All Graduate Theses and Dissertations   Graduate Studies

Integrated Concept Mapping as a Method for Democratically Evaluating a Teacher Preparation Program in the Area of Classroom Assessment Proficiency

Carrie Elizabeth Ashcraft
Utah State University

INTEGRATED CONCEPT MAPPING AS A METHOD FOR DEMOCRATICALLY

EVALUATING A TEACHER PREPARATION PROGRAM IN THE AREA OF

CLASSROOM ASSESSMENT PROFICIENCY

by

Carrie Elizabeth Ashcraft

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

of

DOCTOR OF PHILOSOPHY

in

Education
(Curriculum and Instruction)

Approved:

_____   _____
Marla Robertson, Ph.D.   Amy Piotrowski, Ph.D.
Major Professor   Committee Member

_____   _____
Steven Camicia, Ph.D.   Lisa K. Boyce, Ph.D.
Committee Member   Committee Member

_____   _____
Sylvia Read, Ph.D.   D. Richard Cutler, Ph.D.
Committee Member   Vice Provost of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2022

ABSTRACT

Integrated Concept Mapping as a Method for Democratically Evaluating a Teacher

Preparation Program in the Area of Classroom Assessment Proficiency

by

Carrie Elizabeth Ashcraft, Doctor of Philosophy

Utah State University, 2022

Major Professor: Marla Robertson, Ph.D.
Department: Teacher Education and Leadership

This dissertation reviews current policies regarding Teacher Preparation Program (TPP) evaluation and addresses "accountability era" influences that often minimize factors related to equity and promote a limited view of teacher and school "effectiveness." This study addressed the problem of meaningful program evaluation directly tied to program goals and the desired outcomes for teacher candidates. Integrated concept mapping is proposed as an alternative method of evaluation because of its democratic framework for accountability. Using this method, a TPP in the Mountain West was evaluated in the area of classroom assessment. Classroom assessment was defined as the ongoing informal and formal assessment of students, characterized by a cycle of communication and feedback between the student and teacher. Internal stakeholders within a secondary school of education identified areas of classroom

assessment proficiency, selected desired outcomes for program participants, created an

evaluation tool, and carried out the evaluation.

(213 pages)

PUBLIC ABSTRACT

Integrated Concept Mapping as a Method for Democratically Evaluating a Teacher

Preparation Program in the Area of Classroom Assessment Proficiency

Carrie Elizabeth Ashcraft

This dissertation reviews current policies regarding Teacher Preparation Program (TPP) evaluation and addresses "accountability era" influences which often minimize factors related to equity and promote a limited view of teacher and school "effectiveness." Integrated concept mapping is proposed as an alternative method of evaluation because of its democratic framework for accountability. Using this method, a TPP in the Mountain West was evaluated in the area of classroom assessment.

ACKNOWLEDGMENTS

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Transformative approaches often utilize dialogic qualitative methods, ethically centered in cultural respect, human rights, and reciprocity. Here, the idea is that evaluation is inherently a valuing—and political—activity with the potential for political influence and that evaluators should guard against power imbalances by considering whose interests are served and whose voices are included.

(Cochran-Smith & Reagan, 2022, p. 8)

To begin, I am a faculty member within a Teacher Preparation Program (TPP) in the Western region of the U.S. The topic of TPP evaluation is of interest to me as I have participated in multiple program evaluations. Reviewing current literature about evaluation methods, still primarily rooted in "accountably era" beliefs about education, led me to consider if a more democratic approach to evaluation was possible and how such an approach may impact TPPs.

The "accountability era" began in the 1980s and had a strong influence on education research and practice by the 1990s and well into the 2000s. This era established the teacher as the most critical factor in student learning gains, applied a market-driven ideology to education, and often ignored community considerations and factors relating to equity (Cochran-Smith et al., 2018; McDiarmid, 2019).

A more democratic approach to evaluation, one that is organized and facilitated by stakeholders with a close association to programs, may have greater potential for program improvement by identifying unique program and community considerations, and identifying specific factors related to equity (Cochran-Smith & Reagan, 2022; McDiarmid, 2019;). It may also support the current shift away from outdated

"accountability era" methods, which have proven to be minimally effective, at best, due to a hyper-focus on summative assessment data, a limited view of teacher and school effectiveness, and which generalized or ignored specific equity factors (Ciccone, 2019) Cochran-Smith & Reagan, 2022).

After reviewing topics related to TPP evaluation in this chapter and the associated literature in Chapter II, I will present a mixed methods approach to evaluation currently used within other fields. This method, Integrated Concept Mapping for Evaluation and Planning (Kane & Trochin, 2007), prioritizes local stakeholder involvement within evaluation and aligns with democratic approaches to education delineated by John Dewey (1916) and other recent educational researchers and theorists.

Current Teacher Preparation Program Evaluation Methods

Generally, it is agreed that TPPs need evidence of fidelity in connection with intended outcomes (Kumashiro, 2015). Many also agree that current program evaluation measures lack the local focus needed for accountability to community stakeholders (Cochran-Smith et al., 2018; Ciccone, 2019; Mertens & Wilson, 2019). A lack of local focus, coupled with a hyper-focus on annual student summative assessment data, is often highlighted as a weakness of current evaluation methods within national reports and feedback provided by TPPs (McDiarmid, 2019). Some evaluation organizations have begun to recognize and include more community factors and issues relating to equity within their standards for evaluation; however, there is still a long way to go. For example, AAQEP (The Association for Advancing Quality in Educator Preparation), a

newer accrediting organization that began in 2017, includes "Respect for context and mission" among its seven design principles (AAQEP, 2022).

One exception within current evaluation methods that includes a more local focus and equity considerations is teacher performance assessments such as the Educative Teacher Performance Assessment (edTPA). Performance assessments have been recognized for their impact on the local level because the evaluation is directly tied to the service community. However, teacher performance assessments are often viewed as problematic in terms of implementation on the state and program level due to push-back from students, university faculty, and community stakeholders who are skeptical of new evaluation methods (Apple & Beane, 2007; Cochran-Smith et al., 2018).

It is also important to consider other aspects of the assessment are still tied to long-held beliefs about education forwarded during the "accountability era" (Cochran-Smith et al., 2017). For example, ranking among national performance assessment scoring norms is often required for graduation from TPPs and state licensing (Reagan et al., 2016). In sum, while teacher performance assessments have become more common, they have a limited impact on program improvement. Furthermore, they do not go far enough to shift TPP evaluation away from "accountability era" methods (Cochran-Smith et al., 2018).

Reforming Teacher Preparation Program Evaluation

Because annual student summative assessment data continues to drive accountability initiatives, the validity of current program evaluations is in question (Cochran-Smith et al., 2018; Cochran-Smith & Reagan, 2022). More than 10 years ago

the National Education Policy Center (NEPC) called for a change in accountability

measures for TPPs that prioritized year-end student assessment data and lacked local

considerations. In Holding Teacher Preparation Accountable, Cochran-Smith et al.

(2016), issued the following policy recommendations, detailing a shift away from the

narrow accountability measures:

- Policymakers must acknowledge and address the multiple factors—in addition to teacher quality—that influence student outcomes, including, in particular the impact of poverty, family and community resources, school organization and support, and policies that govern housing, health care, jobs, and early childhood services.

- Systems evaluating teacher preparation must produce results that preparation programs can use to change and improve curricula, practice-based experiences,

- and assessments—not results that simply grade programs without information about why or how particular results occurred or what might improve them.

- Systems evaluating teacher preparation programs must be built on policy mechanisms that have documented capacity to produce usable information for local and larger program improvement within a complex policy and political climate.

- There should be a conceptual shift away from teacher education *accountability* that is primarily bureaucratic or market-based and toward teacher education *responsibility* that is primarily professional and that acknowledges the shared responsibility of teacher education programs, schools, and policymakers to prepare and support teachers. (pp. 4-5)

This national call for expansion of accountability criteria by the NEPC is

juxtaposed against more recent recommendations from organizations, such as the Council

for the Accreditation of Educator Preparation (CAEP), which continues to maintain a

narrower focus on accountability. For instance, CAEP added a standard that "…requires

programs to produce valid and reliable evidence of their graduates' impact on students'

learning as well as other outcome data," which is tied to "accountability era" conceptions

of evaluation (McDiarmid, 2019, p. 118). The new component reads:

> 4.1 REQUIRED COMPONENT The provider documents, using multiple measures, that program completers contribute to an expected level of student-learning growth. Multiple measures shall include all available growth measures (including value-added measures, student-growth percentiles, and student learning and development objectives) required by the state for its teachers and available to educator preparation providers, other state-supported P-12 impact measures, and any other measures employed by the provider. (http://caepnet.org/standards/standard-4)

While the need for evidence such as "value-added measures" and "student-growth percentiles" is in question for many of the reasons stated by NEPC and detailed within their recommendations, so is the practicality of collecting this type of data. Many TPPs do not have the access or resources to gather and compile this data; instead, they rely on outside evaluation entities to provide and explain the data at a high cost to taxpayers and TPPs, in both money and time. Because the benefit and practicality of collecting this information are in question, seeking out new program evaluation methods which provide more complete and useful understanding of program strengths and areas for improvement are needed (Cochran-Smith et al., 2009, 2018; Cochran-Smith & Reagan, 2022).

Complying with Current Evaluation Methods

Due to the cost and complexity of gathering and analyzing summative assessment data, it is often generated and provided to TPPs through contractors, agencies, and accreditors who conduct TPP evaluation or supply data for program evaluation (Zeichner, 2011). This removed data collection, and the manner in which it is gathered and supplied to TPPs, distances the evaluation process from the individuals most closely associated with the program, the faculty, administrators, and the teacher candidates (Cochran-Smith

& Reagan, 2022). The distance makes it difficult for local stakeholders to see the relevance of the data and use it in specific ways to improve programs (Baker et al., 2010; McDiarmid, 2019). Under pressure, some TPPs have looked for ways to comply with the evaluation standards without using the data in meaningful ways that improve the program (McDiarmid, 2019; Peck & McDonald, 2014). Other programs make a concerted effort to use the information but find it difficult to assess the impacts of the data (Brown & Duguid, 2000; Imig & Imig, 2008; Walsh, 2013).

Education policies that focus on annual summative assessment data or minimize the contextual details of the local community lack crucial measures and evidence that should be included within program evaluation. Understanding community factors and current program initiatives should be an important factor within TPP evaluation (Cochran-Smith et al., 2018). When the focus of evaluation and the resources used for evaluation is directed toward compliance in areas that do not prioritize community contextual factors or involve local stakeholders in meaningful ways, it detracts from more impactful accountability methods. As a result, few TPP evaluations impact programs and their associated service communities in ways that will improve K-12 learning (Greenberg & Walsh, 2010).

Cochran-Smith et al. (2016) asserts democratic accountability methods are needed, "… to change and improve curricula, practice-based experiences, and assessments—[and should] not…simply grade programs without information about why or how particular results occurred or what might improve them" (pp. 4-5). Independent reviews, by national and local education stakeholders, have led some to call for changes

in the way TPPs are evaluated (Baker et al., 2010; McDiarmid, 2017, 2019).

Furthermore, education reformers also recommend evaluation criteria for TPPs that

includes data recognizing the relationship and contextual factors between TPPs, local

schools, and the community (Cochran-Smith et al., 2016). In sum, while program

evaluation has the potential to ensure teachers entering the profession are prepared for the

classroom, more meaningful evaluation methods are needed to ensure that TPPs are

meeting their intended outcomes (Coggshall et al., 2012; Cochran-Smith et al., 2018).

Accountability Mechanisms and Theories of Action

Cochran-Smith and Reagan (2022) explain "… it is critical to understand how

power is taken up in the practice of evaluation" (p.449). Considering the Theories of Action

(ToA) within different accountability mechanisms tasked with TPP evaluation may

contribute to this understanding. A ToA is a chain of logic based on specific criteria

believed to lead to improvement. Unpacking an organization's ToA provides increased

clarity about why specific criteria are included within the evaluation, while other criteria

are excluded or minimized (McDiarmid, 2019; Wylie, 2017). Put another way, making

visible a mechanism's ToA highlights what is present and what is missing.

ToAs among local TPP stakeholders and policymakers are at the intersection of

passionate disagreement (Shepard et al., 2009; Walsh, 2006). One common ToA within

evaluation organizations is rooted in a belief that posits the education field has not held

itself accountable and cannot be trusted to make decisions that will benefit students

(Sykes et al., 2009). Justification for this belief is based in the extremely high number of

TPPs that receive accreditation, when compared to stagnant and low annual state

summative assessments for K-12 students and the gap in achievement between poor students of color and other students (Cuban, 2004). Policymakers holding this belief often promote a ToA that relies on mandates and penalties used punitively for the purpose of motivating local stakeholders (McDiarmid, 2019).

In a second ToA, a market-driven component is at play which asserts publishing performance data for TPPs will drive low-performing programs out of the market (Ellis & Smith, 2017; Loeb et al., 2011; Sykes et al., 2009). Assumptions underlying both ToAs are "threat of exposure" and "sanctions." Evaluation entities assert these threats will motivate preparation programs to improve outcomes (McDiarmid, 2019, p. 120).

Entities and initiatives such as the Higher Education Act (HEA), the Council for the Accreditation of Educator Preparation (CAEP), the National Council on Teacher Quality's (NCTQ), and teacher performance assessment organizations such as the edTPA all include specific evaluation measures for TPPs. While the measures vary between each organization and reflect different ToAs, there is one point where they are similar; each agrees public accountability reports will lead to improvement within teacher preparation (Carnoy & Loeb, 2002; Schiller et al., 2020; Wylie, 2017). Furthermore, the HEA, CAEP, and NCTQ initiatives represent a strongly held belief that programs should use annual student summative assessment data to make decisions about program improvement (McDiarmid 2019).

There is a third ToA, which challenges the previous two. This ToA asserts the integrity of members within the TPP community motivates them to gather evidence and engage in ongoing, collaborative efforts to improve programs without interference from

outside entities (Ellis & Smith, 2017; McDiarmid, 2019; Sykes et al., 2009). Examples of data collected for this type of evaluation are classroom artifacts, performance assessment data, case studies, teacher evaluations, and lesson plans (Cochran-Smith & The Boston College Evidence Team, 2009; McDiarmid & Caprino, 2017;).

The assumption underlying this third ToA can be stated as: The level of commitment within TPPs and among community partners is so great, it is all the motivation needed to improve a program (Peck & McDonald, 2013). Of course, the problem with this theory is prior to the last few decades, before mandates by state and national stakeholders, TPPs have not openly shared evidence of internal accountability methods or the standards by which the evidence was gathered (Greenberg & Walsh, 2010; Walsh, 2006). Understanding assumptions and motivations connected to an organization or mechanism's ToA can explain decisions about evaluation criteria and methods (McDiarmid, 2019).

Teacher Preparation Program Evaluation and Classroom Assessment Proficiency

Before alternative methods for evaluation can be considered, the type of evidence programs that are feasibly able to gather should be considered. Likewise, the areas within TPPs which hold the most potential to impact teacher readiness should be of concern. One area of consideration is the classroom assessment proficiency of teacher candidates (Graham, 2005; Guskey, 2003). This section will define and explore classroom assessment as an area for TPP evaluation.

Classroom Assessment Defined

Classroom assessment is defined as the ongoing informal and formal assessment of students within a classroom and is characterized by a cycle of communication and feedback between the student and teacher. This two-way feedback cycle provides both the teacher and student information to inform future instruction and learning and allows for deeper learning (DeLuca & Bellara, 2013; DeLuca et al., 2013).

Classroom assessments should be present throughout the lessons presented to students and is directly related to impactful teaching as it demonstrates to teachers what students currently know and can do. Classroom assessment also allows teachers to know where further instruction, practice, and support is needed (Hattie, 2009). Furthermore, research has established ongoing formative assessment of students is the most important occurrence within a classroom as it relates to learning and ongoing student progress (Wininger, 2005). Ahsan (2018) explains teachers should

> …adjust teaching and learning contemporaneously. Through this process, teachers gauge student learning and reasons students may not be making progress. The information is then used to give students descriptive feedback to facilitate student learning and/or to change instruction. (p. 23)

In other words, to the degree teachers are able to incorporate purposeful and ongoing informal and formal assessment in the classroom and as a result modify instruction, student learning increases. Unfortunately, the classroom assessment practices of new and more experienced teachers often reflect a focus on measuring student summative achievement (Campbell et al., 2002). While summative achievement has a place within classroom assessment, it is an "Assessment of Learning," rather than an "Assessment for Learning" or "Assessment as Learning"; the latter two have a much greater impact on

student progress (DeLuca, & Klinger, 2010; Guskey, 2003; Hattie et al., 2005).

Because established and ongoing research asserts classroom assessment is essential for student learning and because the research shows there is a gap in teacher readiness in this domain, it is an important area for program evaluation (Garrison & Ehringhaus, 2011; Guskey, 2003; Popham, 2011).

Teacher Candidates and Classroom Assessment

Previous research establishes that preparing impactful educators is directly linked to the quality of a TPP and that classroom assessment knowledge of teacher candidates is rarely considered by programs (Brown, 2004; Volante & Fazio, 2007). Of the four evaluation initiatives mentioned previously, three (HEA, CAEP, and NCTQ) do not include a direct connection to classroom assessment practices (Bastian, Henry, et al., 2015). Performance assessments such as the edTPA include an emphasis on ongoing classroom assessment; however, the inclusion of this criteria within the edTPA is likely not enough to evaluate a teacher candidate or a program in the area of classroom assessment, as a low score in this area of the performance assessment section can be overcome by strengths in other areas (Stewart et al., 2015). Furthermore, as previously mentioned, the difficulty of achieving "buy-in" for teacher performance assessments on the program level is difficult, which may limit or stall the teacher performance assessment's impact on program improvements (Bastian, Peterson, et al., 2015).

Volante and Fazio (2007) argue the perpetuation of incomplete and often ineffective classroom assessment practices begins within TPPs. Greenberg and Walsh (2012) found that less than 3% of 180 TPPs provided adequate preparation regarding

classroom assessment. The lack of preparation and practice in the area of classroom assessment is often due to the limited focus in curriculum included within required college courses and the classroom assessment experiences within teacher candidates' student teaching experience. It is during this phase that candidates are mentored by a cooperating teacher or placement school mentor who often demonstrates classroom assessment practices that are inconsistent with best practices within current research (Brownson, 2018; Popham, 2009; Rickenbrode et al., 2018). Unfortunately, even when teacher candidates are exposed to research-based teaching and assessment models within pedagogy courses, they often align their teaching with the placement school, a process that potentially continues the cycle of ineffectual assessment practices (Graham, 2005). Thus, a TPP that prepares teachers in current assessment practices and pedagogy and reinforces the same practices during the student teaching phase, has the potential to strategically improve student learning (Ahsan, 2018; Brookhart, 2001; Brown, 2004; Campbell & Evans, 2010).

**Gathering Evidence of Classroom
Assessment Proficiency**

Gathering evidence of outcome expectations and teacher candidates' perceived proficiency in the area of classroom assessment has the potential to reveal critical program strengths and weaknesses that impact the quality of education on the local level (Campbell et al., 2002; Popham, 2011). However, gathering such evidence on the program level can be problematic and is one of the major arguments for more removed evaluation, such as the use of annual summative assessment data (Cochran-Smith et al.,

2016, 2018).

Program level data collection related to classroom assessment is frequently very limited, and if data is included within evaluation reports, it often lacks focus and depth resulting in few measurable changes. Furthermore, when data is collected at the program level, it is rarely recognized in national reports, which primarily focus on annual summative assessment data (Cochran-Smith et al., 2016; McDirmind, 2019).

In sum, the methods used by TPPs to gather data for program evaluation have not been recognized by national stakeholders. However, there is an evaluation method used within public health, social science, and nursing that includes a local stakeholder focus and community contextual factors. Integrated concept mapping is a mixed method approach for program planning and evaluation that may yield important evaluation and research opportunities within the field of education. However, few TPP studies have used integrated concept mapping as means for program planning or evaluation.

Concept Mapping as a Program Evaluation Method

While the term concept mapping is commonly used to represent ideas in the form of symbols and maps, Trochim and Linton (1986) presented an integrated approach to concept mapping as a specific mixed methods technique. This approach includes brainstorming, statement analysis, generation of maps, and data displayed in charts and graphs (Kane & Trochim, 2007). This integrated approach to concept mapping is used in many research areas due to its ability to represent and visualize systematic processes. This method transforms data from a variety of stakeholders within an organization and

presents it visually (Kane &Trochim, 2007).

While a more detailed explanation of the integrated concept mapping process is presented in Chapter III, briefly stated, concept mapping (Kane & Trochim, 2007) can be used to evaluate a TPP in the area of classroom assessment through the following process:

1. Preparing for Concept Mapping
2. Generating Ideas
3. Structuring the Statements (grouping statements and rating each)
4. Concept Mapping Analysis (multidimensional scaling; hierarchical cluster analysis)
5. Interpreting the Maps
6. Utilization

While some or all of these steps may be included in the more general idea of concept mapping, a specific methodology is included in Kane and Trochim 's (2007) integrated approach. Integrated concept mapping can be completed through the use of proprietary software, through SPSS, or through another similar program. Using one of these programs, a similarity matrix, multidimensional scaling of the similarity matrix (MDS), and hierarchical cluster analysis is used to transform and map the data provided by participants.

Rationale for the Study

Current evaluation systems provide limited insights into the quality of TPPs. Performance assessments for teacher candidates, such as the edTPA, provide specific data about a teacher-candidates readiness for the classroom and by extension the program's impact on its graduates. However, research suggests performance assessments

do not provide timely and direct responses to the strengths and weaknesses of preparation programs, consider community factors or equity to the degree needed, and they forward limited "accountability era" conceptions of teacher effectiveness (Brady et al., 2018).

The evaluation processes currently in place remain too far removed from TPPs to impact programs meaningfully. Furthermore, they lack local community considerations and factors related to equity. Due to these factors, the influence of current TPP evaluation methods on program improvement and, by extension, improvements within schools remains minimal. Reconsidering current program evaluation mechanisms and their impact on the local service area of TPPs allows for thoughtful consideration of what is driving current methods with the potential to shift program evaluation toward more impactful improvements. Moving beyond narrowly focused "accountability era" criteria and toward an era of democratic accountability opens the door to evaluation which may be more collaborative, equitable, and inclusive of a wide variety of stakeholders. Each of these factors may impact TPPs and schools in way that improve K-12 education (Cochran et al., 2018).

