multiple case research study* (Doctoral dissertation, Capella University).

Lynch, D. K. (2018). *The cost of the accreditation process: A case study of North Carolina community colleges*. (Doctoral dissertation, North Carolina State University).

Maki, P. L. (2012). *Assessing for learning: Building a sustainable commitment across the institution*. Stylus Publishing.

Manning, T. M. (2011). Institutional effectiveness as process and practice in the American community college. *New Directions for Community Colleges*, *2011*(153), 13-21.

Mentkowski, M. (1991). Creating a context where institutional assessment yields educational improvement. *The Journal of General Education*, *40*, 255-283.

Meyer, J. W. (1977). The effects of education as an institution. *American journal of Sociology*, *83*(1), 55-77.

Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American journal of sociology*, *83*(2), 340-363.

Middle States Commission on Higher Education (2023). Standards for Accreditation and Requirements of Affiliation, 14th ed. https://www.msche.org/standards/fourteenth-edition/.

Miller, M. G. (2013). *Using Astin's IEO model to explain the development of a culture of assessment: A case study of a four-year university* (Doctoral dissertation, Southeastern Louisiana University).

Morphew, C. C. (2009). Conceptualizing change in the institutional diversity of US colleges and universities. *The Journal of Higher Education*, *80*(3), 243-269.

Murray, F. B. (2009). An Accreditation Dilemma: The Tension between Program Accountability and Program Improvement in Programmatic Accreditation. *New Directions for Higher Education*, *145*, 59-68.

National Architectural Accrediting Board (2020). Conditions for Accreditation. https://www.naab.org/wp-content/uploads/2020-NAAB-Conditions-for-Accreditation.pdf.

New England Commission of Higher Education (2021). Standards for Accreditation. https://www.neche.org/wp-content/uploads/2020/12/Standards-for-Accreditation-2021.pdf.

Nichols, J. O. (1995). A practitioner's handbook for institutional effectiveness and student outcomes assessment implementation. Agathon Press, 100 Newfield Avenue, Edison, NJ 08837.

Northwest Commission on Colleges and Universities (2020). NWCCU 2020 Standards. https://nwccu.org/standards/.

Oden, R. A. (2009). A College President's Defense of Accreditation. *New Directions for Higher Education*, *145*, 37-45.

Oermann, M.H. (Ed.). (2017). *A Systematic Approach to Assessment and Evaluation of Nursing Programs*. National League for Nursing. Wolters Kluwer.

Overbay, J. D., & Aaltonen, P. M. (2001). A comparison of NLNAC and CCNE accreditation. *Nurse Educator*, *26*(1), 17-22.

Pace, C. R. (1979). Measuring Outcomes of College. Fifty Years of Findings and Recommendations for the Future.

Patton, M. Q. (1975). *Alternative evaluation research paradigm*. Center for Teaching and Learning, University of North Dakota. Grand Forks, ND.

Patton, M. Q. (2008). *Utilization-focused evaluation*. Sage publications.

Patton, M. Q., Campbell-Patton, C. E. (2022). *Utilization-Focused Evaluation* (5th ed.). Sage Publishing.

Pavlakis, A., & Kelley, C. (2016). Accreditation in the professions: Implications for educational leadership preparation programs. *Journal of Research on Leadership Education*, *11*(1), 68-90.

Pinkham, F. O. (1955). The accreditation problem. *The ANNALS of the American Academy of Political and Social Science*, *301*(1), 65-74.

Procopio, C. (2010). Differing administrator, faculty, and staff perceptions of organizational culture as related to external accreditation. *Academic Leadership: The Online Journal*, *8*(2), 17.

Ragin, C. C., & Becker, H. S. (Eds.). (1992). *What is a case?: exploring the foundations of social inquiry*. Cambridge university press.

Rickards, W. H., & Stitt-Bergh, M. (2016). Higher education evaluation, assessment, and faculty engagement. *New directions for evaluation*, *2016*(151), 11-20.

Rodriguez, R. (2021). *Regional Accreditation of Southern Colleges and Universities* (Doctoral dissertation, The University of Texas at Arlington).

Saewert, K. J. (2013). Program Evaluation. *Research for Advanced Practice Nurses: From Evidence to Practice*, 283.

Scriven, M. (2000). Evaluation ideologies. In *Evaluation models* (pp. 249-278). Springer, Dordrecht.

Selznick, P. (1957). Law and the Structures of Social Action.