Significance of the Study

This study is based on several hypotheses which have the potential to explain how TPPs can use actionable data for evaluation and improvement. First, this study is based on the belief that a teacher candidate's ability to create and administer classroom assessments, both equitably and at a high level, positively impacts student learning (Chizhik et al., 2018). Second, this study presupposes there is a relationship between how

TPP train students in the area of classroom assessment and a teacher candidate's ability to meaningfully assess students (Slade et al., 2019; Volante & Fazio, 2007). The third hypothesis assumes using integrated concept mapping to evaluate a TPP in the area of classroom assessment outcomes may be more useful and impactful than current evaluation methods. Finally, this study hypothesizes that using integrated concept mapping for TPP evaluation could contribute to the disruption of current accountability methods which still have ties to the narrowly focused "accountability era." This study will add to the body of knowledge about democratic evaluation methods for TPPs that are both meaningful and actionable.

The gaps within the research on TPP evaluation and classroom assessment identified in this study address the following issues: (a) the need to identify and provide evidence of intended outcomes of TPPs (Worrell et al., 2014), (b) current national evaluation requirements which call for TPPs to demonstrate valid and reliable outcome data (McDiarmid, 2019, p. 118), and (c) the lack of meaningful program level data collected as part of the program evaluation (Cochran-Smith et al., 2016).

Purpose of the Study

The purpose of this study is to identify and establish outcomes in the area of classroom assessment for a TPP in the Western region of the United States. This study aims to address the problem of meaningful program evaluation and is directly tied to the program's goals and desired outcomes for teacher candidates (Rickenbrode et al., 2018). This study employs a democratic approach to accountability, which may yield more

useful and actionable data than current methods of program evaluation. Through integrated concept mapping, it may be possible to uncover the program's intended outcomes in the area of classroom assessment and evaluate the program in this area. Evaluation data and the process details, including how the process unfolded within the TPP, will be provided to stakeholders (Kane & Trochin, 2007).

Research Questions

The following research questions assume there is a need to evaluate educator preparation programs, ensure program quality, and understand the impact of programs on participants in the area of classroom assessment. These questions are also based on the assumption that a democratic approach to accountability will yield more useful and actionable data than current, more removed methods of program evaluation (DeLuca & Volante, 2016; Stiggins, 1999).

1. How can integrated concept mapping contribute to the evaluation of a teacher preparation program in the area of classroom assessment outcomes?

2. How can integrated concept mapping shift the focus of TPP evaluation toward a democratic approach to accountability?

The short-term goal of this study is to identify the proficiency level of the program's teacher candidates in the area of classroom assessment. The long-term goal is to explore the impact of integrated concept mapping on TPP evaluation. The data and evaluation model will be shared in joint meetings and trainings within the TPP and possibly with local stakeholders.

Definitions of Terms

The following definitions are intended to orient readers to the ideas presented within this dissertation:

Classroom Assessments—Classroom Assessments are the large variety of activities that happen within classrooms and may include verbal questioning, portfolios, quizzes, performance demonstrations, and more. Data collected through classroom assessments inform ongoing instruction (Popham, 2009).

Democratic Accountability—A theory of accountability within education that prioritizes expanding democratic education into the evaluation methods for TPPs. This theoretical framework views teaching and teacher education as enterprises of public good, challenges market-oriented conceptions of evaluation, and recognizes inequities within schools and the processes by which they are reproduced (Cochran-Smith et al., 2018)

"Accountability Era"—An era that began in the 1980s and lasted through the 1990s and well into the 2000s. This era is characterized by a strong influence on state standardized assessment data and aimed to establish the teacher as the most important determining factor in student learning gains while mainly ignoring factors related to community and equity. It is also characterized by a market-driven approach to education (McDiarmid, 2019).

Integrated Concept Mapping—A mixed-method approach to program evaluation and organizational planning that transforms data gathered from various stakeholders within an organization into a series of maps displayed as charts and graphs (Kane &

Trochim, 2007).

Internal Stakeholder—An individual who has an interest in an organization and the outcomes of the organization's actions and who also has a specific connection to the organization (Kane & Trochim, 2007). This study's internal stakeholders include the administration, faculty, and students within a TPP.

Teacher Preparation Program (TPP)—A program, usually within a college or university, where individuals are trained to become professional educators through academic studies and practice within elementary or secondary schools.

Theory of Action (ToA)—A chain of logic based on specific predetermined criteria that are believed to lead to improvement in an organization or entity, if employed with fidelity (McDiarmid, 2019).

CHAPTER II

REVIEW OF THE LITERATURE

Within this chapter I will review current policies regarding TPP evaluation and present integrated concept mapping as an alternative to more removed evaluation methods. I begin this chapter by detailing the theoretical framework for the study, Democratic Accountability, as explained in *Reclaiming Accountability in Teacher Education* by Cochran-Smith et al. (2018). I chose to start at this point because evaluation within education has been tied to "accountability era" data and standards for over two decades. This approach minimizes factors related to equity and promotes a limited view of teacher and school "effectiveness" (Cochran-Smith et al., 2018). Democratic Accountability directly challenges current evaluation norms associated with the "accountability era," including annual summative assessment data as the standard for "effectiveness" within the field of education. Following the theoretical framework, relevant research findings are presented in the areas of TPP evaluation, preparation of teachers in the area of classroom assessment, classroom assessment design, and integrated concept mapping as a process for program evaluation.

Theoretical Framework

Historically, democratic education has been promoted by educational researchers and philosophers as a standpoint from which to view the purpose and nature of education. Dewey (1916) explained that democracy is "… more than a form of government; it is primarily a mode of associated living, of conjoint communicated experience" (p. 87).

Dewey prioritized diverse members of groups working together as an expression of democratic education. He believed schools should be microcosms for a democratic society; students, teachers, and others should interact with each other to make improvements and provide opportunities for student growth (Kira, 2019). Included within the theory of democratic education is also a democratic approach to school governance and oversight. Kira explains, "Democratic school governance takes the form of direct democracy…as well as of representative democracy… in which representatives of students and staff make decisions" (p. 61).

More recently, Dewey's ideas about democratic education have been linked to efforts to address social inequalities within schools (Apple & Beane, 2007 Knoester, 2012). For example, Barber (2009) asserts that "thin democracy" is more concerned with individual interests while "strong democracy" is most concerned with a self-governing community that prioritizes "civic education" and "participatory institutions" (p. 4). Likewise, Engle (2000) explains that the current education paradigm does not reflect a democratic approach to education and can only be disrupted by challenging market-based ideologies and forwarding democratic education as an alternative. Furthermore, one reason democratic education has not been fully realized within the U.S. is that standards-based reforms, have placed a greater "…emphasis on student achievement on high-stakes standardized tests…" (Kira, 2019, p. 61). Such policies are identified as obstacles to practicing the principles of democratic schooling (Meens & Howe, 2015; Meier, 2003).

Cochran-Smith et al. (2018) explains the concept of democratic education includes notions such as (1) public education as a means of forwarding democratic

principles of equality, (2) teaching students that individual participation in societal affairs is a fundamental aspect of a democratic society, (3) it is the responsibility of community members to reject power relationships that threaten a democratic approach to education, and (4) individuals do not intuitively understand how to live and participate in a strong democratic society; therefore, education should embody, teach, and forward these principles if education is to fulfill its primary function as defined by John Dewey and others who viewed education through the lens of democracy. Rather than viewing education as a means of competition or vying with other countries in a ranked system, a democratic approach to education asserts that the main purpose of education is to prepare students to be part of the larger democratic project where they are expected to be active participants.

In contrast, the current accountability paradigm is based on ideas unrelated to democratic education and democratic accountability (Kira, 2019). Instead, the accountability paradigm promotes "uniformity and compliance" while also limiting action in the areas of equity and forwarding democratic ideals (Cochran et al., 2018, p. 154). It follows that accountability in education, in light of these democratic education principles, is characterized by the following:

> First, democratic accountability is based on the assumption that to survive, 21st-century democratic societies need deliberative and democratic education that teaches all students how to analyze multiple perspectives and engage in deliberative dialogue…. Second, democratic accountability is founded on the assumption that in democratic societies, teaching and teacher education are enterprises of the public good, rather than market-oriented enterprises…part of the job is recognizing inequities in schools and society and working with others to challenge the structures that reproduce inequities. Third, democratic accountability in teacher education is based on dialogue and participation of all stakeholders. (Cochran-Smith et al., 2018, p. 154)

Connected to these principles, the primary role of teacher education is to "prepare teachers who enact deliberative and critical democratic education" (Cochran-Smith et al., 2018, p. 172). Accountability measures for TPPs should aim to situate the purpose and goals of evaluation within the larger concept of democratic education, as defined by Dewey and others, while also challenging the dominant narrative forwarded in the "accountability era," a narrative that states there is a problem within the education system that can only be remedied through structured oversight by accountability experts who rarely understand or include the local contextual factors or meaningfully consider local stakeholder input (Cochran-Smith et al., 2018).

In sum, the concept of accountability within this study will be viewed through the lens of democratic accountability as detailed throughout this section and set forth by Cochran-Smith et al. (2018).

Evaluation of Teacher Preparation Programs

This section, I will present research detailing the history and context of TPP accountability and evaluation mechanisms. I will also describe current evaluation policies, motivations for evaluation processes, and what is still missing within evaluation practices.

Similar to K-12 public school accountability measures, TPPs have been accountable to states since the days of Normal Schools (Fraser, 2007). However, accountability on the national level did not begin until the Elementary and Secondary Education Act of 1965 (ESEA), which provided funds for students in need and offered

grants to school districts who served students living in poverty (Cuban, 2004). The ESEA also provided funds for textbooks, library books, special education, and school quality improvements. This initiative instituted annual evaluations for schools receiving Title I funds, demonstrating a new level of public accountability (Cuban, 2004).

Almost 20 years after ESEA, The Nation at Risk Report ushered in what is known as the "modern era of accountability" (Cochran-Smith et al., 2018). The Nation at Risk: An Imperative for Education Reform was published in 1983. It was commissioned by the Secretary of Education under President Ronald Reagan (Kamenetz, 2018). The report warned of mediocre education and highlighted calls for reform and accountability. The report was published in the shadow of the Cold War amid fears that the US was falling behind other countries and would not be able to compete globally. This fear was laid at the feet of the U.S. education system and spurred numerous state and federal reform measures (Strauss, 2021). Today, much of the evidence presented in the report has been discredited, and the report is viewed as a highly politicized investigation of public education containing many factual errors (Kamenetz, 2018). Despite the backlash over time, the document's influence remained paramount and still impacts education reform and accountability measures through the precedent of extreme criticism set by the document and its call for greater accountability (Strauss, 2021).

Since *A Nation at Risk's* publication, government agencies and policymakers have been steadily focused on the ability of schools, teachers, states, and TPPs to hold themselves accountable (Feuer et al., 2013). Private entities entered the realm of educational oversight during this "accountability era," publishing public rebukes on the

quality of education in the U.S. and the lack of accountability (Fleener & Exner, 2011).

For example, The Carnegie Foundation for the Advancement of Teaching (CFAT) issued

a public statement calling for action after a report (generated by the same corporation)

reported that over fifty percent of private corporations had to provide employees

instruction in math and reading (Eurich & Wade, 1986). Likewise, the publication of

What Matters Most: Teaching for America's Future by the National Commission on

Teaching and America's Future focused on weaknesses within TPPs. The publication

focused on the "irrelevance" of teacher licensure as a marker of readiness to teach in light

of K-12 student outcomes (Feuer et al., 2013).

Legislation aimed at education accountability, including the reauthorization of the

ESEA Act in 1994 and the HEA in 1998, is still in force today (Walsh, 2006). The

reauthorization of the ESEA included a strong focus on standards-based accountability of

public schools (Baker et al., 2010). Similarly, under the reauthorization of the HEA in

1998, accountability measures for TPPs were added to Title II of the Act, requiring states

and preparation programs to report the following each year.

TPP Requirements

- Graduates' scores on licensure examinations
- Enrollment numbers within preparation programs
- The length of required school experience before licensing
- Program approval requirements as outlined by the state
- The program's status with the state indicating if the TPP is considered low performing.

State Level Requirements

- Indicators developed by the state for identifying low-performing TPPs
- Public reporting of the status of TPPs within the state
- State licensing requirements

- State alignment between K-12 content standards and licensure requirement
- Licensure examination results for the candidates within preparation programs
- Licensure waivers approved by the state for teachers.
- Descriptions of alternate routes to licensure and test pass rates for the participants
- Details of required subject-matter tests (Earley, 2001)

Tracing the history and the process of evaluation mechanisms forwarded throughout the "accountability era" is critical to understanding current policies regarding TPP evaluation. The formalizing of evaluation methods within this era, including linking evaluation to annual summative assessment data, established summative yearly assessments as the critical measure of teacher, school, and TPP effectiveness. Following education reform measures of the 1990s and facing continued pressure from the public and private entities, states began to work on their public accountability measures (Braun, 2005). Independent of each other, several states sought to develop and align standards and summative assessments and to link accountability measures to this alignment (Carnoy & Loeb, 2002). State-level initiatives used these accountability methods to identify struggling schools. If a school was identified as struggling, additional support and resources for the school were provided, a practice which was often reported as an effective practice on the state level.

Following these efforts forwarded by the states, the No Child Left Behind (NCLB) movement gained momentum, and federal legislation was passed in 2001. The NCLB act increased accountability measures for teachers, schools, and TPPs. Unlike the state-level models, NCLB included sanctions for schools that did not achieve "Adequate Yearly Progress" (AYP; McDiarmid, 2019). While the NCLB legislation was seen as a national initiative modeled after state programs that identified lower-performing schools,

the advent of sanctions was a direct departure from state models.

Annual State Summative Assessments and Teacher Preparation
Program Evaluation

After education accountability was more formally transferred from the state to the national level through NCLB and earlier legislation, K-12 student summative assessment data played a central role in evaluation and accountability. At the same time, local contextual factors and considerations were minimized within evaluations (McDiarmid, 2019).

Each of the accountability mechanisms (e.g., HEA, NCLB, Every Student Succeeds Act [ESSA], Race to the Top [RTTP]) and CAEP, organized over the past 30 years, call for the collection and publication of K-12 student outcome data from annual summative assessments as a significant aspect of TPP accountability (Bastian, Patterson, et al., 2015). This data is considered the most significant measure of program "effectiveness." The term "effectiveness," as it relates to education, has become synonymous with K-12 summative assessment gains and evaluation. Evaluation entities have portrayed yearly student gains on such assessments as an objective means for evaluation (Darling-Hammond, 2000). However, the terms "effective" and "effectiveness" are most often considered in relation to summative assessment data and exclude other critical aspects of evaluation such as community and school contextual factors and factors related to equity (Cochran-Smith et al., 2018).

Today, TPP evaluation is often conducted by outside entities who link student

summative assessment data to teachers and then to the TPPs where they graduated. This process is completed using a model developed at the University of Tennessee (McDiarmid, 2019). The Tennessee Value-Added Assessment System (TVAAS) was created to measure "…student academic growth on state standardized assessments [and] to provide evidence of teachers' impact on student learning over time "(State Collaborative on Reforming Education [SCORE], 2017, p. 2). The model "isolat[ed] a teacher's influence on student achievement from other variables that could affect student performance" (Sanders & Horn, 1994, p. 304). The model is based on earlier research which claims teacher-effectiveness is the most important determining factor of student success; however, this claim is solely based on value-added, annual, or semi-annual summative assessment data (Sanders & Rivers, 1996). Interestingly, even researchers who support the use of the TVAAS model, explain the data should only be used as a marker of "effectiveness… when used along with other evaluation measures" (SCORE, 2017, p. 2).

With the TVAAS model accepted as the paramount example for isolating a teacher's influence, and with the critical element of better technology emerging in the early 2000s, it became possible and popular to efficiently link student data to individual teachers (Brady et al., 2018). In a relatively short time, data collected through value-added summative assessments became the most valued measure of student achievement and the means most often used for measuring a teacher's influence (Darling-Hammond, Amrien-Beardsley et al., 2011). The next step was to link "teacher effectiveness" to the TPP where teachers received their training. Once summative assessment data was

isolated from other factors, linked to teachers, and then to TPPs, data-based reform measures and accountability grew to its present state (Brady et al., 2018; Noell & Burns, 2006).

By pioneering this evaluation model and directly linking K-12 student growth to individual teachers and to programs, the TVAAS model paved the way for other technology companies to follow a similar model for evaluation (McDiarmid, 2019). Today, more than 22 states, contracting with a number of education technology companies, provide similar teacher performance data. However, data is rarely available or is difficult to disaggregate for more detailed analysis. Instead, more removed data analysis reports are supplied to TPPs from evaluation entities (Data Quality Campaign, 2014). In one study of the Deans for Impact organization, which included TPPs for more than 17 states, it was reported that only six programs had access to the information connecting teachers from their program to state summative assessment data (Ciccone, 2019).

Beyond simply having access to the data, most TPPs lack the resources needed to use these data in meaningful ways. Because program improvement is the major impetus for national accountability measures and legislation, the lack of direct access to the data, data collection methods, and contextual details regarding the data collected, evaluation aims cannot be addressed by TPPs in meaningful ways that may more directly impact K-12 students (Peck & McDonald, 2014). Additional data such as teacher placements within school districts, Praxis tests for content and pedagogy knowledge, and/or teacher performance assessment scores, and teacher retention rates are also collected as part of

accountability measures. However, this data is secondary to annual student summative data and is not considered in connection to summative data (Cochran-Smith et al., 2018).

Before moving on, it is important to note what has not been included in this review. I have not included the financial cost of accountability measures, those who may benefit financially from accountability measures, and the associated consequences of both. While these considerations may yield important understandings and implications, they are beyond the scope of this study.

Following this historical summary of education evaluation, legislation, and current models of evaluation I will now turns to the Theories of Change and Theories of Action inherent within the evaluation methods previously detailed and which are forwarded by the agencies and organizations involved with TPP evaluation; specifically, those who report, publicize, and publicly comment with authority on evaluation findings.

Evaluation Models: Theories of Change and Theories of Action

This review now turns to evaluation models and their associated Theories of Change (ToC) and Theories of Action (ToA), both of which are important to uncovering an understanding of the motivations and complexities of educator accountability mechanisms. A critical investigation of an organization's logic model and subsequent ToA provides this insight (Cochran-Smith et al., 2018; McDiarmid, 2019).

A logic model is a "[g]raphical depiction of logical relationships" which are demonstrated "between resources, activities, outputs, and outcomes of a program" and are used to ascertain program outcomes and assess causal relationships (Becker, n.d., p.

4). Logic models are intended to shed light on gaps in logic, clarify assumptions, build understanding, and focus attention on what needs to be evaluated. Importantly, each model is a summary of a complex program and provides a means for understanding the process and motivations (Mathison & Rogers, 2007). However, there are also critical limitations to logic models when considering their relationship to TPP evaluation. First, a logic model is not a complete plan "…for designing and/or managing a program or policy" (Becker, n.d., p. 6). Second, it is not an evaluation of the plan itself. Third, logic models should include external factors. Finally, if the individuals creating the model are too far removed, do not fully understand which contextual factors should be considered, and/or ignore factors, the model itself is faulty (Becker, n.d.). These limitations affect the evaluation process and the implications for organizations using the data generated from the evaluation. When a logic model is organized for the purpose of evaluating a program, the program's stakeholders should be included in the creation of the logic model, due to the factors previously mentioned and because it "enhances stakeholder buy-in to the model…[which] may be key to their motivation to undertake activities" (Chen et al., 2018, p. 62).

A ToA results from causal process(es) stated as *If…Then… Because…*within the logic model and connects activities to outputs. It should also be based on assumptions supported by research. When the assumptions articulated in the model are directly linked to the *If…Then…Because…* statements, an intervention or treatment can be instituted and measured. The consequences of the intervention or treatment are then connected to program outcomes and considered intentional (Bastian, Fortner, et al., 2015). The

"effectiveness" of a program depends on the "truthfulness" of the assumptions on which

the treatment or intervention activities are based. As Becker (n.d) suggested, "[i]f invalid

assumptions dictate the strategies of a program, it is unlikely to succeed" (p. 5).

In the case of TPP evaluation and current complex systems of accountability, the

evaluation mechanisms in place are independent and disconnected (Finnigan & Gross,

2007). Each entity aims to improve education or education outcomes through separate

and distinct logic models with specific requirements and different ToAs (McDirmind,

2019). A ToA outlines distinct parts of an organization or system. It also identifies

relationships within and among the parts and is based on a goal as well as the processes

for how the intended goal is achieved (Mathison & Rogers, 2007). While a mechanism's

logic model is rarely published or easily accessible, the underlying assumptions, causal

processes, and ToAs can be uncovered by identifying the resulting consequences and

then working backward through a restructuring of the logic model (Cochran-Smith et al.,

2018).

An exercise in unpacking the current mechanisms of evaluation for TPPs and their

ToA may lead to the following insights.

1. Title II of the HEA requires TPPs (and states) to track specific data and report annually to the federal government. The ToA related to the entity's logic model holds that if negative information about the TPP is published, fewer students will enroll in the program and instead enroll in another TPP. This negative report will affect the program's public reputation. Programs are fined if information is not reported accurately and in a timely manner. The threat of financial penalties is also intended to motivate programs to provide this information.

2. CAEP requires TPPs to submit reports annually and pass accreditation every seven years to maintain accreditation status. Accreditation also affects Title II accountability measures (Finnigan & Gross, 2007). The theory of ToA related to this mechanism holds that if programs do not pass accreditation measures,

public reports about probationary status or the loss of accreditation will affect the program's standing with the state, enrollment, and public confidence. Therefore, this threat will motivate the program to meet accountability demands.

3. Performance assessments such as the edTPA, required in about half of US states, are often connected to licensure requirements for teachers and affect program status on the state and national level due to Title II and accreditation requirements (Bastian, Henry et al., 2015). The ToA is as follows: If teacher candidates are not prepared within the program they enroll in, they will not pass the assessment at a high rate. The student pass rate is publicly reported and used within accreditation measures, state licensing requirements, and HEA requirements. Again, loss of public confidence and enrollment in the program is meant to motivate programs to better prepare teacher candidates for the assessment.

The information about TPPs published in HEA, CAEP, and edTPA reports is used by many public and private organizations that republish the data along with secondary analyses and further commentary (Fleener & Exner, 2011). Of course, the more removed an entity is from the data gathered, the less likely they are to understand the process and particulars of the data, its relevance, and the contextual features of the community (Finnigan & Gross 2007).

As mentioned previously, policymakers hold beliefs about education rooted in the assumption there is a historical lack of accountability within teacher preparation (McDiarmid, 2019). Entities representing this belief agree teacher preparation, and by extension teacher quality, can be improved through mandates, penalties, and publication of specific evaluation data which will affect the program's success and force lower-performing TPPs to improve or exit the marketplace (Bastian, Patterson et al., 2015). In *Holding Teacher Preparation Accountable: A Review of Claims and Evidence*, Cochran-Smith et al. (2016) explains how these theories reveal critical policy considerations:

Initiatives reflect different accountability mechanisms and theories of change, and

they are governed by different institutions and agencies, including governmental offices, professional associations, and private advocacy organizations. Despite differences, each assumes that the key to teacher education reform is accountability in the form of public assessment, rating, and ranking of states, institutions, programs, and/or teacher candidates. This brief addresses two questions for each initiative: What claims do proponents of the initiative make about how it will improve teacher preparation and thus help solve the teacher quality problem in the U.S.? What evidence supports these claims? The first question gets at the theory of change behind the initiative and its proponents' assumptions about how particular mechanisms actually operate to create change. The second involves the validity of the initiative as a policy instrument—that is, whether or not there is evidence that the initiative actually meets (or has the capacity to meet) its stated aims. (pp. 3-4)

Uncovering the ToA within an evaluation organization or policy allows for a more candid discussion about current and future policymaking and evaluation. Specifically, it allows for consideration of missing, misdirected, and misunderstood assumptions within logic models which heavily influence and direct TPP accountability.

Classroom Assessment

Because classroom assessment is the focus of the program evaluation for this study, this review now moves to the topic of informal and formal classroom assessments.

Classroom assessment includes the large variety of activities that take place within classrooms such as verbal questioning, portfolios, quizzes, performance demonstrations, and more. Data collected through classroom assessment should be used to inform instruction and can be defined as the "… formal and informal procedures teachers employ in an effort to make accurate inferences about what their students know and can do" (Popham, 2009, p. 6). Further research defines additional features of classroom assessment as:

1. Having the potential to be formative, with the goal of improving learning
2. Being linked to ongoing classroom instruction
3. Mainly created by teachers, rather than standardized
4. Having a focus of feedback to the students and teacher
5. Typically informal, but sometimes they are more formally organized
6. Low stakes in nature
7. Usually not formally graded, but at times they may be more formally graded.
   (Garrison & Ehringhaus, 2011; Popham, 2009; Stiggins & Chappuis, 2005).

The term classroom assessment is often used interchangeably with formative assessment throughout the research, although the two are not the same (Bennett, 2011). For example, classroom assessment may be formative, meaning the goal of the assessment is to improve student learning, provide information about student learning to the teacher, and is not formally graded. However, the goal of a classroom assessment may also be to collect data about where students are at a given point in time in connection with learning targets or standards, while at the same time being connected with formal grades (Wininger, 2005).

While teachers have many reasons to employ assessment techniques in the classroom, it is the assessments intended to inform future instruction and/or provide feedback to students and teachers that research overwhelmingly supports in connection with student progress (Birenbaum et al, 2015). Research also suggests that most classroom assessments have the potential to be formative in nature, even if they were intended as summative and can impact student achievement when used by the teacher to inform future instruction (Black & Wiliam, 2011). Furthermore, most classroom assessments are often blended, they are both formative and summative by design (Banta, 2007). The practice of blending classroom assessments increases "the reliability, validity, and utility of …assessment data" and is supported by assessment experts (Ahsan, 2018, p.

24). Bennett (2011) explains that the purpose of classroom assessments, both summative and formative, is not as far apart as supposed and the two often work together to support student learning.

In a review of classroom assessment literature, which included a wide range of ages and countries, Black and Wiliam (2004) found specific features were associated with classroom assessment. Feedback, inquiry, and self-assessment practices were each identified as unparalleled tools for student learning; each resulted in significant learning gains, when compared with other classroom practices. Classroom assessment was also shown to close gaps in achievement between students within the same classroom (Stiggins & Chappuis, 2005). When compared to more removed state level summative assessment, classroom assessment allowed for direct and timely feedback for both students and teachers, which contributed to ongoing cycle of instruction and learning in a classroom (Garrison & Ehringhaus, 2011). Accordingly, Shepard, et al. (2018) found assessment practices that include feedback impact student learning and achievement at a higher rate.

Diagnostic Power of Classroom Assessment

When classroom assessment is formative or blended, teachers are able to use assessment is as a tool to understand what a student knows, is able to demonstrate, and areas where they need more practice and support. This diagnostic aspect of assessment is what makes classroom assessment a critical factor in the classroom. It is characterized by a cycle of two-way feedback, which moves from the student to the teacher through the assessment, and from the teacher to the student in the form of corrective and supportive

communication (William, 2011). Each component of classroom assessment should work together to strengthen student learning goals and allow for the critical element of feedback (Garrison & Ehringhaus, 2011).

Classroom assessment is often designed to identify specific proficiencies, which are captured through discussion, short writing assignments, observation, structured practice, and more traditional quizzes/tests (Ahsan, 2018). Following the analysis of assessments, teachers are able to use data to inform instruction quickly and respond to individual students, patterns, and trends within groups of students, and the class as a whole (Brookhart, 2001).

Because classroom assessment practices should be embedded within instructional practices in the classroom, it is critical teachers are knowledgeable in areas of assessment and its connection to instruction (Brookhart, 2001).

Assessment and Learning

There are three ways to define assessment: (1) assessment of learning, (2) assessment for learning, and (3) assessment as learning. Understanding the difference is important for this study as "assessment of learning" is characterized by stakeholder accountability data and other measures of learning that attempt to capture a summative view of all the learning that takes place within a course (Garrison & Ehringhaus, 2011). This type of assessment is often associated with final grades, passing a course to move on to another course or grade, diplomas, certificates, or other markers of summative achievement (Campbell et al., 2002). As mentioned previously, this type of assessment rarely helps students progress in their learning.

Classroom assessments intended to be "Assessment for Learning" support students in the classroom as they continue to learn. Furthermore, "Assessment as Learning," happens when an assessment is designed to be the learning tool itself (William, 2011). In other words, the process of assessment is the means for continued learning. As mentioned previously, there can be cross-over between formative and summative classroom assessments. Crossover assessments are still connected to student learning gains and provide a "rich learning process for the students" (Ahsan, 2018, p. 62). A few examples of cross-over assessments are self-assessments, peer assessments, games, projects, presentations, performances, and classroom dialogues.

In the area of classroom assessment, Hattie and Yates (2013) discussed the role of feedback in connection to student learning within the classroom. Hattie uses the term "maintenance of learning" to describe explicit strategies teachers should develop. Maintenance takes place when teachers help students evaluate their own learning and as students learn to respond to instruction and feedback more formally. Hattie's research on Visible Learning is based, in part, on the concept that as teachers evaluate their impact on student learning in the classroom, they become more knowledgeable teachers. Closely connected to this concept, is the role of feedback and its relationship to maintenance:

> We require high levels of maintenance in learning and thus the ability of teachers to diagnose where the student is relative to the criteria of success is critical. This is where notions such as assessment for learning…[and] student assessment capabilities are all invoked…. (Hattie & Yates, 2013. p. 32)

In sum, teachers should know where students are within their learning, so they may choose the right intervention at the right time. The concept of maintenance within this context of classroom assessment is connected to ongoing learning and the observation of

learning by the teacher. Knowledge retention requires maintenance because it takes multiple points of learning to move from superficial levels of learning to deeper levels (Hattie & Yates, 2013).

Given the research supporting classroom assessment in student gains and deepening learning, it is surprising classroom assessment has not gained as much attention as other areas of classroom instruction. Focusing research on classroom assessment may reveal effective cycles for instruction, patterns of student learning, development within the classroom assessment and feedback process, greater support for classroom assessment instruction within TPP, and ongoing professional development (Ahsan, 2018).

Frequency of Classroom Assessment

Classroom assessment is practiced at a lower-than-expected rate, given its effect on student learning (Black & Wiliam, 2009). Furthermore, many teachers who employ classroom assessment strategies are less effective than they could be, as some teachers misunderstand the feedback aspect of classroom assessment and focus more on issuing a grade in connection with the assessment (Campbell et al., 2002; Chen, 2005).

Classroom assessment is less effective when it is focused on lower-level learning such as memorization and recall, instead of thinking critically and applying concepts in different contexts (William, 2011). Black and Wiliam (2004) also found that even when teachers have the intent to teach higher-order skills, they often mistakenly end up measuring lower-order thinking skills related to factual recall and more basic demonstration of skills. Furthermore, teachers do not often "critically review"

assessments to ensure they are aligned with the curriculum as defined by the standards and objectives, at a high rate (Ahsan, 2018, p. 25). When classroom assessment is not used in conjunction with meaningful feedback and instead is simply connected to grades, it can negatively affect student motivation in the classroom (Birenbaum et al., 2015)

Black and Wiliam (2004) describe a "Poverty of Practice" among teachers related to classroom assessment and that stems from three sources. The first results from focusing on high-stakes assessments and annual summative assessments. A focus on high-stakes assessments over the last few decades has resulted in teachers feeling pressure to concentrate on specific aspects of the curriculum while ignoring or diminishing other aspects, commonly referred to as teaching to the test. This practice often reduces the practice of teachers providing "descriptive feedback" to students in the areas that are minimized, which impacts potential student learning (Ahsan, 2018, p. 30).

The second reason classroom assessment occurs less often than expected stems from the instruction received during educator preparation (Popham, 2011). When teachers leave a TPP and enter the classroom, they are often not prepared in classroom assessment, nor do they adequately understand the impact of classroom assessment practices (Wininger, 2005). Popham's research demonstrates that TPPs do not often train teachers to use classroom assessment and the professional development which occurs after teachers leave the program, does not often focus on classroom assessment in ways that impact student learning. One final reason teachers may not use classroom assessment effectively is the belief that student IQs or other markers of intelligence cannot be changed through instruction in the classroom. Therefore, classroom assessment is not

viewed by the teacher as having an important impact on student learning (Birenbaum et al. 2015; Black & Wiliam, 2004).

Teacher Preparation and Classroom Assessment

The literature concerning teacher preparation and classroom assessment is limited; however, the available research suggests several areas TPPs should emphasize within the program. First, teacher education programs should purposefully connect assessment with instruction. Second, programs should teach students to use the right assessment method at the right time, and third, TPPs should prepare teachers to develop quality assessments that are linked to scoring criteria. Programs should also teach future educators how to avoid assessment bias, effectively communicate students learning, and how to use assessments effectively as a classroom instructional strategy (Stiggins, 1999; Volante & Fazio, 2007). Each of these areas requires teachers to integrate assessment throughout instruction and to consider assessment theory within their instructional decisions (DeLuca & Bellara, 2013).

Suggestions for assessment curriculum within preparation programs have been made by researchers. Stiggins (1999) advocates for the inclusion of assessment methods across a variety of education classes where students are taught concepts multiple times in relation to different education courses, while also including independent courses in assessment. DeLuca and Klinger (2010) agree there should be a focus on classroom assessment within TPPs and a goal of training teachers to be assessment literate. Greenberg and Walsh (2012) suggest assessment literacy could take place within specific

courses or be included across many classes already in place. Furthermore, additional courses where students apply assessment methods and theory in a clinical setting under the supervision of a university professor and/or professional teacher who is an expert in assessment should be included within programs (Stiggins,1999). DeLuca and Klinger (2010) found assessment should also be part of ongoing professional training within clinical and student teaching experiences.

Throughout the literature, researchers assert current assessment instruction within TPPs is not adequate and does not provide enough explicit instruction to impact prospective teachers' knowledge and skills (Volante & Fazio, 2007). This is true, even if a TPP has an assessment course in place (MacLellan, 2004). Teacher candidates who complete at least one assessment course report higher confidence levels and skills when they begin teaching, but they still did not possess the skills or knowledge at a high enough level to "…engage in deep and complex learning about linkages between assessment, teaching, and learning" (Ahsan, 2018, p. 35). This finding suggests TPPs need to address classroom assessment, and that it may impact K-12 student learning (DeLuca & Klinger, 2010; Graham, 2005; MacLellan, 2004).

Stiggins (1999) maintains TPPs need to examine where classroom assessment is currently taught and modeled within the program and identify specific graduation and licensure requirements in the area of classroom assessment. Finally, Ahsan (2018) asserts there is enough evidence to demonstrate the need for TPPs to address classroom assessment instruction,

> Indeed, the duration of required assessment education courses is short, typically one semester (which is 3 hours), leaving little instructional time to provide teacher

candidates with a strong theoretical and practical foundation in assessment processes, assessment fairness, and measurement theory, let alone providing adequate coverage of more integrated and complex concepts of assessment for learning, communication of assessment information, and linkages between classroom environment and assessment. (p. 68)

Classroom Assessment and Curriculum within Teacher Preparation Programs

The content and curriculum taught within TPPs has been considered and debated from the beginning of teacher education institutions (Fraser, 2007). General education knowledge, content area knowledge, and general as well as specific pedagogy have become the teacher education curriculum standard. While it is generally agreed that TPPs should include courses in curriculum, assessment, teaching methods, classroom management, multicultural studies, technology, and pedagogy, pedagogical knowledge is viewed as critical and is generally part of all courses. Throughout the research available in the area of TPP and classroom assessment, it is suggested assessment instruction be linked to overall pedagogical knowledge throughout courses and as mentioned above, should also be taught in standalone courses (Darling-Hammond, 2000; Hawk & Schmidt 1989).

Black and Wiliam (2004) suggest two reasons national policymakers have ignored classroom assessment. First, annual summative assessments have diminished the focus of classroom assessment and shifted attention away from classroom assessment practices. Second, because policymakers are too far removed from local education needs, it is easy for classroom assessment to be overlooked within teacher preparation and program evaluation.

More recent research suggests it is possible to connect classroom assessment to student learning (Hattie & Yeats, 2013). Therefore, evaluation entities looking for valid and reliable measurements of student learning may simply be unaware of the issue or not fully understand the impact of classroom assessment on student learning. There is one exception, performance assessments such as the edTPA include specific references to informal and formal classroom assessment practices and score teacher candidates in the area of assessment (Stewart et al., 2015).

In many instances, the nature of TPP evaluation within the U.S. has situated accreditation entities and the evaluation data they require from TPPs too far away from the local service area to be useful in terms of program improvement (Cochran-Smith et al., 2014). In other words, the local needs of students and districts are not considered to the degree they should be to impact programs and K-12 education. Within program evaluation, classroom assessment practices have not garnered attention in the same way annual summative assessment data has captured the attention of those calling for accountability (Ahsan, 2018). If national evaluation does not emphasize classroom assessment as part of the evaluation criteria for teachers or TPPs, it is not an important focus on the state or program level regardless of the effect on student learning (DeLuca & Bellara, 2013). To this point, only about half of states include competency requirements or require assessment courses for teacher candidates in the area of classroom assessment (DeLuca et al., 2010).

Little research is available on "the content, effectiveness and nature of…teacher preparation program [coursework] as it relates to classroom assessment," (Ahsan, 2018,

p. 72). However, DeLuca and Bellara (2013) found classroom assessment instruction within TPPs varied greatly and lacked alignment. Even within courses designated as assessment courses, the curriculum lacked standardization across programs (DeLuca et al., 2010). A greater national focus on classroom assessment as it relates to teacher preparation may strengthen the local focus on classroom Assessment within programs and go a long way toward strengthening the classroom assessment curriculum within TPP for teacher candidates (DeLuca & Bellara, 2013).

Mertler and Campbell (2004) found that teachers report inadequate training within TPP in the area of classroom assessment. The low levels of assessment literacy among teachers are connected to teacher perceptions of not being prepared to assess their students (Maclellan ,2004). DeLuca and Klinger (2010) and Graham, (2005) explain that much of the instruction within TPPs in the area of classroom assessment is limited in scope and frequency, is mainly theory-driven, and lacks a local focus. The result is a lack of confidence when it comes to classroom assessment practices and an inability to connect course material to the classroom during clinical experiences and after graduation (Stiggins, 1999).

In order for teachers to become more skilled in classroom assessment and impact student learning to a greater degree in their classrooms, specific proficiencies need to be taught within TPPs and be supported through program evaluation requirements (Wenglinsky, 2002). A greater emphasis on classroom assessment on the state and national level will help forward this initiative on the local level, likewise, the inclusion of local classroom assessment data related to teacher candidates' classroom practices should

be an aspect of internal program evaluation for TPPs (DeLuca, Chavez, Bellara & Cao, 2013).

Research supports the connection between the inclusion of a specific assessment course and teacher assessment confidence levels in the area of classroom assessment (Chen, 2005; DeLuca & Klinger, 2010). Furthermore, Koh (2011) found additional assessment instruction, beyond one course, dramatically improved assessment literacy. DeLuca and Volante (2016) found that teacher effectiveness in many areas including classroom assessment is linked to national policies for teacher education. Stiggins (1999) asserts the National Council on Measurement in Education (NCME) of the U.S. who advocates for improvement in the quality of annual summative assessments, should also lobby for improvements in classroom assessment. When an organization does not focus attention and resources on classroom assessment, they are ignoring "the most important piece of assessment that happens in the classroom…" (Ahsan, 2018, p. 31).

Classroom Assessment and Student Teaching

The role of student teaching or a teaching practicum within educator preparation is a standard aspect of program design (Kamens, 2007). The role of student teaching is to allow students, at the end of their course of study, to practice what they have learned in an authentic environment. Research demonstrates when students teach under the supervision of an experienced teacher, their proficiencies and attitudes toward many classroom aspects mirror that of their cooperating teacher (Kamens, 2007). Because research demonstrates classroom assessment does not currently occur at the frequency needed to impact student achievement at a high degree, the chances that a student teacher

will be paired with a cooperating teacher who demonstrates the attitudes, skills, and knowledge needed to shape the student-teacher's perceptions and competencies in this area is low (DeLuca & Volante, 2016).

Continued learning within the TPP during student teaching may impact how teacher candidates continue to develop their teaching skills and knowledge while they are teaching in an authentic setting (Ball & McDiarmid, 1990). If a teacher candidate feels disconnected from the program or has a difficult student teaching experience, they may have little opportunity to practice classroom assessment in meaningful ways. In turn, this may hinder the transfer of classroom assessment skills to their own classrooms after graduation. Although some research suggests modifying the student teaching practicum to connect more directly with the TPP during student teaching would benefit program participants in the area of classroom assessment, few programs have modified their program to include a more direct connection (Clandinin & Connelly, 1998).

Concept Mapping and Program Evaluation

The final section of this literature review describes the method that will be used to evaluate a TPP in the area of classroom assessment for this study. I decided to introduce prior to Chapter IV due to its novelty within the field of education.

Integrated concept mapping offers an alternative to current evaluation methods and is used within the fields of social science, nursing, and counseling for planning and evaluation. It has been used minimally within the field of education and is not currently used as a standard evaluation method within TPPs.

Concept mapping may be a good option for TPP evaluation for a few reasons. First, it is already used successfully within other fields. Second, it is a mixed method that is both user-friendly, manageable for novice evaluators, and transforms qualitative data into quantitative data that is relatively easy for the organization to understand and use for future planning and reporting (Kane &Trochim, 2007). Third, it captures contextual factors of organizations through stakeholders who take an active part in the process, it is essentially a democratic process for evaluation. Reviewing literature related to concept mapping and integrated concept mapping within this section is intended to provide background and context for the study method presented in Chapter III.

Concept Maps

Concept maps can be defined simply as conceptual diagrams depicting relationships between concepts (Kane &Trochim, 2007). They were originally used in qualitative research and were not always formally structured. They are often used to demonstrate research participants' understanding of a topic or issue (Wheeldon, 2010). The general structure of concept maps has remained intact over time and the process is sometimes characterized by more specific steps which result in the creation of reliable maps (Novak & Caas, 2008). Generally, the steps include listing concepts through brainstorming, creating a hierarchy of the concepts, and using words to link concepts intended to demonstrate relationships (Novak & Caas, 2008).

Today, concept mapping has grown into a more formal mixed methods technique. The maps generated can be used for program planning and evaluation (Kane & Trochin, 2007). Throughout this more structured process, stakeholders provide information which

is refined, organized, and presented visually. In the integrated concept mapping approach presented by Kane and Trochin, maps can be layered to represent different stakeholder groups' understandings and ideas. Visual maps are created by entering data into a digital platform for analysis (Wheeldon, 2010). The imported data provides specific quantitative results that can be presented to an organization for review and further analysis. As mentioned in Chapter I, the maps generated can include similarity matrixes and multidimensional scaling of similarity matrixes (MDS) which result in "Point Maps" and hierarchical cluster analysis which results in "Cluster Maps" (Kane & Trochim, 2007).

Concept mapping can be an effective method for developing evaluation criteria when there is a close association between the researcher and those participating in the research (Marshall & Rossman, 2011). It is also good choice when there is a "shared endeavor between professional researchers and those of a community…[that] may be defined conventionally or [as] people who are associated with a particular organization or initiative" (Kane & Trochin, 2007, p. 67). Furthermore, concept mapping is a good choice when a concept is out of focus and needs clarification by stakeholders to determine the criteria for evaluation (Rossman, & Rallis, 2012).

Concept Mapping as a Methodology

The Center for Disease Control (CDC), the National Cancer Institute (NCI), the Hawaii Department of Health, a number of psychology and social science research studies, and a handful of education research studies have used Kane & Trochim's method for integrated concept mapping to gather evidence for evaluation and/or planning (Abrahams, 2004; Bedi, 2004; Davis, 2003; Edwards, 2002).