Shadish, W. R., Cook, T. D., & Leviton, L. C. (1991). *Foundations of program evaluation: Theories of practice*. Sage.

Shawen, N. M. (1983). The evolution of regional accreditation in higher education: The role of the North Central Association, 1895–1934. *Assessment and Evaluation in Higher Education*, *8*(1), 1-18.

Shibley, L. R., & Volkwein, J. F. (2002). Comparing the Costs and Benefits of Re-Accreditation Processes. AIR 2002 Forum Paper.

Stake, R. E. (2006). *Multiple Case Study Analysis*. Guilford Press.

Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage Publications.

Stufflebeam, D. L. Madaus, G., & Kelelghan, T. (Eds.). (2000). *Evaluation models*. Kluwer Academic.

Southern Association of Colleges and Schools Commission on Colleges (2017). *The principles of accreditation: Foundations for quality enhancement* (6th ed.).

Southern Association of Colleges and Schools Commission on Colleges (2021). *Most frequently cited* principles *in decennial reaffirmation reviews: Class of 2021*. https://sacscoc.org/app/uploads/2022/04/Most-Frequently-Cited-Principles_2021.pdf.

Southern Association of Colleges and Schools Commission on Colleges (2024). *The Principles of Accreditation: Foundations for Quality Enhancement* (2024 Edition).

Suskie, L. (2004). What are good assessment practices. *Assessing student learning: A common sense guide*.

Swing, R. L., & Coogan, C. S. (2010). Valuing assessment: Cost-benefit considerations. *NILOA Occasional Paper*, *5*.

Thelin, J. R. (2019). A History of American Higher Education (3rd ed.). Johns Hopkins University Press.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of educational research*, *45*(1), 89-125.

Trifts, J. W. (2012). The direct and indirect benefits and costs of AACSB accreditation. *SAM Advanced Management Journal*, *77*(1), 20-27.

U.S. Department of Education. (n.d.). *Accreditation in the United States*. https://www2.ed.gov/admins/finaid/accred/accreditation_pg2.html.

Volkwein, J. F., Lattuca, L. R., Harper, B. J., & Domingo, R. J. (2007). Measuring the impact of professional accreditation on student experiences and learning outcomes. *Research in higher education*, *48*(2), 251-282.

Walton, A. V. (2022). *Applying a Systems Thinking Approach to the Experience of Self-Study for Institutional Accreditation: A Phenomenological Study* (Doctoral dissertation, Drexel University).

WASC Senior Colleges and University Commission (2023). Standards of Accreditation. https://www.wscuc.org/handbook2023/#standards-of-accreditation.

Weick, K. E. (1976). Educational organizations as loosely coupled systems. *Administrative science quarterly*, 1-19.

Welsh, J. F., & Metcalf, J. (2003a). Cultivating faculty support for institutional effectiveness activities: Benchmarking best practices. *Assessment & Evaluation in Higher Education*, *28*(1), 33-45.

Welsh, J. F., & Metcalf, J. (2003b). Faculty and administrative support for institutional effectiveness activities: a bridge across the chasm?. *The Journal of Higher Education*, *74*(4), 445-468.

Whittemore, R., Chase, S. K., & Mandle, C. L. (2001). Validity in qualitative research. *Qualitative health research*, *11*(4), 522-537.

Wholey, J.S. & White, B.F. (1973). Evaluation's impact on Title I elementary and secondary education program management. *Evaluation, I,* 73-76.

Willis, C. R. (1994). The cost of accreditation to educational institutions. *Journal of allied health*, *23*(1), 39-41.

Wolcott, H. F. (1990). On seeking-and rejecting-validity in qualitative research. *Qualitative inquiry in education: The continuing debate*, pp. 121-152.

Wolcott, H. F. (1994). *Transforming qualitative data: Description, analysis, and interpretation*. Sage.

Wood, A. L. (2006). Demystifying accreditation: Action plans for a national or regional accreditation. *Innovative Higher Education*, *31*, 43-62.

Woolston, P. J. (2012). *The costs of institutional accreditation: A study of direct and indirect costs*. University of Southern California.

Yarbrough, D.B., Shula, L.M., Hopson, R.K., & Caruthers, F.A. (2010). *The Program Evaluation Standards : A guide for evaluators and evaluation users* (3rd ed.). Joint Committeee on Standards for Educational Evaluation. Sage Publications.

Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.). Sage Publications.

Young, A. L. (2013). The regional accreditation process at community colleges: A case study of effectiveness.