As a methodology, concept mapping can be used separately in the planning and evaluation stages of a program (Kane &Trochim, 2007). However, it is often used during both and "provides a quantitative framework" that stakeholders can use to understand important issues and goals within a program (p. 2). When employed with fidelity, concept mapping is a systematic approach to the development of measures for evaluation that is considered both valid and reliable and can be used in conjunction with other qualitative and quantitative study designs (Marshall & Rossman, 2011; Rossman & Rallis, 2012).

Within mixed method research, concept maps represent a transformation of quantitative data to qualitative data. The resulting maps can be used to construct measures that are "unique and novel" within a program (Wheeldon, 2010, p. 88). Furthermore, concept mapping combines the "reliability" of quantitative methods with the "credibility of participant perception" (Wheeldon, 2010, p. 98). As mentioned in Chapter I, there are six phases of concept mapping beginning with (1) Generating statements from stakeholders, (2) Sorting the statements generated, (3) Rating the statements, (4) Analyzing the sorting data, (5) Analyzing the rating data, and (6) Visualizing the findings (Kane & Trochin, 2007).

Within the context of this study, integrated concept mapping was used to evaluate one aspect of a TPP (Kane &Trochim, 2007; Trochim & Linton, 1986). The topic of interest, classroom assessment, was already an established topic among the organization's stakeholders. This consideration, along with the close connection of the evaluator to the organization and the need to clarify the topic, make concept mapping a good fit for this study.

Conclusion

Classroom assessment is discussed and included in this study as an example of a high-impact area of teacher preparation that should be included within program evaluation and is rarely considered by national evaluation entities. This study contends that a democratic approach to accountability furthers the concept of democratic education. Furthermore, it may have the potential to shift the focus of evaluation away from "accountability era" measures which have not resulted in meaningful improvement of TPPs and by extension, K 12 student learning. Concept mapping, by design, involves local stakeholders in a mixed methods process of evaluation.

CHAPTER III

RESEARCH DESIGN AND METHODS

In Chapter II, I presented literature in the areas of democratic accountability, TPP evaluation, classroom assessment, teacher preparation and classroom assessment, and integrated concept mapping, In this chapter, I will present the research design and methods, as well as a description of the study. This study addressed the following research questions.

1. How can integrated concept mapping contribute to the evaluation of a teacher preparation program in the area of classroom assessment outcomes?

2. How can integrated concept mapping shift the focus of TPP evaluation toward a democratic approach to accountability?

The chapter begins with a description of the specific mixed method research design chosen for the study, an overview of integrated concept mapping and how it integrated within the study design. The section articulates the setting and participants as well as a description of the study.

Mixed Methods

In this section, I will explain why I chose a mixed method design. A mixed method approach to evaluation has the advantage of combining at least one aspect of qualitative and quantitative research within the study design. This can happen during data collection, data analysis, making inferences, or while confirming findings (Johnson & Morgan 2016). Mixed methods are often chosen as a research design when there is a need for deeper understanding and the researcher(s) have reason to believe a strictly

quantitative or qualitative design will ignore or diminish critical factors needed to understand the phenomena (Creswell & Plano Clark, 2018).

I chose a mixed methods design for three reasons; first, I hoped to understand the issues of program evaluation and classroom assessment at deeper level. Second, integrated concept mapping is a mixed method by design. The "…qualitative and quantitative components are inexorably interwoven" as data is collected from stakeholders, transformed into statistical representations, and displayed within a series of maps which make the data visible and more useful to an organization (Kane & Trochim, 2007, p. 1). The third reason I chose a mixed method design for this study is that it allowed for the inclusion of qualitative research strategies to understand the experiences of the stakeholders as they go through the integrated concept mapping process.

Mixed methods design has been debated and refined over the decades. Creswell & Plano-Clark (2018) categorized design methods within 12 areas based on specific study features, from which three designs later emerged Convergent, Explanatory, and Exploratory (Creswell & Plano-Clark, 2018, p. 64). The strength of a mixed methods study depends on purposeful decisions about how, when, and why data is mixed. Studies are characterized by their design and model structure. If the data is not mixed as part of the study design, the study is "a collection of multiple methods" and not a mixed methods study (Creswell & Plano-Clark, 2018). This study fits the parameters of a mixed methods design, as explained by Creswell (2008). The survey data and data gathered during the concept mapping process were mixed at various points, which are detailed in later sections of this chapter and the next chapter.

Validity within Mixed Methods

Within mixed methods research, validity is grounded in both the quantitative and qualitative components of the study. Additionally, the validity of a mixed methods study depends on "…employing strategies that address potential threats to drawing correct inferences and accurate assessments from the integrated data" (Creswell & Plano Clark, 2018, p. 251). Finally, validity within mixed methods also needs to be considered in light of the study design.

This study used the Exploratory design; I chose this design because I began the study by collecting qualitative data from stakeholders. The data went through a "development phase" where the qualitative findings were translated and tested using a quantitative tool, Concept Systems Inc. (Creswell & Plano Clark, 2018, p. 94). One characteristic of the Exploratory design is that new measures may be included at a later point in the study. The new measures should be a result of collecting qualitative data that may require the addition of a survey or new experimental activities (Creswell & Plano Clark, 2018). In this study, qualitative data were collected throughout the study and a survey was administered to teacher candidates who make up the third internal stakeholder group.

Threats to validity within Exploratory design include: "Not building the quantitative feature based on qualitative results," "Not developing rigorous quantitative features," and "Selecting participants for the quantitative tests that are the same as the qualitative sample" (Creswell & Plano Clark, 2018, p. 253). The first threat, "Not building the quantitative feature based on qualitative results" was addressed through the

integrated concept mapping process. The quantitative features of the integrated concept

mapping process are a result of collecting qualitative data in the form of open-ended

statements, and the categorizing and ranking of statements.

The second threat, "Not developing rigorous quantitative features" is also

addressed through the Integrated Concept Mapping methods which includes specific tools

for translating qualitative data. The quantitative features of the study include multivariate

statistical techniques such as multidirectional scaling and hierarchical cluster analysis.

These techniques are commonly used within quantitative research (Kane & Trochim,

2007). The third threat to validity, "Selecting participants for the quantitative tests that

are the same as the qualitative sample" was minimized as the stakeholders who

completed the survey belong to a different stakeholder group, and not the stakeholder

groups where qualitative data was gathered. Finally, an advisory group reviewed the

research process to assure the protocols for addressing threats to validity were followed

(Kane & Trochim, 2007).

Qualitative Validity Considerations

Qualitative research is defined by the process of understanding an issue or

problem from a constructivist, advocacy, or participatory standpoint (Creswell, 2008).

The researcher conducts the study in a natural setting and analyzes participant behaviors

and their words. Analysis of data is often rooted in the values and beliefs of study

participants, with the goal of understanding issues in context (Charmaz, 2009).

I begin this study with a constructivist approach to the problem of program

evaluation and used semistructured interviews, a series of discussions, responses to open-

ended statements, and stakeholder categorizing of open-ended statements to study the process. Beginning a mixed methods study using qualitative techniques is a facet of Exploratory mixed methods design (Creswell & Plano Clark, 2018). The semistructured interviews and discussions within this study were recorded and transcripts were generated and edited. I watched and listened to the recordings, reviewed and edited the transcript, and documented statements and interactions of stakeholders in a three-column researcher journal. Grounded theory is often used within the qualitative aspects of Exploratory design (Creswell & Plano Clark, 2018). In this study, grounded theory was used to understand the phenomena that emerged from the process of Integrated Concept Mapping (Creswell & Plano Clark, 2018). Transcriptions within the three-column researcher notes were coded, from which concepts, categories, and finally a "core category" emerged (Charmaz, 2009).

Threats to validity within qualitative research are connected to research methods. There is also a general approach to validity within qualitative research which aims to,"…employ accepted strategies to document the accuracy of …studies" (Creswell & Poth, 2018, p. 259). Grounded theory is an accepted method within qualitative research and will be used within this study, as explained previously. Beyond the use of common methods, Creswell and Poth recommend choosing two of nine suggested validation strategies. In this study, I will use three validation strategies: (1) "Clarifying the researcher bias or engaging in reflexivity," (2) "Member checking or asking for participant feedback," and (3) "Having a peer review or debriefing of the data and research process" (Creswell & Poth, 2018. p. 260).

The first validation strategy, "Clarifying the researcher bias or engaging in reflexivity" was satisfied by writing in the first person throughout this study and by explaining my role within the context and setting of the study. The second strategy, Member Checking, was satisfied through feedback from the stakeholders and advisory groups. The third validation strategy, "Having a peer review or debriefing of the data and research process," was satisfied as I I worked with the advisory group to debrief and understand the data throughout each step in the process (Creswell & Plano Clark, 2018, p. 253). The advisory group was made up of faculty members I work with who also took part in the study. They made decisions related to the concept mapping process and provided advice and feedback. The role and responsibility of the advisory group is detailed in a later section of this chapter.

Quantitative Validity Considerations

While this study began with a constructivist approach to the problem of program evaluation, I moved to a post-positivist view of the problem as the qualitative data was translated using quantitative analysis (Creswell & Plano Clark, 2018, p. 84). This change in worldview from constructivism during the qualitative elements of the study, to a post-positivist worldview during the quantitative features of the study is common within Exploratory design (Creswell & Plano Clark, 2018).

Quantitative analysis requires the researcher to rely on numerical data to demonstrate information and make claims about phenomena (Creswell, 2008). Positivist claims such as "…cause and effect thinking, reduction to specific variables, hypotheses and questions, use of measurement and observation, and the test of theories" designate

boundaries for the development of knowledge (Ivankova, 2002, p. 43). Within, quantitative methods connections are limited to isolating variables and finding causal relationships through magnitude and frequency.

Validity and reliability are strong characteristics of quantitative research due to the selection of variables and measurement tools by the researcher. In this study, qualitative data gathered through the concept mapping process was translated using the Concepts Mapping Inc. software. As previously mentioned, multidirectional scaling and hierarchical cluster analysis were used in this study. These measures have internal reliability and test-retest reliability. Furthermore, data were transformed into a series of "maps," the details of which are explained in a later section of this chapter.

Integrated Concept Mapping

This study used integrated concept mapping (Kane & Trochin, 2007) as an alternative means of TPP evaluation when compared to current evaluation methods. Current evaluation methods often lack a strong local focus, fail to consider contextual considerations of the community, and minimalize or overlook factors related to equity. Furthermore, while some changes have been made in recent years in regard to TPP evaluation, there have been limited results with regard to program improvement and, by extension, improvements in K-12 education. Implementing new program evaluation methods with a democratic focus may have the potential to impact program improvement at a higher rate (Cochran-Smith et al., 2018).

Integrated concept mapping for evaluation is a process grounded in local

stakeholder participation. When local stakeholders drive the evaluation process, they are more likely to be aware of and address strengths and weaknesses within programs (Kane & Trochim, 2008). For example, blind spots within a program are more likely to be uncovered when ideas, suggestions, and the experiences of local stakeholders from diverse areas are the focus of the evaluation process as well as the evaluators (Kane & Trochim, 2008). Within TPP evaluation stakeholders are categorized as internal and external stakeholders. For example, internal stakeholders may be the faculty, staff, and students within the TPP. External stakeholders may be K-12 students and parents, teachers, and administrators. Due to the scope and feasibility of this study, only internal stakeholders took part in the evaluation. In future studies, including external stakeholders may expand the democratic focus of evaluation.

While integrated concept mapping allows for multiple areas of a program to be evaluated at the same time, this study limited evaluation to one aspect of teacher preparation, classroom assessment. Classroom assessment was chosen because it is a high impact area of teacher preparation and is often overlooked within current evaluation systems (Ahsan, 2018). Furthermore, focusing on one area of evaluation allows for a more specific examination of the integrated concept mapping process. Throughout the evaluation process, stakeholders within a secondary school of education identified areas of classroom assessment proficiency agreed on desired outcomes for program participants, created an evaluation tool, and carried out an evaluation of the program in this area.

To date, integrated concept mapping has been used to compare, influence, and

guide evaluation and planning in the areas of social work, social sciences, nursing, and, in a more limited way, the field of education (Marshall & Rossman, 2011; Rossman, & Rallis, 2012; Wheeldon, 2010). The method can be used separately or together in the planning and evaluation stages of a program (Kane &Trochim, 2007). For this study, integrated concept mapping was used for evaluation, although the results may also impact future planning goals.

Integrated concept mapping is recommended if there is a close association between the participants and researcher(s) and requires a facilitator. It is common within integrated concept mapping for the facilitator to be a stakeholder who is taking part in the process (Kane & Trochim, 2008). In this study, the researcher acted as the facilitator. The facilitator's role and responsibilities are explained in a later section of this chapter.

Understanding how integrated concept mapping could work as an evaluation method for this TPP included recording meetings with stakeholders, listening to recordings, and noting the process and stakeholder interactions within a digital researcher notebook. Additionally, the first and last meeting with two stakeholder groups included a pre and post semistructured interview protocol (see Appendices B and C). These questions addressed stakeholder expectations prior to the concept mapping process and their conclusions after completing the process. The questions were designed to help uncover stakeholder understandings, their view of the process, changes in perceptions, and their experiences throughout the process.

This study collected and analyzed data through an integrated concept mapping process (Kane & Trochim, 2007). Following is a description of the setting and

participants and a detailed description of the steps and processes of concept mapping enacted for this research. Each step in the process contains important details that were followed closely.

Setting and Participants

The setting for this study is a School of Education (SOE) within a large, metropolitan university in the western region of the U.S. The school of education is responsible for training elementary, secondary, and special education teachers through three programs. Between 350 and 400 students graduate from the program each year. This study is limited to the secondary education program, which prepares between 120 and 150 students annually. Most graduates remain in the region and accept employment within the service area or close to the service area, which includes cities of 100,00 (or more) residents, smaller towns, and more rural areas. The university has a variety of students, and many students are older than traditional college students, work full or part time, and have families. A significant number of students have a professional career in another field before entering the program (Trotter, 2019).

Concept mapping requires working with stakeholder groups. The first internal stakeholder group was made up of administrators within the SOE, as detailed in Table 1. The second internal stakeholder group was made up of the teaching faculty. The faculty within the secondary school of education includes fourteen full-time faculty members and one to two part-time adjunct faculty members, depending on the semester.

A third internal stakeholder group was made up of secondary education students within the TPP. Student stakeholders did not take part in the preliminary concept

Table 1

Setting and Participants

Setting	Participants	Advisory group
A secondary school of education within a large regional university in the western region of the U.S.	Stakeholder Group 1: Four administrators within the university's school of education.	One or two members of stakeholder group 1 (administrators)
	Stakeholder Group 2: Eleven secondary education faculty members.	
	Stakeholder Group 3: 58 secondary education students.	Two members of stakeholder group 2 (faculty)

mapping sessions; instead, they provided evaluation data in the form of a survey after the evaluation criteria were established by the first two internal stakeholder groups. Students in this program are working toward their first bachelor's degree or are post-baccalaureate students returning to college to earn a teaching certification or endorsement(s).

Each semester, between 60 and 90 secondary education students participate in a student-teaching or internship experience, as part of their final term in the education program. During this time, students are responsible for the instruction and assessment of secondary students within their assigned classes. Because this evaluation is focused on the secondary education program within the university's school of education, it is important to evaluate students across different content areas and include at least forty percent of students enrolled in the program (Johnson & Morgan 2016). The following content areas were included in the sampling: Math, English, Sciences, History/Social Studies, Health, Physical Education, Visual Arts, Theater, World Languages, Business/ CTE, and Music.

Sampling Techniques

Purposeful sampling will ensure the administration and faculty groups represent a range of individuals (Johnson & Morgan, 2016). Each member of the administration and secondary faculty was invited to participate in the study. Also, a variety of teacher candidates, across a broad cross-section of content areas were invited to complete the survey generated in the Utilization Step, which is detailed in a later section of this chapter.

An individual who does not have a connection with the students or courses at the university invited all secondary education students who were student teaching or interning and enrolled in a capstone course, to participate in the study through a Qualtrics survey embedded in the course announcements. All students received ten points of extra credit if they filled out the form, either opting in or out of the study. The Qualtrics survey was routed to the outside individual, so I did not know who opted into the study until after the semester ended and final grades were posted. More than sixty percent of the student group agreed to participate, a threshold of forty percent was needed for the study to be considered valid (Johnson & Morgan 2016).

Eighty percent of administrators and faculty would participate in the study as they routinely participate in program evaluation. Furthermore, this is a small, close-knit group that has historically worked together to improve student learning and who have expressed interest in this topic and the process.

Data for this program evaluation was collected over one semester from January to May. An advisory group independently reviewed each step in the process, including the

survey data (Creswell & Poth, 2018). The advisory group discussed and reviewed all data sources, except the recordings, which were stored in a password-protected site.

Potential Bias

One potential bias within this study is that of *Selection Bias*. If not enough students across content areas had opted into the study, there was a potential that the program evaluation would not reflect the variety of students enrolled in the TPP (Denzin & Giardina 2018). Another potential bias is interviewer bias. This happens if the researcher subconsciously provides clues to interviewees through body language or voice inflection which may result in participants providing answers that align with the researcher's personal beliefs. I made sure to be aware of this bias and made those in the interview aware of this potential bias in order to reduce or eliminate this occurrence (Creswell & Poth, 2018).

Protection of Human Subjects

Participation in the study was voluntary. Within the researcher notes, members of internal stakeholder groups one and two were referred to as faculty member 1, 2, 3 and administrator 1, 2, 3, and so forth. The labels were randomly assigned. An individual who does not have a connection to the university invited all members of internal stakeholder groups one, two, and three to participate through and email and Qualtrics surveys and members of group three to participate through a course announcement and Qualtrics survey. Additionally, the identity of the stakeholders within group three was protected through password protected access to data generated through the surveys. Participant

identity was protected by using codes such as Math, English, Music, and so forth on the surveys. The Institutional Review Board (IRB) approved the protections (see Appendix A).

Structured Concept Mapping Steps and Facilitation

Within this section, a detailed explanation of the integrated concept mapping process is provided. If the process is familiar, a cursory reading of the steps may be all that is needed. Following each step, a description of how the process unfolded within the context of this study is provided.

Facilitating the Process

Key to the process of integrated concept mapping is the facilitator(s) who guides stakeholders through the six-step process. The facilitator can be an "outside consultant [or consultants]" or "an internal member [or members] of the group" (Kane & Trochim, 2007, p. 7). I was the facilitator for this process for this study. As the facilitator, the researcher organized each phase of the process by choosing and inviting all the key stakeholders to participate and facilitating the synchronous and/or asynchronous brainstorming sessions. I also structured the statements and transferred the data collected into the chosen software.

While it was my job as facilitator to organize and manage the concept mapping process, the stakeholders within this study created the content, interpreted the maps, and determined how the data was used to make decisions (Kane & Trochim, 2007, p. 8).

The six steps of the integrated concept mapping are listed below, a more detailed

explanation of each step follows (Kane & Trochin, 2007).

1. Preparing for Concept Mapping

2. Generating Ideas

3. Structuring the Statements (grouping statements and rating each)

4. Concept Mapping Analysis (multidimensional scaling and hierarchical cluster analysis)

5. Interpreting the Maps

6. Utilization

Step 1: Preparing for Concept Mapping

Step 1 includes preparation that will ensure meaningful data is generated through the integrated concept mapping process (see Figure 1). First, an individual or group of individuals initiates the process by defining an issue. This is done by "…identifying the core need, issue, or interest" (Kane & Trochim, 2007, p. 27). After the issue is defined, one or more interested parties move the process forward by making others aware of the issue. This can happen in an organizational meeting, through discussions about future planning or evaluation, or by presenting preliminary data in an informal or formal setting.

Once a decision is made by the initiators to go forward with the integrated concept mapping process, one or more facilitator(s) is confirmed by the initiators. The facilitator(s) may be chosen from within the organization or be an outside entity. Next, the goals and purposes for initiating the process is identified and the focus is further defined by the initiators, which may include the facilitator.

Figure 1

Step 1: Preparing for Concept Mapping Flowchart (Kane & Trochim, 2007)
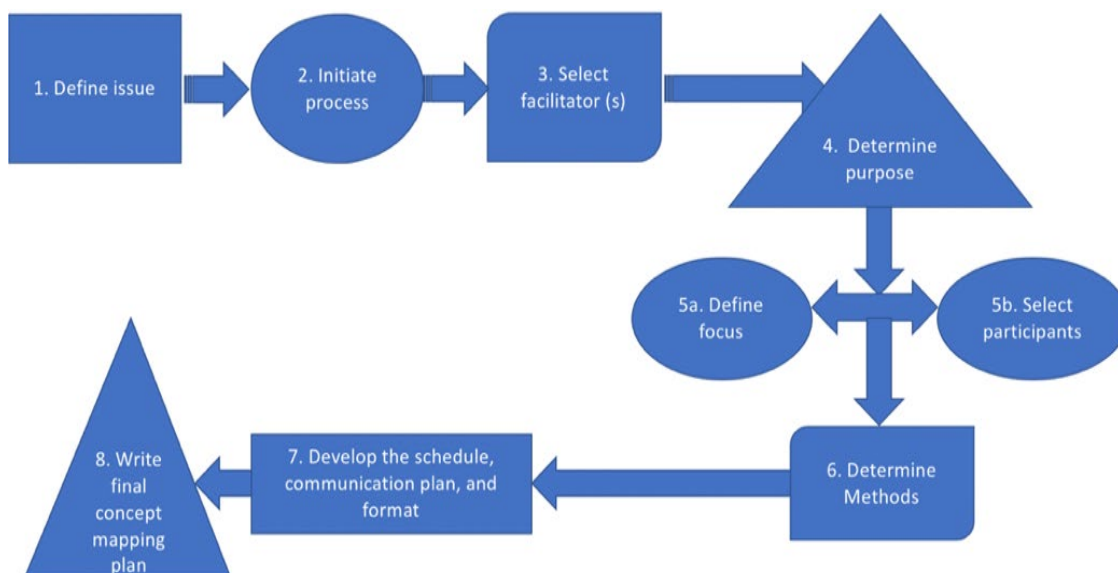
In this study, classroom assessment practices were identified as an area of interest for the TPP due to its impact on learning gains. For this study, the goal was to identify expected program outcomes in the area of classroom assessment and to evaluate program participants' perceived competency in this area. As mentioned previously, the internal stakeholders in this study included administrators, faculty, and current students associated with the TPP. The administrators within this program are charged with overseeing the goals of the TPP. The faculty is responsible for carrying out instruction within the program. The students were enrolled in the TPP and taught in a student-teaching placement or internship.

The first two internal stakeholder groups generated ideas related to classroom assessment, decided how the ideas were connected to the program, and considered the

frequency of occurrence within the program. Subsequently, the third group (students) took part in a survey where they assessed their proficiency in each area identified. The first two groups interpreted the concept maps generated as a result of the data collected and decided how it would be used going forward (Kane & Trochim, 2007).

An advisory group was chosen among the members of stakeholder groups one and two. The advisory group offered guidance and advice to the facilitator throughout the process and was a point of contact for questions and information. The advisory group was also tasked with creating specific, open-ended focus statements that were used in the idea generation, sorting, and rating activities with the larger stakeholder groups. In this study, I oriented the first two internal stakeholder groups to the integrated concept mapping process during the first meeting. Following the orientation, statements were generated by the same stakeholders. Because Kane and Trochim (2007) suggest practicing with the advisory group first, I completed a practice session with the advisory group prior to introducing the activity to the larger stakeholder group.

Step 2: Generating Ideas

In Step 2 of the Integrated Concept Mapping process, ideas are generated with one or more stakeholder groups. The ideas resulting from these sessions should "…describe the conceptional domain of interest" (Kane & Trochim, 2007, p 49). In this study, the "conceptional domain of interest" was the classroom assessment proficiency outcomes for students completing the program.

The facilitator's job during the brainstorming sessions is to manage the process by keeping the group on track (see Figure 2). For example, the facilitator may need to point

Figure 2

Step 2: Generating Ideas Flowchart (Kane & Trochim, 2007)
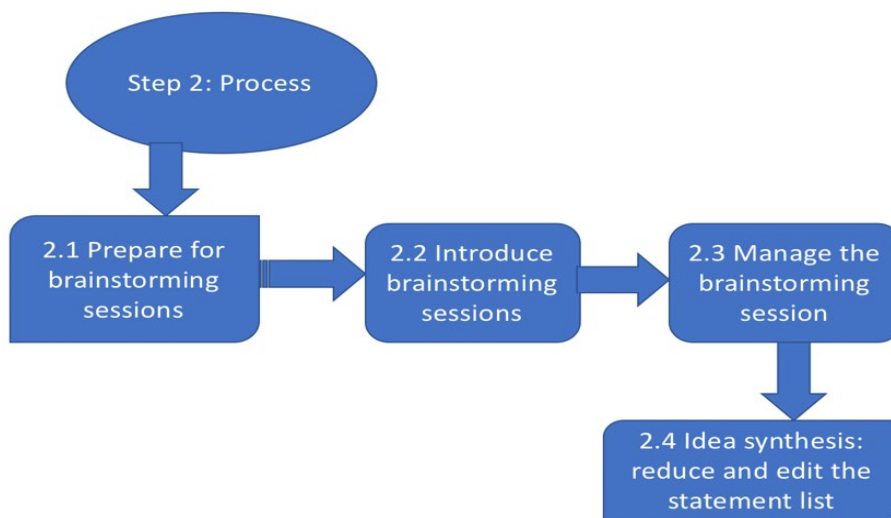
out if an idea is "outside the scope of the brainstorming, while [also] avoiding the role of conceptional gatekeeper" (Kane & Trochim, 2007, p. 57). The facilitator manages ongoing discussions, possible conflicts, and decides when to end the session. Ideas generated in the sessions result in a statement list that will be sorted and rated by the same stakeholders in a later step.

In this study, stakeholder participants were provided a definition and examples for the topic "Classroom Assessments." The PowerPoint slide (PPT) demonstrates materials created by the advisory committee (see Figure 3). Both the definition and examples were previously approved by the advisory committee and included feedback provided by the stakeholders in groups one and two. The advisory committee was made up of three internal stakeholders, one from the administrative group and two from the faculty group.

Using Concept Systems Incorporated software, members of internal stakeholder groups one and two generated multiple statements in response to the following prompt:

Figure 3

PPT Slide: Classroom Assessment Definition and Examples

"Secondary teacher-candidates are considered proficient in the area of classroom assessment if… ." The prompt was designed by the advisory committee. Each of the fifteen internal stakeholder participants in groups one and two completed the statement generation step, also called "Brainstorming" (see Figure 4).

Figure 4

Brainstorming Activity (Groupwisdom, 2022)

Next, statements were synthesized, and a statement list was compiled by the internal stakeholders in groups one and two. The list was reviewed by the advisory committee (Kane & Trochim, 2007). Statements were reviewed for duplication and clarity. Additionally, compound ideas were split apart and "off" topic statements were removed. The advisory committee voted on each statement submitted, a majority vote was required to keep or remove a statement from the list. Of the statements submitted, approximately 38 of 143 were identified as duplicates or "off" topic. The Edited Statement list (Appendix G) of 105 statements was presented to the internal stakeholders in groups one and two for sorting and rating.

Step 3: Structuring the Statements

Step 3 involves two separate tasks. First, the same stakeholders who generate the statements, identify similarities between ideas in the final statements list and rate each of the statements "…by answering the rating focus question for each idea" (Kane & Trochim, 2007, p. 67). The resulting data from these two tasks become the "raw data needed to execute a concept mapping analysis and generate the concept maps. Second, stakeholder demographic information can be collected at this point, which "…allows for subgroup analysis later in the process" (p. 68).

After generating statements in Step 2, stakeholders sort statements into "piles" digitally, according to themes (see Figure 5). As part of the sorting process, each stakeholder names the piles they organized. Stakeholder participants are asked to categorize statements alone if the statement was unrelated to other statements. They were also asked to avoid creating piles according to "…[dissimilarity], priority, or value, such

as "Hard To Do" …or "Other" (Concept Systems, 2022). To be included in the analysis,

at least 75% of statements had to be sorted into piles.

Figure 5

Step 3: Structuring the Statements Flowchart (Kane & Trochim, 2007)
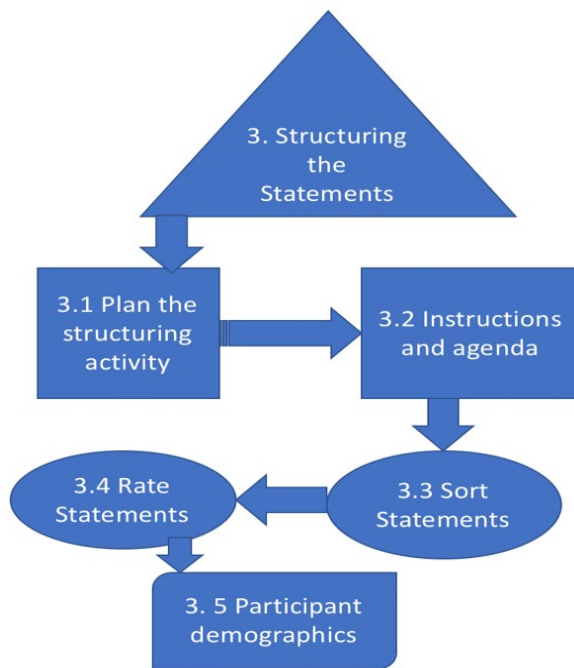
In this study, 14 of the 15 internal stakeholders completed the sorting activity at

the rate required (see Figure 6).

Following the sorting activity, participant stakeholders rated each statement.

During the rating activity, statements appeared randomly on a list, not as they had been

sorted previously (see Figure 7). The entire statement list was rated twice. Internal

stakeholders first rated statements according to importance and then according to

emphasis within the program. Each statement was presented randomly with a Likert scale

below the statement.

Figure 6

Sorting Activity (Groupwisdom, 2022)
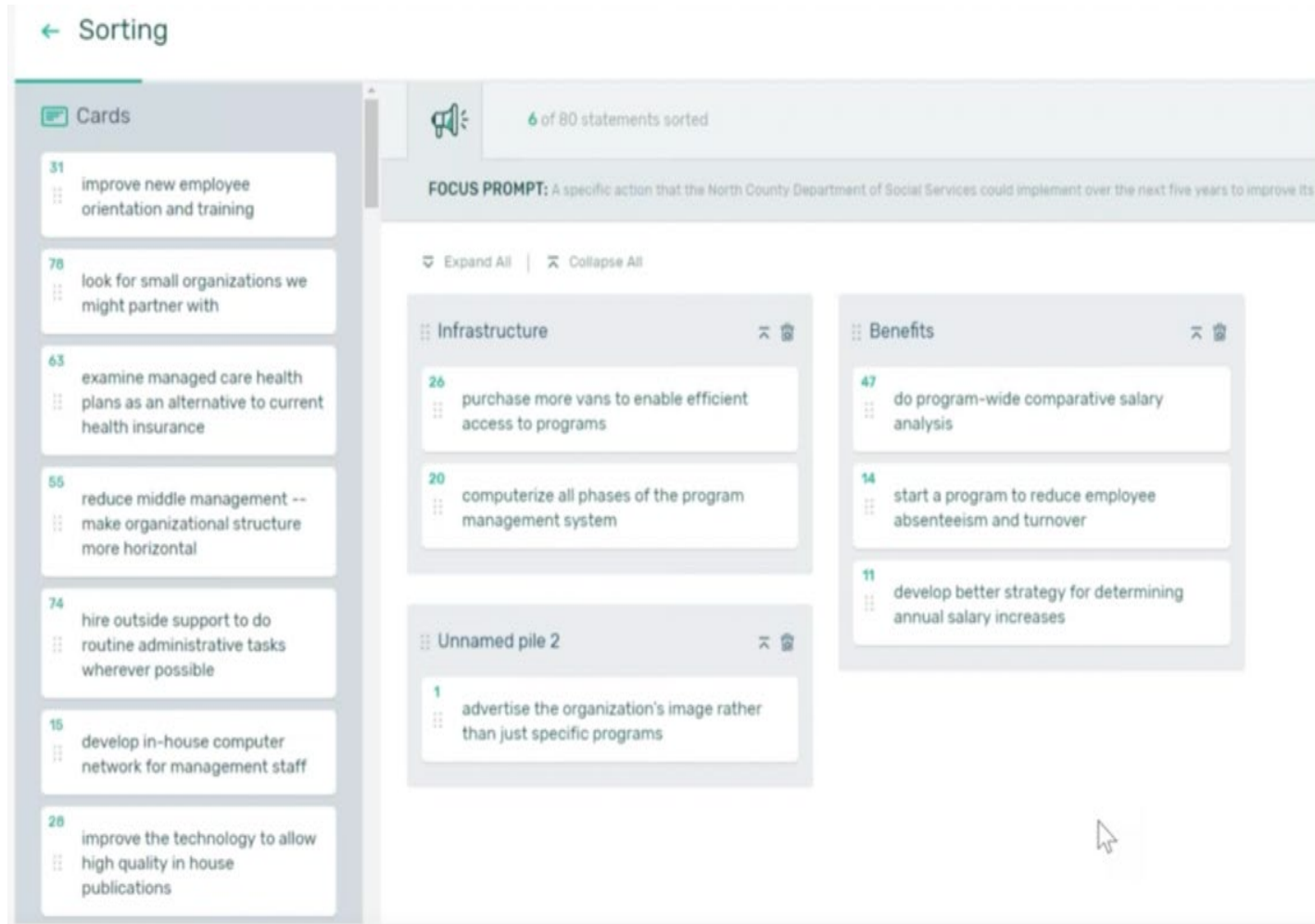
Figure 7

Rating Activity-Importance (Groupwisdom, 2022)

Out of 15 stakeholder participants, 13 completed the *Importance Rating Activity* at a rate of 75% or higher, the required percentage for inclusion within analysis. For rating question two, regarding *frequency* within the program, 11 of 15 stakeholder participants completed the ratings (see Figure 8). Two more could not be included

Figure 8

Rating Activity-Frequency (Groupwisdom, 2022)

because they did not meet the 75% threshold. Both the rating prompt and the scale were

designed by the advisory committee and approved by the internal stakeholders in groups

one and two.

Step 4: Concept Mapping Analysis

During Step 4, the data from Step 3 is displayed in maps through a tool that

transforms the data using quantitative measurements (see Figure 9). The data are

represented as a "geography of thought across multiple communities of interest" resulting

in "a new, quantitative framework" which can be used for planning and/or evaluation of a

program After the statements were sorted and rated, the data were analyzed using

Concept Systems Incorporated (2022; Kane & Trochim, 2007, p. 2). Concept maps can

be generated using computer software programs. Both the Statistical Package for the

Figure 9

Step 4: Concept Mapping Analysis Flowchart (Kane & Trochim, 2007)
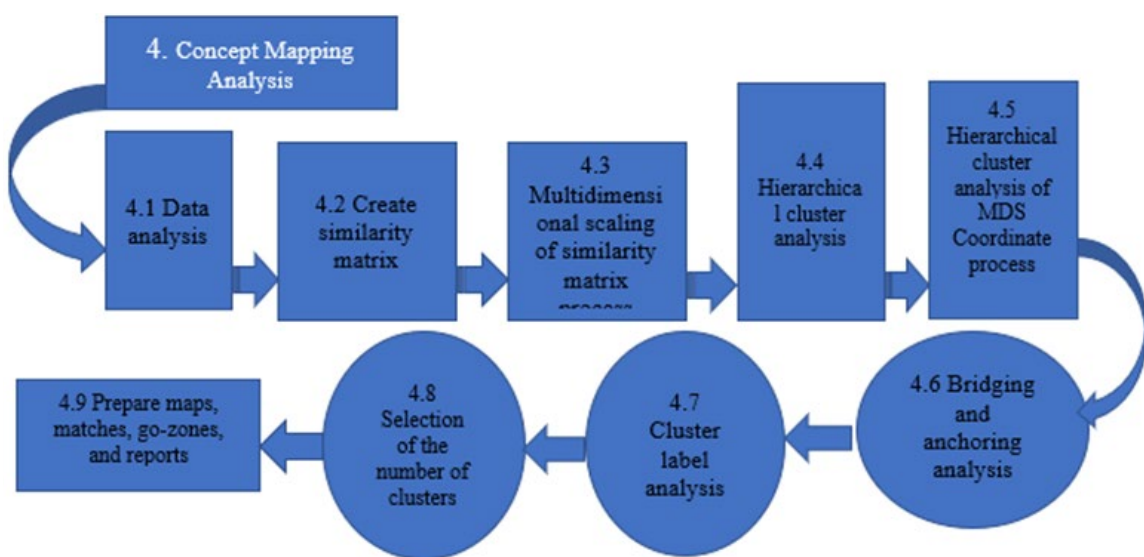
Social Sciences (SPSS) and the Statistical Analysis System (SAS) can be used for analysis. Additionally, other digital programs are available on the internet. However, each require a level of programing skill and advanced statistical proficiency (Kane & Trochim, 2007). Another option, and the option selected for this study, is The Concept Systems software (Concept Systems Inc., 2005), which was developed "...to accomplish the sequence of analyses" described by Kane & Trochim (2007).

When the data are transformed, it appears in a "rectangular data matrix" which demonstrates the ratings from each participant, "The cells are the rating values of each person (row) for each statement (column)…The average values for each statement can then be calculated across participants simply by obtaining summary statistics for each column" (Kane & Trochim, 2007, p. 90). Demographic data may also be considered and represented in a table with rows for each stakeholder and columns for each variable. The summary statistics for each column are used to create a group similarity matrix. From this point a "…a two-dimensional nonmetric multidimensional scaling of the similarity matrix" is completed using the data (Kane & Trochim, 2007, p. 93).

Multidimensional scaling is a type of multivariate analysis used to demonstrate the distance between things, in this case it was the distance between the ideas generated by the stakeholders. The similarity matrix data is used to demonstrate the relative distance between concepts and are shown as points on a map, essentially coordinates. Hierarchical cluster analysis groups individual statements on the point map into clusters of similar concepts. While most calculations are completed using software, decisions need to be made about the number of clusters to be included in the resulting map (Kane

& Trochim, 2007, p. 99).

After the previous analyses have been completed, it is time to prepare for the interpretation of the maps. Choosing the number of clusters to appear on the final map is related to the purpose, focus, and goals of the organization. There is not one correct number. More clusters are advisable if the organization wants to look at every aspect of the issue. Similarly, an organization may want a smaller number of categories if they will be focusing on the statements within each cluster, rather than the many aspects of the issue (Kane & Trochim, 2007, p. 102).

During multidimensional scaling analysis, points were placed on a map. At times a statement was located at a particular point on the map, "…because it was sorted with statements that are immediately adjacent to it," a statement is an "anchor" for a section of the map "…because it reflects well the content in its vicinity" (Kane & Trochim, 2007, p. 10). At other times, statements are placed within other sections of the map because the content is dissimilar. These statements are considered as bridging statements because they bridge between two or more ideas on the map that are more distant based on the sorting (Kane & Trochim, 2007). Understanding where bridges and anchors appear on a map is connected to interpreting the meaning of the different areas on the map as well as the nuances within each area.

The final aspect of Step 4 is to prepare materials for the interpretation session(s). Materials were presented to the first two stakeholder groups in Step 5. Four maps represent major ideas and how they are interrelated, and two additional maps show comparisons of rating results across the different criteria. Each of the maps and their

significance is as follows.

1) *Point Map*: A map showing statements on the map by multi-dimensional scaling.

2) *Cluster Map*: A map showing how statements were grouped by the cluster analysis.

3) *Point Rating Map*: A numbered point map with average statement ratings overlaid.

4) *Cluster Rating Map*: The cluster map with average cluster ratings overlaid.

5) *Pattern Matches*: Pairwise comparisons of cluster ratings across criteria such as rating variables or points in time which uses a ladder graph representation.

6) *Go-zones*: Bivariate graphs for statement values. Two rating variables within a cluster, divided into quadrants above and below the mean of each variable, showing a "go-zone quadrant for statements…above the average on both variables" (Kane & Trochim, 2007, p. 13).

The analysis resulted in a number of concept maps. The most important maps for this study, due to usefulness with regard to creating the survey, were the point maps and cluster maps. Both maps are graphical representations of how the ideas provided relate to each other and are "quantitatively derived" (Kane & Trochim, 2007, p. 1). They were used to help the participant stakeholders develop an "…awareness of issues" within the TPP and move toward "agreement on how to proceed" with the evaluation (Kane & Trochim, 2007, p. 1).

In this study, the advisory group limited the focus and number of clusters which were the presented to internal stakeholders in groups one and two for feedback and voting. The advisory group examined the map for bridges and anchors and used a process suggested by Kane and Trochim (2007) to finalize the number of clusters to be presented to stakeholders for voting.

The goal was to choose the number of clusters that preserves "the most useful detail between clusters" and merges other clusters which "sensibly belong together" (Kane & Trochim, 2007, p. 103). The advisory group decided on the minimum and the maximum number of clusters from a practicality standpoint and provided three choices for a stakeholder vote. The advisory group voted to combine, or not to combine, clusters that appeared closer together on the map until the number of clusters ranged between five and fifteen clusters. While many maps were available, we focused on two of the maps, the point map and cluster maps for evaluation. In the future, we may use additional maps for program planning; however, that was beyond the scope of this study.

Point Maps and Cluster Analysis

The location of a particular point on the map was determined during the sorting activity. The distance between points reflects how frequently the statements were sorted together (see Figure 10). Each point on the map represents a statement and the number appearing by the point is a reference to a specific statement which can be located on the list or by clicking on the number within in Concept System, Inc. software (2022). Points appearing close together represent ideas often sorted together by participant stakeholders, points further apart were not often sorted together.

Within this study, the point map represents faculty and administrator beliefs about what teacher candidates should know and be able to do with regard to classroom assessment. The advisory committee examined different point maps and cluster analysis scenarios. They prioritized the relevance of the ideas collected from the participant stakeholders, in light of the previously agreed-upon definition of classroom assessments.

Figure 10

Transformation of Ideas to a Point Map (Groupwisdom, 2022)
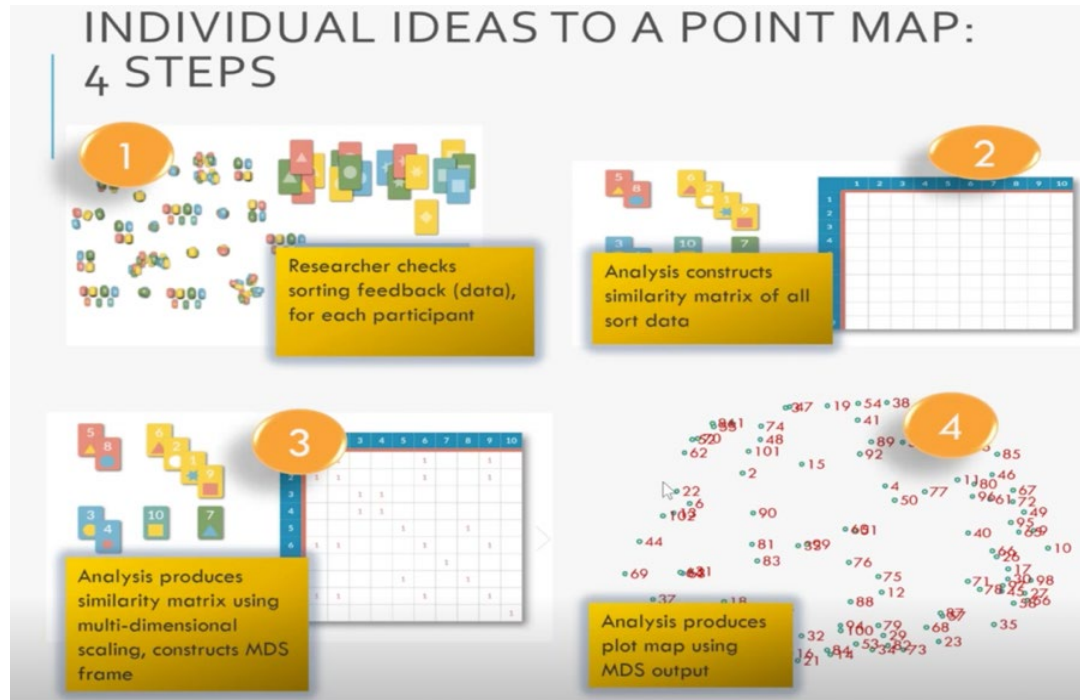
The advisory committee presented scenarios to the internal stakeholders in groups one and two, explained how they were generated, and answered any questions, see Figure 13.