APPENDIX

INTERVIEW PROTOCOL

Interview Protocol Project: Theorizing Best Practice in Program Effectiveness Planning	
Time of interview	
Date	
Place	
Interviewer	
Interviewee	
Briefly describe the project	This project is aimed at discovering the best practices in program effectiveness planning by exploring the practices that practitioners employ.

Questions:

Participant Background:

1. How long have you been in your current role?
2. How long have you been involved in program evaluation and/or accreditation efforts?
3. Do you have any experience with these types of efforts at other institutions or in different programs?
4. If so, what differences do you see in practice between them?

Program Background:

1. Describe the history of your program.
2. What are the demographics of your students, administration, and faculty?
3. How many administrators are in your unit?
4. How many faculty are in your unit?
5. How many students are in your programs?
6. How many students are in your programs?

Effectiveness Plan's Formation & Creation

1. How do you understand how your program's plan for program effectiveness has come into its current state? (RQ1; Plan Structure)
2. Who was/is involved in the creation of the plan? (RQ1; Plan Structure)
    a. What level of engagement do you see by the various parties? Do they have differing motivations? (RQ1; Buy-in)
3. How does your plan integrate within the context of the institution? (RQ1, RQ2; Plan Structure)
    a. How do your practices contribute to institutional accreditation practices (RQ2; Plan Structure, Best Practices)
    b. How do your practices contribute to institutional assessment and evaluation processes (RQ2; Plan Structure, Best Practices)
    c. Does the integration of your program's practices create duplicative work with institutional requirements? (RQ2; Plan Structure, Best Practices)
4. Can you think of examples of important changes to your program effectiveness practices? (RQ1; Plan Structure, Utility)
    a. What were the motivations for those changes? (RQ1; Utility, Plan Structure)
    b. Who, or what, led the change? (RQ1; Buy-In. Communication)

Effectiveness Plan's Implementation

1. Who is involved in the implementation of the plan? (RQ1; Plan Structure)
2. What level of engagement do you see by the various parties? (RQ1; Buy-In)
3. Do they have differing motivations? (RQ1; Buy-In)
4. How are responsibilities communicated? By whom and to whom? (RQ1; Communication)
5. Would you say the process is driven more by compliance or by improvement? (RQ1; Plan Structure, Utility)
6. Would you say the process is more administratively-driven or faculty-driven? (RQ1; Plan Structure)
7. How would you rate the overall buy-in of the stakeholders? (RQ1; Buy-In)
    a. Is there any pushback to your program effectiveness practices?
8. Would you say that stakeholders perceive these efforts as being for compliance or for improvement? (RQ1; Plan Structure, Communication, Utility)
    a. How do you know?
9. How do you document work that's been completed? (RQ1; Plan Structure)
10. How is the work that's been done get communicated? (RQ1; Plan Structure, Utility, Communication)
    a. Who communicates to whom? (Communication)
    b. Structured or unstructured? (Plan Structure)
    c. What format or forums? (Plan Structure, Utility)

Institutional Context & Influence

1. How does the context of the institution you're in influence your practices? (RQ2, Plan Structure, Buy-In, Communication, Utility, Best Practices)
2. How does your program's effectiveness practices contribute to the institution's culture of assessment (RQ2: Best Practices, Buy-In, Utility)
3.

Lessons & Best Practices

1. How effective would you say your plan for effectiveness has been? (RQ3; Best Practices)
2. How would you define effectiveness? (RQ3; Best Practices)
3. Would you say your plan leads to changes? (RQ3; Utility; Best Practices)
    a. Can you site one or two specific examples in greater detail? (RQ3; Utility, Best Practices)
    b. What was it about your plan that you think facilitated the change process (RQ3; Plan Structure, Utility, Best Practices)
4. Would you say all parties involved understand the utility of the evaluation and assessment practices? (RQ3; Utility, Best Practices, Communication)
5. How do you evaluate the effectiveness of your practices (RQ3; Best Practices, Utility)
6. What are some of the challenges that you face in your effectiveness practices? (RQ3; Best Practices)
    a. Are there any ways you've been able to mitigate those challenges? (RQ3; Best Practices)
7. What are some of the things that work best with your effectiveness practices (RQ3; Best Practices)

Closing

1. Is there anything else you'd like to tell me about your experience with program effectiveness planning?

Thank the participant for participating in the interview. Assure them of the confidentiality of

responses and possible future interviews

Table 4: Interview Protocol