The internal stakeholders in groups one and two attended meetings, participated in discussions, and asked relevant questions before moving forward to vote on the final scenario. The survey, which is the critical component for stakeholder group three's contribution to the evaluation, was a direct result of the scenario chosen.

Multiple iterations of the point maps were generated, based the demographic criteria of the stakeholder participants (see Figure 11). The point map found to be most useful for this evaluation included all participant stakeholders in groups one and two, rather than maps including only a subset of faculty or administration. This map was the

Figure 11

Point Map Scenarios (Groupwisdon 2022)
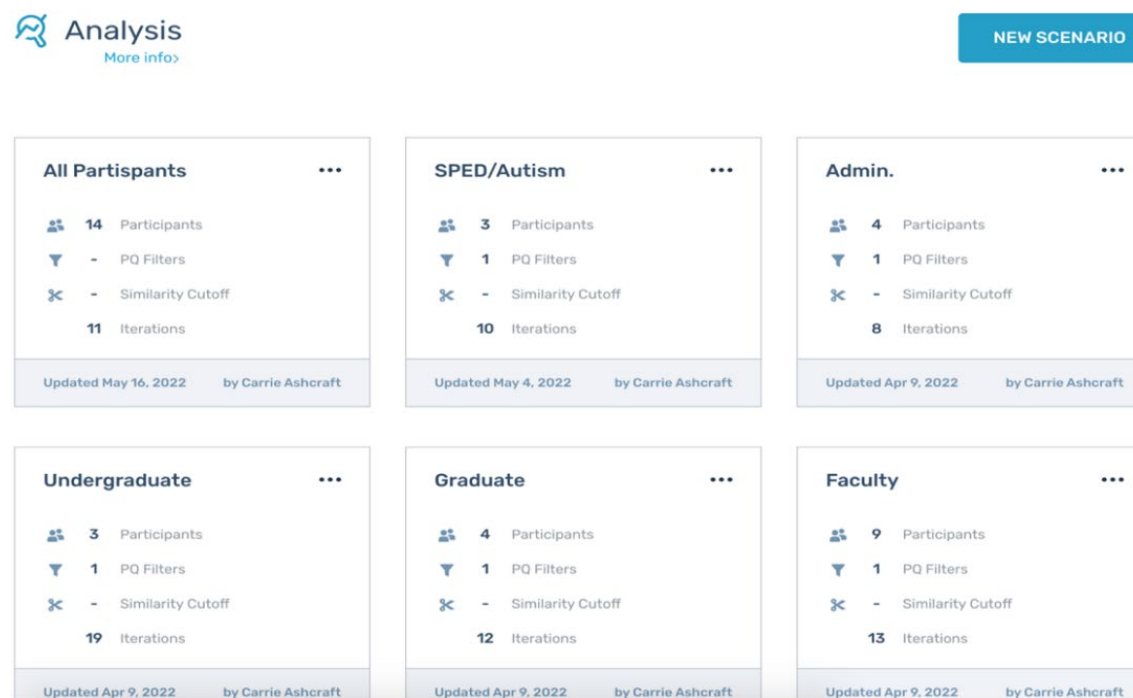
most useful as it demonstrated consensus among the faculty and administration which

was also visible in maps sorted by program demographic, such as *SPED* (special

education) or *Administrators*. In other words, the statements were sorted by theme in a

very similar way by the majority of stakeholders across the stakeholder groups (see

Figure 12). Likewise, there was also strong agreement across faculty and administration

in the rating of the statements.

Point Maps

In the point map (see Figure 13), "bridging" is included. There are two reasons a

point may be placed at a specific location on a grid during multidimensional scaling.

First, it may be there because it was "…sorted by many people with statements that are

Figure 12

Point Map for all Stakeholder Participants in Groups One and Two
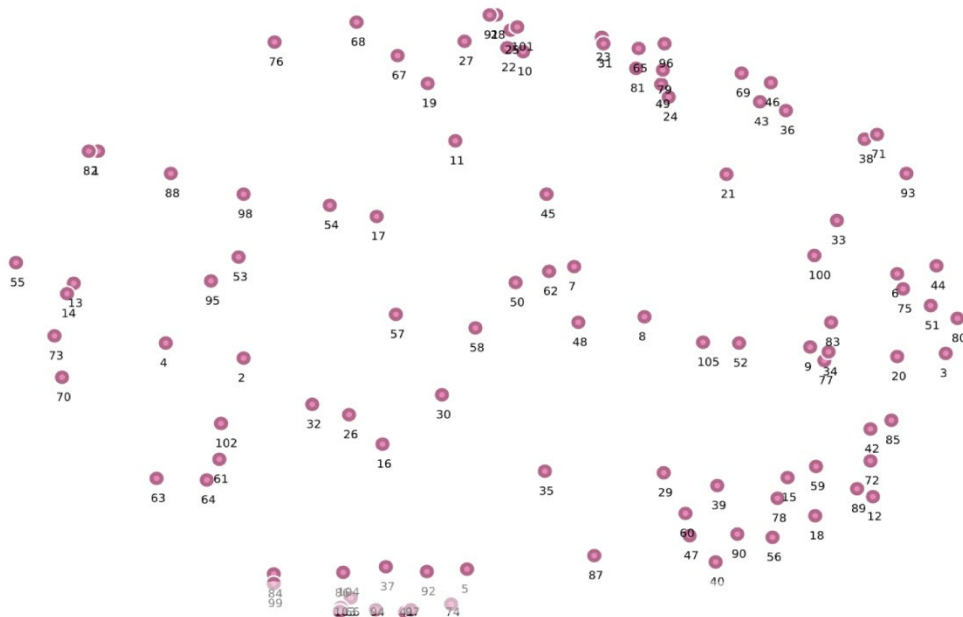
Figure 13

Point Map for all Stakeholder Participants with Bridging
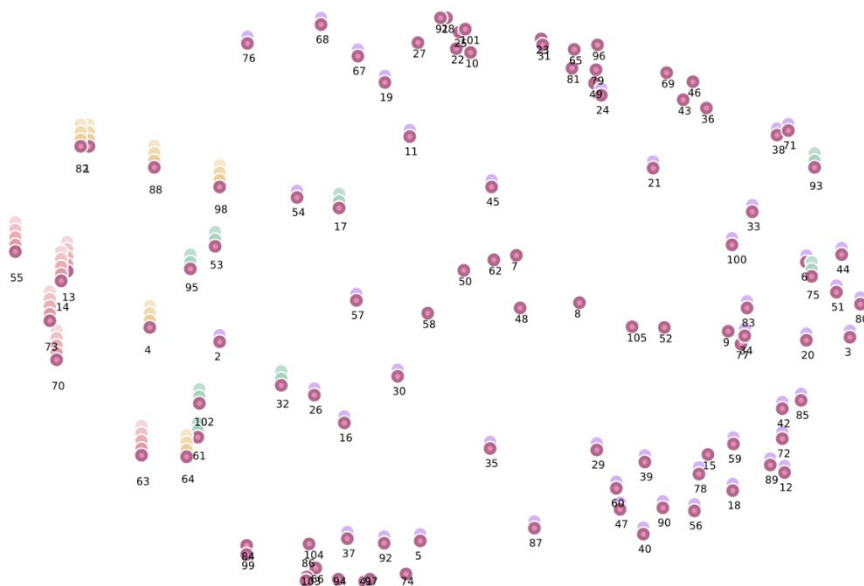
immediately adjacent," such a point is considered an anchor because it "…reflects well the content in its vicinity" (Kane & Trochim, 2007, p. 101). The second reason a point may be placed on a map between two "somewhat distant" points is that "…the algorithm has to pace it somewhere, so it locates an intermediate position" (Kane & Trochim, 2007, p. 101). In the second case, where a point is placed between two points "somewhat distantly," the point is considered a "bridging statement," as it links two ideas on the map.

Within the Concept Systems Incorporated software, there is a "proprietary index for calculating" bridging and anchoring values. The software uses "original sort data" from the project and the results of the multidimensional scaling to calculate the values (Kane & Trochim, 2007, p. 101).

Cluster Maps

Using hierarchical cluster analysis, individual statements from the point map were grouped into clusters of similar concepts within the Integrated Concept Mapping Inc. software. Essentially, clusters are defined as shapes that include the concepts most often sorted together as they appeared on the point map. Choosing a number of clusters ranging between 4 and 20 is recommended (see Figure 14). This range has been found to support more meaningful and actionable evaluation criteria within a program or organization (Kane & Trochim, 2007). Understanding the differences within the point maps is key to understanding how integrated concept mapping works within evaluation.

As the facilitator, I ran multiple scenarios within the software platform, in consultation with the advisory group. As mentioned previously, the map scenario found

Figure 14

PPT Slide: Cluster Map Details for Participant Stakeholders in Groups 1 and 2

to be most useful for this study included all of the faculty and administrators. While the calculations were completed using the software mentioned, decisions were made by the advisory group about the clusters presented to participant stakeholders in groups one and two.

When the number of clusters was reduced below eight, key concepts represented on the map were consumed within unrelated areas. Additionally, choosing nine clusters did not result in significant changes to the map, when compared to ten clusters. The advisory group also decided to include the choice of 15 clusters, as they determined more detailed nuances could be identified within the map. Clusters of 11 through 14 did not group concepts in a way that was meaningfully different from the maps chosen.

In the end, the advisory committee settled on three cluster map scenarios, an 8-cluster map, a 10-cluster map, and a 15-cluster map (see Figure 15).

Figure 15

Cluster Map Scenarios (Groupwisdom, 2022)

The titles for the clusters were generated by the internal stakeholders during the sorting process. The advisory committee decided on the most relevant titles and assigned one to each cluster (Figure 16). Later, the titles were edited by the committee for additional clarity within the survey.

Figure 16

Cluster Map with Eight Concepts (Groupwisdom, 2022)
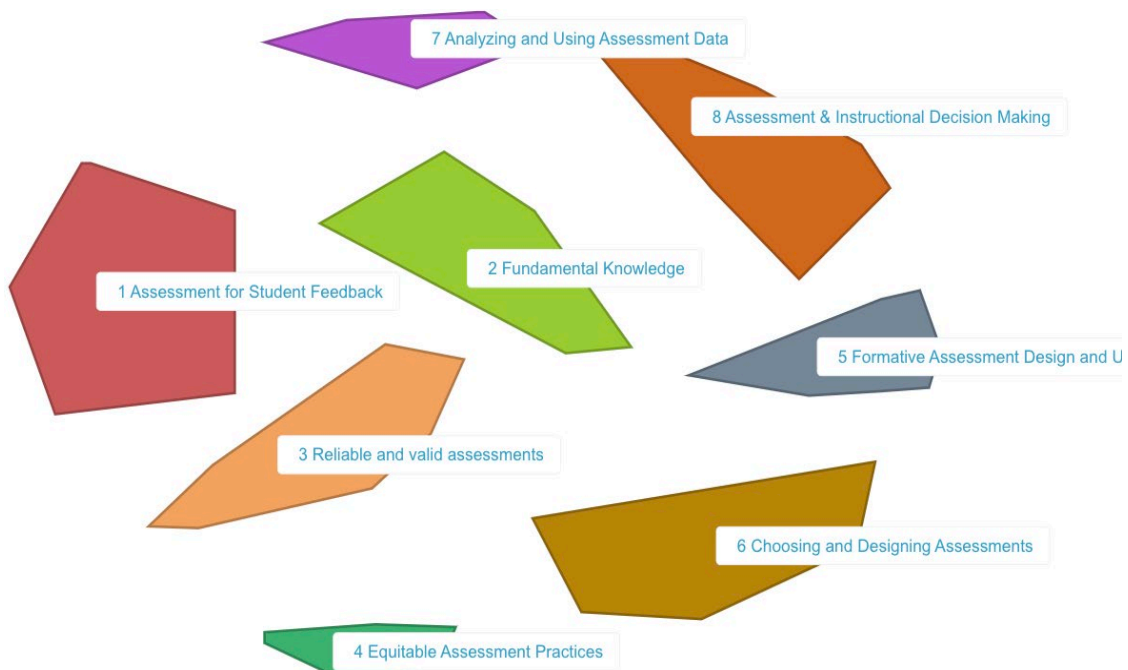
There are a few differences between the concept clusters appearing in Figure 16 (the 8-cluster map) and Figure 17 (the 10-cluster map). Within the 10-cluster map, the addition of a cluster representing "Differentiating Assessments" appears on the lower left side of the map. While differentiation is included in the eight-cluster map, within the concept of "Choosing and Designing Assessments." Also, within the ten-cluster map three clusters at the top and top right represent "Assessment and Data Analysis," "Analyzing and Using Assessment Data," and "Planning for Assessments in Lessons." Within the eight-cluster map, these concepts are combined into two clusters rather than three.

Figure 17

Ten Cluster Map (Groupwisdom, 2022)
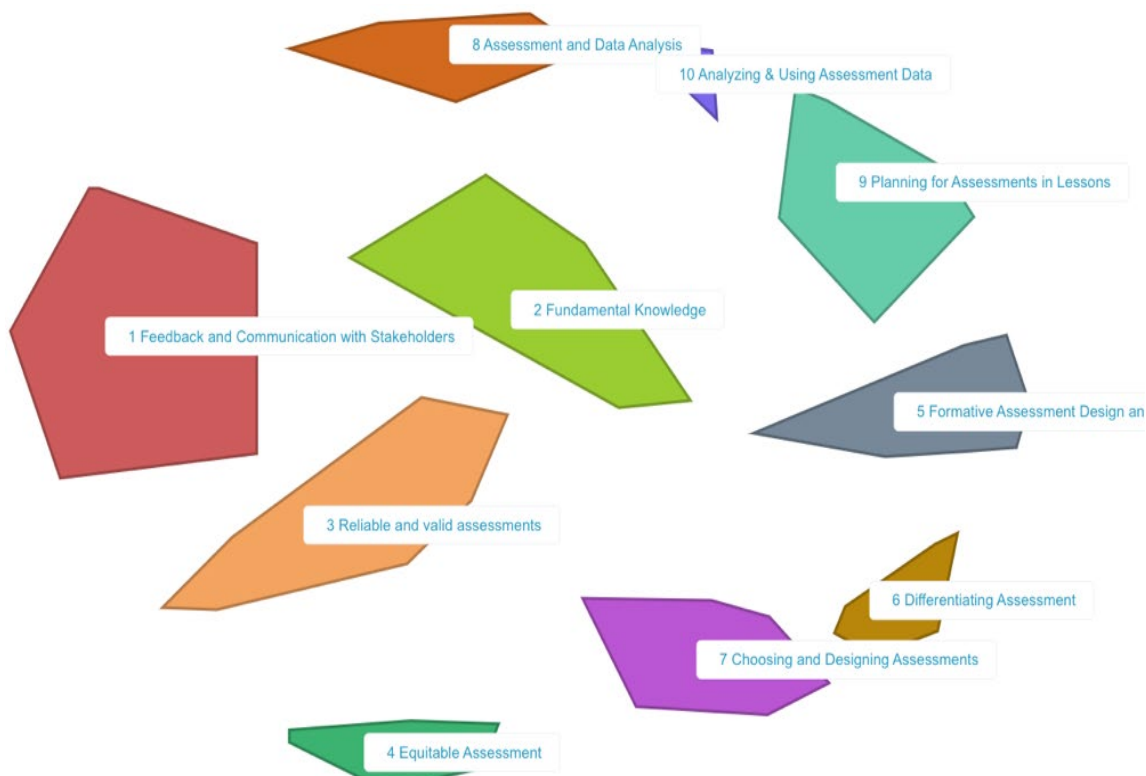
There was some debate within the advisory committee about including a third map with 15 clusters as a choice for the participant stakeholders (see Figure 18). When comparing the 10-cluster map and the 15-cluster map, new clusters such as "Reliability and Validity," "Common Assessments," "Assessment Collaboration," "Validate Assessment Tools," and "Assessment and Teacher Self-Reflection" appear on the 15-cluster map, these topics were not identified as separate clusters within the 10-cluster map.

Figure 18

Cluster Map with 15 Concepts (Groupwisdom, 2022)
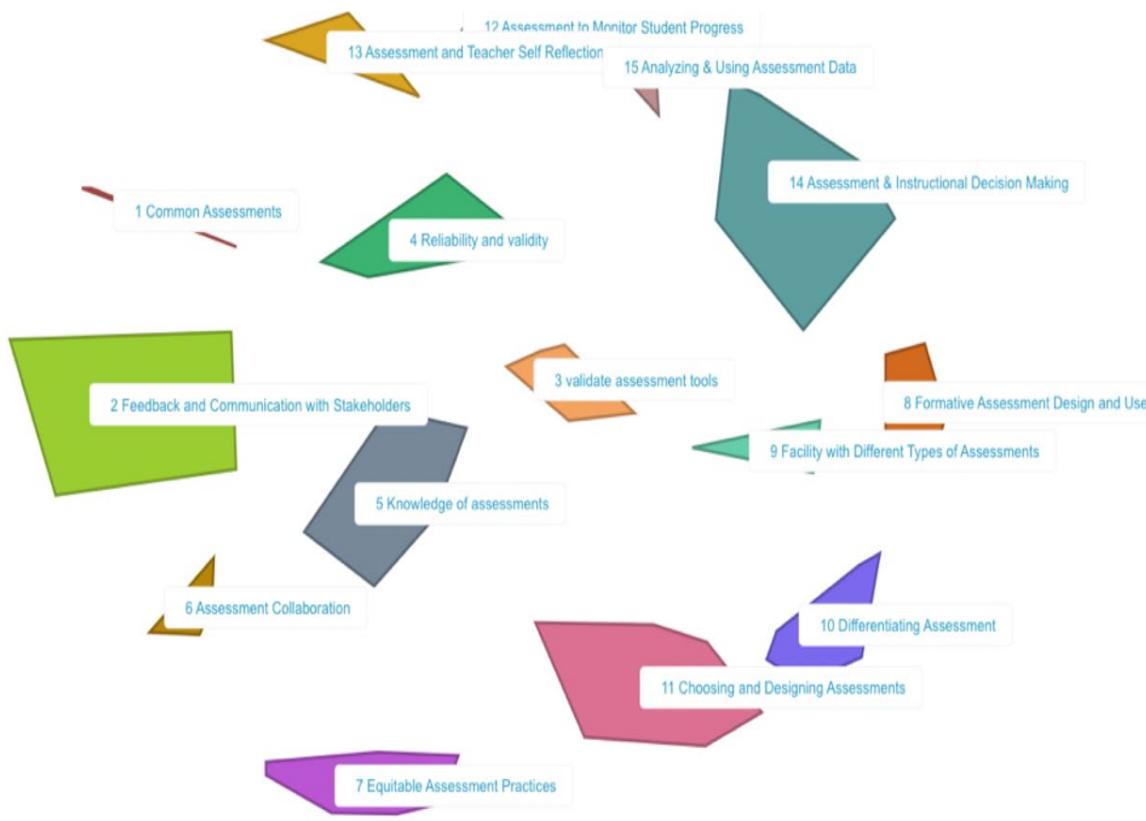
One member of the advisory group was of the opinion 15 areas of evaluation would not be as manageable or meaningful as eight or ten. Another member of the advisory committee believed it would be helpful for participant stakeholders to see and compare the nuances included within the 15-cluster map. A third member of the advisory committee could see value in both arguments. After some deliberation, the 15-cluster map was also presented to the participant stakeholders in groups one and two. The stakeholders in groups one and two voted for one map, the final map was chosen through an anonymous Google Poll (see Figure 19). In the end, the eight-cluster map was chosen by the participant stakeholders, with almost 60% of internal stakeholders who participated identified it as their preference (see Figure 20).

Figure 19

PPT Slide: Cluster Map Voting and Results
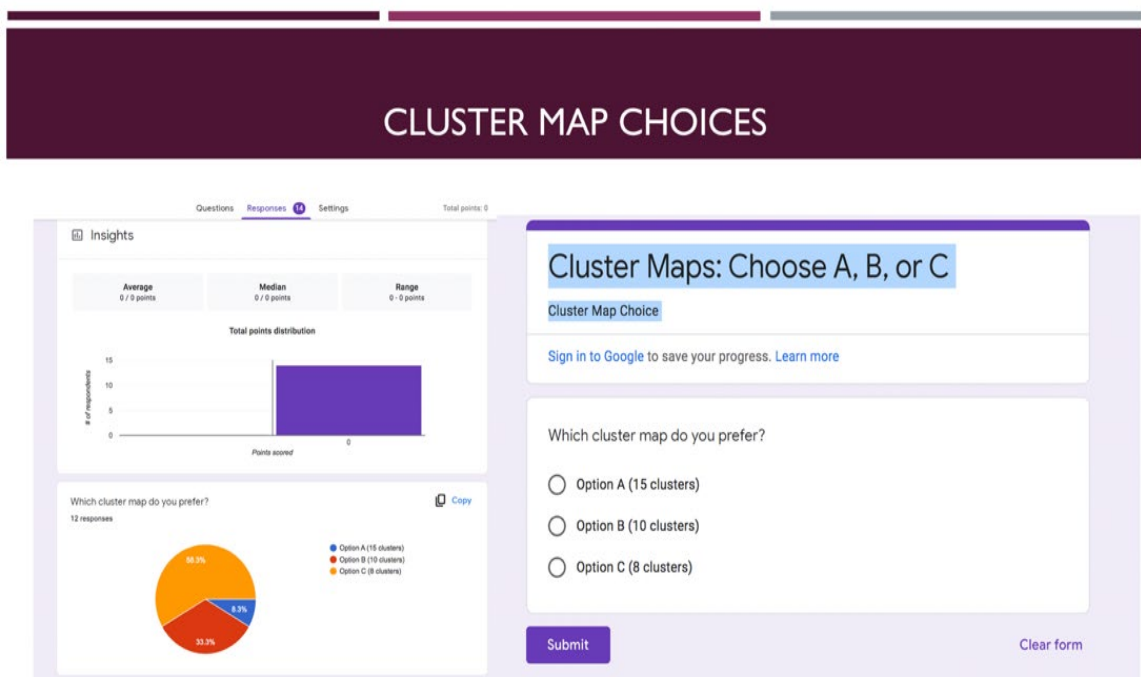
Figure 20

PPT Slide: Final Map Chosen for Evaluation
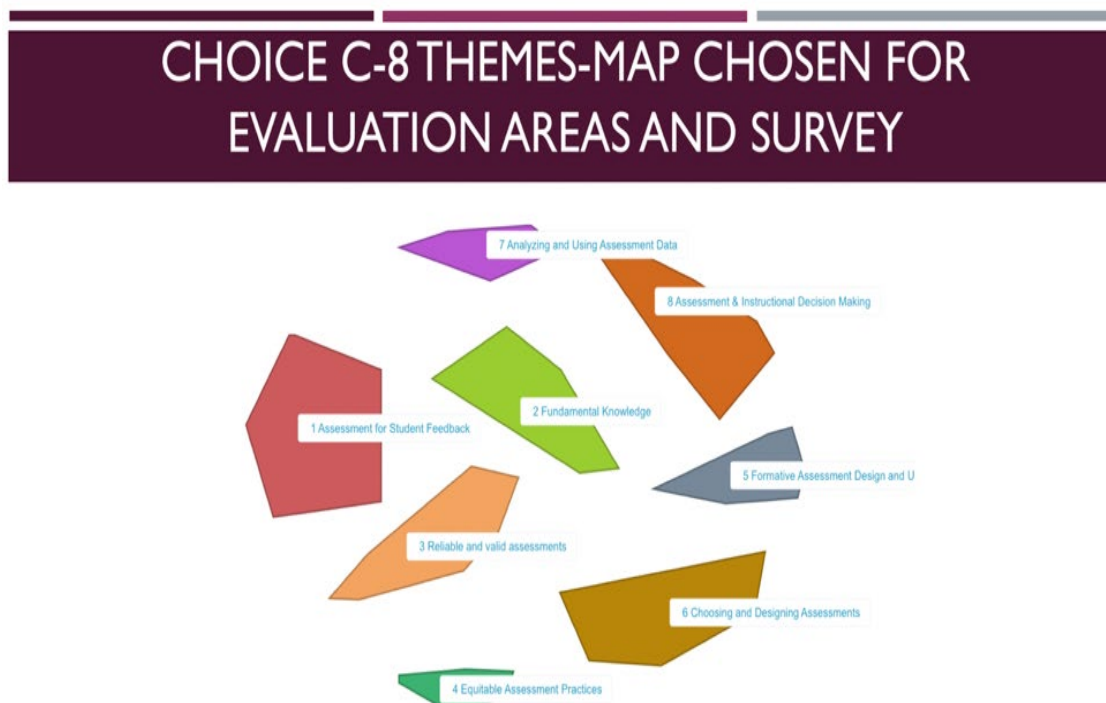
Step 5: Interpreting the Maps

The purpose of concept maps is to clarify the views of a larger group and to allow the group a way to make changes or measure something that is of interest and importance to the group. In step 5, participants take part in an interpretation session where the **co**mpleted maps are presented. Ideally, by the end of the interpretation session, participants should understand the information and come to an agreement about how they may be used (Kane & Trochim, 2007).

In this study, I conducted the interpretation session. First, stakeholders were introduced to the process and provided with an agenda. The groups reviewed the final list of generated statements. Next, I presented the Point Maps, which demonstrated

relationships between statements by their placement on the map. Relationships were determined by the stakeholders through the sorting and rating of statements in an earlier step. After stakeholders became familiar with the data, I provided "a visual tour of the point map and its underlying ideas" (Kane & Trochim, 2007, p. 117), at which point we moved toward consensus. Next, cluster maps were presented to the stakeholders. I explained ideas that appeared closer together conceptually and also appeared closer together on the map. Any concerns with maps were addressed at this point. Once there was agreement, the session was moved into discussion about what the maps demonstrate about the stakeholders' "…ideas for evaluation or planning" (Kane & Trochim, 2007, p. 123). At this point, the Cluster Rating map was presented to stakeholders. This map looks just like the Cluster Map, except there are layers demonstrating the "average cluster ratings" (Kane & Trochim, 2007, p. 124).

Step 6: Utilization

In the final step in the process, some or all of the stakeholders decided how they will use the concept maps for future planning or evaluation within the program. The maps could be used to guide a framework for a planning report or act as an organizational tool for evaluation (Kane & Trochim, 2007, p. 14). The concept maps generated through this study were used to generate a tool for program evaluation. Logic models will be discussed before the evaluation tool is presented because Integrated Concept Mapping includes consideration of the planning or evaluation tool(s) as part of Utilization.

The development of logic models for evaluation within program theory has evolved since the 1970s and is seen in contrast to other evaluation models which use a

single comparison variable to evaluate a program (Kane & Trochim, 2007). Program theory includes the development of a model about how the program, or an aspect of the program works, how it influences the "immediate outputs," and how the outputs impact the long-term outcomes of the program (Kane & Trochim, 2007, p. 160).

After a logic model is generated, there are different options for further development of the methods for evaluation criteria. For example, a list of open-ended questions can be created and used with focus groups. Another option is to develop measures and scales in the form of a survey instrument which may be created by some or all of the stakeholder groups (Kane & Trochim, 2007). In other studies, data relating to long-term goals were organized with graphs showing changes over time in specific areas; in longer studies, it may be possible to demonstrate causal effects (p. 172).

In this study, due to the smaller scope of the initiative, one survey was generated for Stakeholder group three. The survey asked this group of internal stakeholders to rate their proficiency in different areas of classroom assessment identified throughout the concept mapping process and to provide details of their experiences through an open-ended question within each section.

Survey Design and Response Data

The eight-cluster map was used to generate the survey for the teacher candidates within the program. Each cluster was presented as a conceptional area for classroom assessment proficiency and a selection of the statements contained within each area became the rating criteria for the teacher candidates within the program. The title for each conceptual area was edited by the advisory committee during the survey development

process.

Figure 21 shows the edited list of conceptional areas for evaluation included within the survey and presented to the teacher candidates within the program.

Figure 21

Classroom Assessment Evaluation Areas

Concept 1-General Assessment Knowledge (appearing as light green on the map)

Concept 2-Designing and Choosing Assessments (appearing as brown on the map)

Concept 3-Validity and Reliability (appearing as peach on the map)

Concept 4-Analysis, Monitoring, and Tracking (appearing as purple on the map)

Concept 5- Instruction and Interventions (appearing as orange on the map)

Concept 6-Equitable Assessment Practices (appearing as dark green on the map)

Concept 7-Variety and Pacing of Assessments (appearing as grey on the map)

Concept 8-Community and Feedback (appearing as red on the map)

After the title for each area was finalized, the advisory committee, in consultation with internal stakeholders in groups one and two, developed an anonymous survey for teacher-candidates within the TPP.

During the survey design process, the following was considered: First, what format for questions and responses would be the most useful for this survey and the evaluation? Second, how could the length of the survey be balanced with the need to gather specific details relating to the concepts chosen for evaluation? Third, how could we ensure teacher-candidates could understand the specific meanings of words and phrases within the context of this evaluation? Fourth, how could we understand teacher candidates' survey ratings at a deeper level?

For the first consideration, "Which format for questions and responses would be

most useful for this survey and the evaluation," the advisory committee discussed several question and rating designs. It was decided ratings would appear below each statement rather than at the top of each page or conceptional area. This decision was made because survey research suggests including ratings below questions reduces confusion about ratings (Johnson & Morgan, 2016).

The committee decided to use radio buttons for survey response options, rather than checkmarks or another response option as respondents would be limited to one response within this survey. Checkmarks and other response designs are recommended if multiple responses to questions are desired. Additionally, radio buttons correlate to a higher response rate (Johnson & Morgan, 2016). Finally, the response choices appeared as negative choices on the left and moved toward positive choices on the right, this design minimizes positive choice response bias (Johnson & Morgan, 2016). Additionally, response bias has not been found to be significant when using negative language in the response items Johnson & Morgan, 2016).

Finally, the advisory committee debated using five or seven rating choices and decided on a 5-point Likert scale. The 5-point scale includes two options considered "extreme" on the extreme right and left, two options considered "intermediate," moving toward the center, and one option considered "neutral" which appears in the middle (Johnson & Morgan, 2016). While a 7-point scale is often preferred due to accuracy and for purposes of statistical analysis, a 5-point scale is easier for respondents to understand and increases survey repose rates (Johnson & Morgan, 2016). While research demonstrates there is no evidence to support even- or odd-numbered scales (Johnson &

Morgan 2016). One advantage of using an even-numbered scale is the avoidance of a "…neutral response" (Johnson & Morgan, 2016 p. 83). However, within this survey, teacher candidates' perceptions of their skills and knowledge were being measured, and the research suggests a middle response may be needed if respondents are "comparing quality" or assessing their own "perceptions" (Johnson & Morgan, 2016 p. 84).

For the second consideration, balancing the length of the survey with the need to collect relevant data, the advisory committee decided the survey should be uniform in terms of the number of questions included in each section (Johnson & Morgan, 2016). The advisory committee also determined that limiting the questions to six per section would allow us to gather the data needed within each conceptual area, without deterring teacher-candidates from completing the survey due to length (Johnson & Morgan, 2016). The advisory committee identified the five most relevant statements within each cluster on the map and edited each statement for clarity, if needed. Additionally, some statements were combined. Following this process, statements were sent to participant stakeholders within groups one and two for review and feedback, see Figure 22.

Participant stakeholders provided minor editing suggestions, recommended changes in the ordering of questions, and suggested more nuanced changes to questions in two instances. The advisory committee agreed to each of the changes, although in one situation the suggestion came too late, and could not be included. Although the committee was satisfied with the final survey, there are changes we would have made in retrospect. For example, there are some minor editing details we missed. We also would have considered the order of the questions within each section to a greater degree.

Figure 22

Survey Design

```
Concept 1: Classroom Assessment-General Assessment Knowledge

    1.  I know how to include classroom assessments in lesson planning.

Not Proficient     Somewhat Proficient     Proficient     Very Proficient     Extremely Proficient

    2.  I know how to use classroom assessment as part of an ongoing cycle of instruction and
        feedback.

Not Proficient     Somewhat Proficient     Proficient     Very Proficient     Extremely Proficient

    3.  I know the difference between formative and summative assessments.

Not Proficient     Somewhat Proficient     Proficient     Very Proficient     Extremely Proficient

    4.  I know how to align classroom assessments with state standards.

Not Proficient     Somewhat Proficient     Proficient     Very Proficient     Extremely Proficient

    5.  I know how to use classroom assessments to measure student progress against state
        standards.

Not Proficient     Somewhat Proficient     Proficient     Very Proficient     Extremely Proficient
```

Finally, we would consider if the first concept "General Assessment Knowledge" is needed or if the ideas included within this concept are already considered within other conceptional areas.

As part of this study design, the survey was anonymous; however, because the evaluation was for a secondary education program, teacher candidates were asked to identify their teaching content area. Content area choices appeared at the beginning of the survey. However, identifying content areas proved to be a nominal concern within this study as the results were very similar across content areas. This information may be helpful within a future study. For example, combining content area data with responses

provided for the open-ended questions in each area of evaluation may reveal specific classroom assessment proficiency within content areas that may be useful for program planning (for both the education courses and content area courses; see Figure 23).

Figure 23

Content Areas

Identify your content area (choose one):

☐ English Language Arts   ☐ History/Social Studies
☐ Math   ☐ Science
☐ Business/Engineering,   ☐ Dance
☐ Theater   ☐ Visual Art
☐ Music   ☐ World Languages
☐ Physical Education   ☐ Health
☐ SPED

For the third consideration, ensuring respondents understand specific meanings of words and phrases within the survey, we decided to include a brief "Explanation of Terms" within the survey. Inclusion of definitions was a concern for both the advisory committee and the participant stakeholders in groups one and two (see Figure 24).

The advisory committee designed the "Explanation of Terms," and it was then presented and approved by the participant stakeholders, who also provided feedback. Again, while there was general satisfaction with the terms provided, in hindsight, additional terms may have been added or the definitions may have been modified.

For the fourth consideration, understanding teacher candidates survey ratings at a deeper level, we decided to include an open-ended question as the final question within each section (Johnson & Morgan, 2016). The additional question asked teacher-candidates, the participant stakeholders within group three, to provide further examples and explanations related to each conceptual area (see Figure 25).

Figure 24

Explanation of Terms

> **Explanation of Terms:**
>
> *Classroom assessments include a large variety of activities that take place within classrooms such as verbal questioning, portfolios, quizzes, performance demonstrations, and more. Data collected through classroom assessment should be used to inform ongoing instruction and can be defined as the "...formal and informal procedures that teachers employ in an effort to make accurate inferences about what their students know and can do" (Popham, 2009, p. 6).*
>
> *Validity, in the context of classroom assessments, refers to the accuracy of an assessment, "...whether or not the assessment measures what it is supposed to measure" (Florida Center for Instructional Technology).*
>
> *Reliability, in this context, is the ability of the assessment to produce accurate results; typically an assessment is reliable if the results can be replicated.*
>
> *Quantitative analysis is associated with the numerical analysis of data, where information is collected and classified. The goal of quantitative analysis is to understand the occurrence of events and describe them through statistical methods.*
>
> *Qualitative analysis involves identification, examination, and interpretation of patterns and themes within the data collected. It seeks to understand the meaning of within the patterns and themes identified.*

Figure 25

Additional Open-Ended Question for Each Section

> 6. Review the concept and the associated questions in this section. If possible, provide examples and/or explanation of your experiences during student teaching or an internship, related to the concept, questions, and your proficiency ratings.

The data collected through the addition of this open-ended question were extensive and proved to be a critically important aspect of the survey. Although it was an optional question, a good percentage of teacher candidates added details about their classroom experiences or explained their survey response choices. While these data are

helpful within the current study, it also may be useful as part of a larger study related to teacher candidates' perceptions about their skills and knowledge.

Data from the survey were presented to stakeholder groups one and two. These groups decide if the process ends with the evaluation, or if it will be used for future planning.

Data Sources

In this section, I will explain how the research questions were answered through the data sources collected. The first research question in this study is *How can integrated concept mapping contribute to the evaluation of a teacher preparation program in the area of classroom assessment outcomes?* This question was answered at the close of the concept mapping process when the first two stakeholder groups interpreted and used the data generated to create an evaluation tool for the program. It was also answered through survey data received from stakeholder group three. Additionally, this question was addressed as the first two stakeholder groups met to discuss the process of integrated concept mapping in meetings and the semistructured interviews.

The second research question for this study is *How can integrated concept mapping shift the focus of TPP evaluation toward a democratic approach to accountability?* The process of integrated concept mapping includes strong stakeholder collaboration and community considerations which are illustrative of Dewey's democratic ideals defined as, "…primarily a mode of associated living [and] conjoint experience (Dewey, 1916, p. 87). This question was answered through qualitative

analysis of data regarding the process, participation, and the interaction of stakeholders.

**Discussions with Advisory Groups and
Stakeholders**

Throughout the concept mapping process, internal stakeholder groups one and two and the advisory group met to discuss the unfolding concept mapping process and to complete tasks associated with concept mapping (see Table 2). As part of the concept mapping process, multiple meetings took place with Stakeholder groups one and two and the advisory group. Additionally, questions connected to a semistructured interview protocol were asked in a pre/post format to better understand stakeholder perceptions about the integrated concept mapping process, including possible changes in perceptions. These discussions were recorded.

**Transcripts, Recordings, and Three-
Column Coding**

As part of the study design, the process of concept mapping was documented using transcripts and recordings, see Table 3. The purpose of the journal was to document the integrated concept mapping process, to discover how the process may contribute to evaluation of this TPP in the area of classroom assessment outcomes and how it may shift the focus of TPP evaluation toward a democratic approach to accountability.

Statement List

Statements were generated by stakeholders who typed responses onto a google form during live Teams meetings to complete an open-ended statement such as: *If this program is successful in the area of classroom assessment proficiency, teacher*

Table 2

Research Question and Data Source Alignment

Research question numbers	schedule and data source	Data analysis and question alignment
Research Questions 1 and 2	**Week 1** **Discussion**: Advisory group semistructured interview protocol pre/post (see Appendices B and C) **Documents:** Open-ended statement Recording Researcher journal **Artifacts**: PPT	1. Discussion in the advisory group describes **how** the process could be used as an evaluation tool for classroom assessment outcomes. Notes in the Researcher journal describe how the process unfolded within the SOE. Grounded theory will be used to analyze behaviors and discussion and to uncover possible themes. 2. Discussion in the advisory group, in response semistructured protocols questions result in descriptions *of* **how** the process may shift the focus of TPP evaluation toward a democratic approach to accountability. Notes in the Researcher journal describe how the process unfolded within the SOE. Grounded theory will be used to analyze behaviors and discussion and to uncover possible themes.
Research Questions 1 and 2	Week 1 **Discussion:** Stakeholder groups 1 and 2 semistructured interview protocol pre/post (see Appendices B and C) Recording Researcher journal **Artifacts**: PPT	1. Discussion in the Stakeholder groups in response to questions, ***will result in descriptions of how*** the process could be used as an evaluation tool for classroom assessment outcomes. Notes in the Researcher journal describe how the process unfolded within the SOE. Grounded theory will be used to analyze behaviors and discussion and to uncover possible themes. 2. Discussion in the Stakeholder groups, in response to questions, ***will result in descriptions of how*** the process could shift the focus of TPP evaluation toward a democratic approach to accountability. Notes in the Researcher journal describe how the process unfolded within the SOE. Grounded theory will be used to analyze behaviors and discussion and to uncover possible themes.
Research Question 1	Week 2 **Document:** Statements list (from stakeholder groups 1 & 2) Recording Researcher journal **Artifacts:** PPT	1. The Statement list is a artifact identifying each area of importance for classroom assessment outcomes ***and will demonstrate how*** this step in the concept mapping process leads to the generation of evaluation criteria. Researcher journal describe how the process unfolded within the SOE. Grounded theory will be used to analyze behaviors and discussion and to uncover possible themes.
Research Question 1	Week 2 **Document**: Edited statement list (advisory group) Recording Researcher journal **Artifacts:** PPT	1. The Edited Statement list is an artifact identifying each area of importance for classroom assessment outcomes ***and demonstrates how*** this step in the concept mapping process leads to the generation of evaluation criteria. Researcher journal describe how the process unfolded within the SOE. Grounded theory will be used to analyze behaviors and discussion and to uncover possible themes.

(table continues)

Research question numbers	schedule and data source	Data analysis and question alignment
Research Question 1	Week 2 **Documents:** Sorted statements lists Ratings for sorted statements (each member of stakeholder groups 1 and 2 sorts the statements and rates each one). Recording Researcher journal **Artifacts**: PPT .	1. The sorted lists are an artifact identifying the categories for classroom assessment evaluation and their ranked order of importance and *demonstrates how this step* in the concept mapping process leads to the generation of evaluation criteria. Researcher journal describe how the process unfolded within the SOE. Grounded theory will be used to analyze behaviors and discussion and to uncover possible themes.
Research Question 1	Week 3 **Discussions:** Stakeholder discussion of maps; Advisory group discussion of outcomes and creation of survey **Documents:** Concepts maps Evaluation outcomes Recording Researcher journal **Artifacts:** PPT	1. The concept maps are the result of the edited statement lists, the sorted lists, and the demographic data. These maps are the basis for creating the final evaluation outcomes and *will demonstrate how* this step in the concept mapping process leads to the generation of outcomes. The researcher journal describes how the process unfolded within the SOE. Grounded theory will be used to analyze behaviors and discussion and to uncover possible themes.
Research Question 1	Week 4 **Documents:** Program evaluation survey and short answer questions to be completed by Stakeholder group 3 (students) created by advisory groups from concept maps Researcher journal **Artifacts**: PPT	1. The survey is the result of evaluation outcomes and *will demonstrate how* this step in the concept mapping process leads to evaluation of the program in the area of classroom assessment outcomes. Summary statistics will be used to analyze survey data. The researcher journal describes how the process unfolded within the SOE. Grounded theory will be used to analyze short answers within the surveys, as well as behaviors and discussion and to uncover possible themes.
Research Question 1 and 2	Week 5 **Documents:** Survey ratings and short answers from stakeholder group 3 (students) Researcher journal **Artifacts:** PPT	1. Evidence gathered through survey responses and ratings **will demonstrate how** this step in the concept mapping process leads to evaluation of the program in the area of classroom assessment outcomes *and how the process could shift the focus* of TPP evaluation toward a democratic approach to accountability. Summary statistics will be used to analyze survey data. The researcher journal describes how the process unfolded within the SOE. Grounded theory will be used to analyze short answers within the surveys, as well as behaviors and discussion and to uncover possible themes.
Research Question 1	Week 6 **Documents**: Analysis of rating data from surveys (Advisory group). Researcher journal **Artifacts:** PPT	1. Discussion in the advisory group will i*dentify how the survey data could be used to evaluate* the TPP in the area of classroom assessment. The researcher journal describes how the process unfolded within the SOE. Grounded theory will be used to analyze short answers within the surveys, as well as behaviors and discussion and to uncover possible themes.

(table continues)

Research question numbers	schedule and data source	Data analysis and question alignment
Research Question 1 and 2	Week 7 **Discussions:** Utilization of findings and semistructured interview protocol pre/post discussion with stakeholder groups 1 and 2 (see Appendices B and C) Recording Researcher journal **Artifacts**: PPT	1. Discussion within stakeholder groups will result in *descriptions of how* the process and outcomes will be used as an evaluation tool for classroom assessment outcomes, evaluation of the program, and how data will be used going forward. Researcher journal describe how the process unfolded within the SOE. Grounded theory will be used to analyze behaviors and discussion and to uncover possible themes.
Research Question 1 and 2	**Week 8** **Discussion:** Faculty Meeting with Stakeholder groups 1 and 2 (final report is presented) **Documents:** Final report-summary of concept mapping analyses, program evaluation findings from the survey, and process description (as a result of the recordings and researcher journal) Recording Researcher journal **Artifacts:** PPT	1. Discussion within stakeholder groups ***will result in descriptions of how*** the process and outcomes will be used as an evaluation tool for classroom assessment outcomes, evaluation of the program, and how data will be used going forward. Researcher journal describe how the process unfolded within the SOE. Grounded theory will be used to analyze behaviors and discussion and to uncover possible themes.

candidates will be able to... Following statement generation, the statements were reviewed and edited by the advisory group, resulting in the Edited Statement List.

Edited Statement List

An edited statement list was created by the advisory group. Once statements were sorted by keywords, statements were sorted again into categories by the ideas. Compound ideas within statements were split apart and placed in the correct list. Statements remaining after this process were edited for clarity, resulting in the final list.

Sorted Statement List

Stakeholders in groups one and two sorted statements within the final edited statement list. Stakeholders worked remotely and on their own to sort statements. They sorted statements into piles according to similarity.

Table 3

Overview of Data Collection by Research Question

Research question	Source	Data collection procedure
Question 1: How can….	Advisory Group (sub members of Stakeholder Group 1 and 2)	Discussion 1 with pre-research semistructured interview questions and digital recording
		Documents – meeting PPT or other materials, open-ended statement for concept mapping procedures
		Discussion 2 digital recording
		Document - Editing statement list collected from previous concept mapping procedure to pass to the next step
		Discussion 3 digital recording
		Document – meeting PPT or other materials, program evaluation survey
		Discussion 4 digital recording
		Document – meeting PPT or other materials, evaluation findings to be presented to stakeholder groups
	Stakeholder Group 1	Discussion 1 with pre-research semistructured interview questions and digital recording
		Documents – meeting PPT or other materials, statement list created for concept mapping process
		Discussion 2 using concept mapping protocols, digital recording
		Documents – meeting PPT or other materials, individual sorted list from participants (digital) as part of concept mapping procedures, participant stakeholder self-reported demographic information
	Stakeholder Group 2	Discussion 1 with semistructured interview questions and digital recording
		Documents – meeting PPT or other materials, statement list created for concept mapping process
		Discussion 2 using concept mapping protocols, digital recording
		Documents – meeting PPT or other materials, individual sorted list from participants (digital) as part of concept mapping procedures, participant stakeholder self-reported demographic information
	Stakeholder Group 1 and 2 (includes Advisory Group)	Discussion 3 digitally recorded following concept mapping procedures
		Documents – meeting PPT and other materials, concept maps created from software of the previous step in the process, multiple documents created from this step of the process (e.g., classroom assessment outcomes for evaluation, ideal resource allocation for classroom assessment instruction and focus), evaluation survey of teacher competency (created by the advisory group based on this meeting)
		Discussion 4 Part 1 digital recording following concept mapping procedures
		Documents – meeting PPT and other materials, action items and descriptions of how data will be used within the program

(table continues)

Research question	Source	Data collection procedure
		Discussion 4 Part 2 with post-research semistructured interview questions and digital recording
		Documents – meeting PPT or other materials
	Stakeholder Group 3	Evaluation survey of teacher competency created from Discussion 3
	Researcher Journal	Record of researcher activities and decisions throughout the research study Document divided into 3 columns sides with notes on process (what is going on), researcher notes, (preliminary and ongoing analysis), and open-codes.
Question 2: How can….	Advisory Group (sub members of Stakeholder Group 1 and 2)	Discussion 1 with pre-research semistructured interview questions and digital recording
		Documents – meeting PPT or other materials, open-ended statement for concept mapping procedures
	Stakeholder Group 1	Discussion 1 with pre-research semistructured interview questions and digital recording
		Documents – meeting PPT or other materials, statement list created for concept mapping process
	Stakeholder Group 2	Discussion 1 with pre-research semistructured interview questions and digital recording
		Documents – meeting PPT or other materials, statement list created for concept mapping process
	Stakeholder Group 1 and 2 (includes Advisory Group)	Discussion 4 Part 2 with post-research semistructured interview questions and digital recording
		Documents – meeting PPT or other materials

Ratings for Sorted Statement List

After sorting each statement, the first two stakeholder groups rated each statement using a rating focus question generated by the advisory group during the planning stages for Step 3.

Demographic Information for Stakeholder
Groups One and Two

The same stakeholders who sorted and rated the statements answered demographic information related to their professional role and responsibilities within the TPP. This information may be used to understand possible differences in sorting and

ratings by each stakeholder group for future planning, which is beyond the scope of this study.

Concept Maps

Point Map: A map showing statements as they were placed by multi-dimensional scaling.

Cluster Map: A map showing how statements were grouped by the cluster analysis.

Point Rating Map: A numbered point map with average statement ratings overlaid.

Cluster Rating Map: The cluster map with average cluster ratings overlaid.

Evaluation Outcomes

A final list of classroom assessment outcomes was used for program evaluation. It was generated through discussion with Stakeholder groups one and two. Outcomes were based on maps generated through the concept mapping process. The maps represented ideas within the same stakeholder groups. After the discussion, the advisory group created the final list of edited evaluation outcomes, which were checked with members of stakeholder groups one and two before the final list was completed.

Survey With Open-Ended Questions

A survey with additional short answer questions was created by the advisory group and feedback was provided by the stakeholders in groups one and two. Stakeholders within group three rated their proficiency in different areas of classroom

assessment.

Data From Surveys

Data from survey ratings and short answers provided by stakeholder group three were compiled by the facilitator and presented to stakeholders in groups one and two. While some summary statistics were used to compile and present the survey data, grounded theory was used to uncover themes within short-answer responses connected to the survey.

Conclusion

Democratic methods of accountability challenge evaluation norms and current TPP evaluation methods. These methods are tied to evaluation processes put in place during the "accountability era" and are heavily reliant on yearly student summative assessment data and other decontextualized factors. Furthermore, evaluators are too far removed from the community and program to include a variety of stakeholders, who may understand community factors and factors related to equity.

While Dewey viewed diverse members of groups working together within educational organizations as an expression of what should happen within a larger democratic society (Kira, 2019), recent theorists have expanded this view of democratic education to include issues of equity within schools. This conception of "strong democracy," is associated with a self-governing community that values the participation of stakeholders (Apple & Beane, 2007; Cochran-Smith & Reagan, 2022) explains high stakes standardized assessments and associated accountability methods are a critical

reason the United States has not been able to forward the concept of democratic education. Cochran-Smith et al. (2018) theory of democratic accountability asserts TPPs should aim to situate evaluation within the larger concept of democratic education, as defined by Dewey and others.

The proposed method for TPP evaluation, integrated concept mapping, has the potential to shift evaluation from more removed and ineffective methods to more local and impactful methods. By extension, this shift may impact learning within secondary schools due to the potential improvement within TPPs resulting from the evaluation. In sum, the use of integrated concept mapping as a method of evaluation has the potential to disrupt current evaluation methods associated with "accountability era" methods and, at the same time, provide a strong alternative to such methods.

CHAPTER IV

RESULTS

This study had two purposes. The first purpose was to explore how integrated concept mapping may contribute to program evaluation within a TPP, and the second was to determine if Integrated Concept Mapping has the potential to shift TPP evaluation toward a more democratic approach to accountability (Cochran-Smith et al., 2018). For this study, I focused on one area of evaluation, the classroom assessment proficiency of teacher candidates. The results of this study are presented as follows: (1) Survey Results, (2) Analysis of the Concept Mapping Process, and (3) Axial Coding Paradigm. Findings are presented with details of stakeholder decisions which demonstrate how the interactions between stakeholders unfolded. Examining these interactions is critical to understanding the impact if democratic evaluation methods.

Stakeholder questions, experiences, and responses were gathered through semistructured interviews conducted at the beginning and end of the study and informal discussions during meetings throughout the concept mapping process. Analysis of the data included coding of stakeholder responses to interview questions and discussions from the meetings. I used grounded theory to develop a theoretical understanding of the process (Creswell & Poth, 2018).

Survey Results

Data presented in this section were collected through a survey generated by stakeholders in groups one and two and was completed by teacher candidates (internal

stakeholders in group 3). The survey was generated as a result of the Integrated Concept

Mappings Process and is recommended by Kane and Trochim (2007) when concept

mapping is used for program evaluation. The survey was designed to evaluate eight areas

related to classroom assessment proficiency. Survey respondents rated their proficiency

in each conceptual area through five questions and one additional open-ended question

within each area. The open-ended question was designed to provide examples and details

related to each area of evaluation (see Appendix F, Survey Administered to Group 3).

The survey was made available to teacher-candidates through a link in a course

announcement. All teacher-candidates were enrolled in the course during their final

semester in the program. Of the 63 teacher-candidates invited to participate, between 41

and 45 teacher-candidates answered the survey questions. The range of survey responses

for questions varied as a few questions were skipped by respondents. The survey

response rate was between 65% and 71%, a response rate of over 50% is considered

strong (Johnson & Morgan, 2016).

Survey results for the first conceptual area are presented by question and by

average for the area (see Figure 26). Results for conceptual areas two through eight are

provided by average for the area only (see Figures 27 through 31). In each of the five

questions related to general assessment knowledge, teacher-candidates overwhelmingly

chose a rating of Proficient, Very Proficient, or Extremely Proficient, with one teacher-

candidate choosing Somewhat Proficient for one question. The supporting information

within question six included responses for 21 teacher-candidates, just over 50% of

respondents, see Figure 32.

Figure 26

Conceptional Area 1 Question 1: General Assessment Knowledge

Concept 1-Question 1
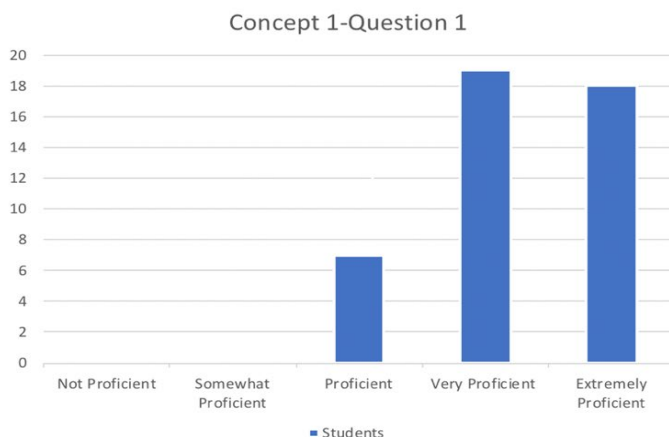
1. I know how to include classroom assessments in lesson planning.

Not Proficient-0
Somewhat Proficient-0
Proficient-7
Very Proficient –19
Extremely Proficient-18

Figure 27

Conceptional Area 1 Question 2: General Assessment Knowledge
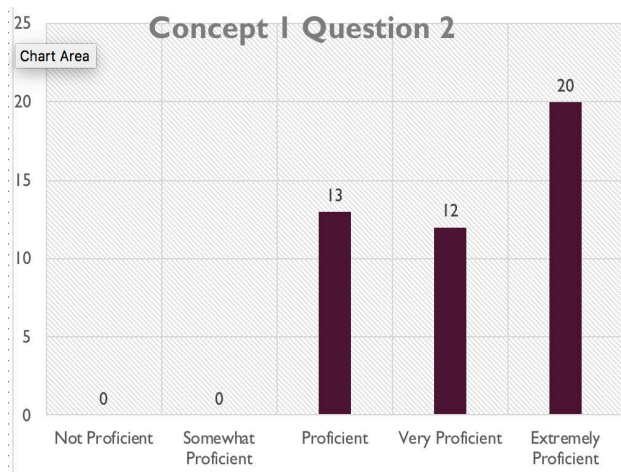
Concept 1 Question 2

2. I know how to use classroom assessment as part of an ongoing cycle of instruction and feedback.

Not Proficient-0
Somewhat Proficient-0
Proficient-13
Very Proficient-12
Extremely Proficient-20

Figure 28

Conceptional Area 1 Question 3: General Assessment Knowledge

3. I know the difference between formative and summative assessments.

Not Proficient-0
Somewhat Proficient-0
Proficient-1
Very Proficient-12
Extremely Proficient-31
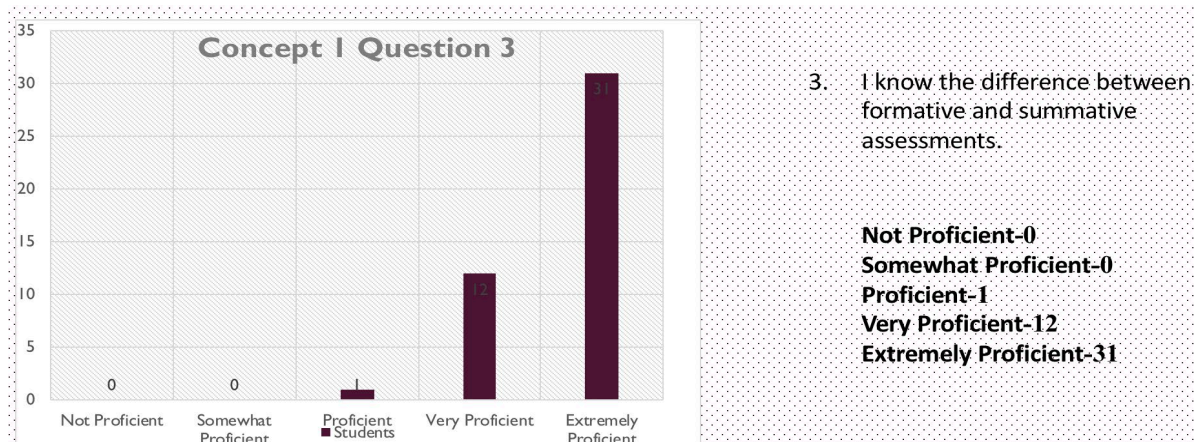
Figure 29

Conceptional Area 1 Question 4: General Assessment Knowledge

4. I know how to align classroom assessments with state standards.

Not Proficient-0
Somewhat Proficient-0
Proficient -4
Very Proficient-12
Extremely Proficient-29
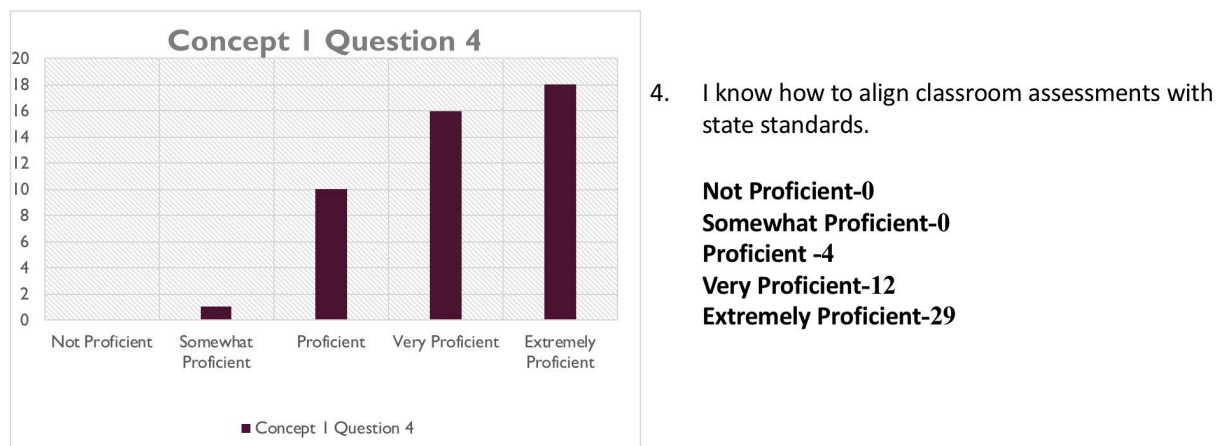
Figure 30

Conceptional Area 1 Question 5:General Assessment Knowledge

Concept I Question 5

5. I know how to use classroom assessments to measure student progress against state standards.

Not Proficient-0
Somewhat Proficient-1
Proficient-10
Very Proficient-16
Extremely Proficient-18
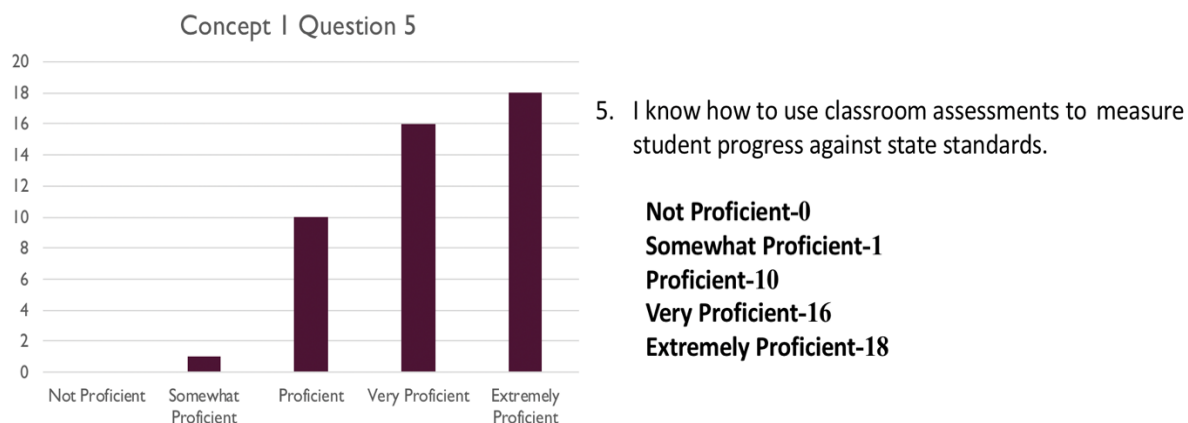
Figure 31

Conceptual Areas 1-2 (overall percentages)

CONCEPTUAL AREAS 1-2: OVERALL PERCENTAGES

Concept 1: General Assessment Knowledge

Extremely Proficient-52%
Very Proficient-32%
Proficient-16%
Somewhat Proficient-0
Not Proficient-0

Concept 2: Designing and Choosing Assessments

Extremely Proficient-43%
Very Proficient-43%
Proficient-18%
Somewhat Proficient-4%
Not Proficient-1%
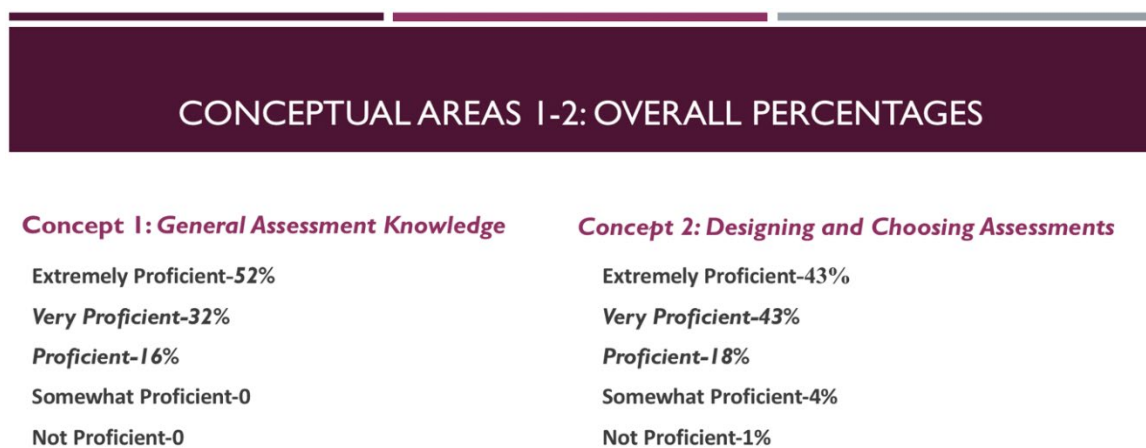
The open-ended responses overwhelmingly aligned with the teacher candidates' ratings within this area of evaluation. The responses also provided insight into teacher candidates' awareness of what they know and were able to do within this area. For example, one respondent stated,

Figure 32

Conceptional Area 1: Selection of Open-Ended Question Responses

CONCEPT 1: CLASSROOM ASSESSMENT-*GENERAL ASSESSMENT KNOWLEDGE*

- I was able to use the standards to teach my students the content that they needed to learn. I also was able to create assessments based on those standards in order to see their progress in learning the material and how I could be a better teacher.
- During my education I never designed a lesson without first having to consult the state standards. It feels like common practice to me now. All of my evaluations for students are designed purposefully to help me understand what they are learning or not learning. My summative assessments are in line with standards and prove that what I have taught has helped students meet those standards.
- I know how to create assessments in that are both formative and summative. I understand how they all work together to understand a current students understanding. I can understand how they build on each other and how they should be based upon the state standards. I will say I struggle with knowing how to take that information in order to make a change or aid students.
- *21 respondents provided comments for Concepts 1*

> I know how to create assessments that are both formative and summative. I understand how they all work together to understand a current student's understanding. I can understand how they build on each other and how they should be based upon the state standards.

However, the teacher-candidate also added, "I will say I struggle with knowing how to [use] that information to make a change or aid students."

In sum, while the survey results may demonstrate the TPP is preparing future teachers well in this area, details from the open-ended questions may demonstrate where teacher-candidates could benefit from additional support.

Likewise, the internal stakeholders in group three (teacher-candidates) overwhelmingly rated their proficiency level within conceptual areas two through eight as Proficient, Very Proficient, or Extremely Proficient; however, there was some variety within the responses. For example, fewer teacher-candidates provided ratings of

"Extremely Proficient" within conceptual areas three, four, five, and eight, opting instead for "Proficient," "Very Proficient" at a higher rate (see Figure 33).

Figure 33

Conceptual Areas 3-4 (overall percentages)

CONCEPTUAL AREAS 3-4: OVERALL PERCENTAGES

Concept 3: *Validity and Reliability*

Extremely Proficient-30%
Very Proficient-37%
Proficient-25%
Somewhat Proficient-7%
Not Proficient-1%

Concept 4: *Analysis, Monitoring, and Tracking*

Extremely Proficient-29%
Very Proficient-29%
Proficient-30%
Somewhat Proficient-8%
Not Proficient-3%
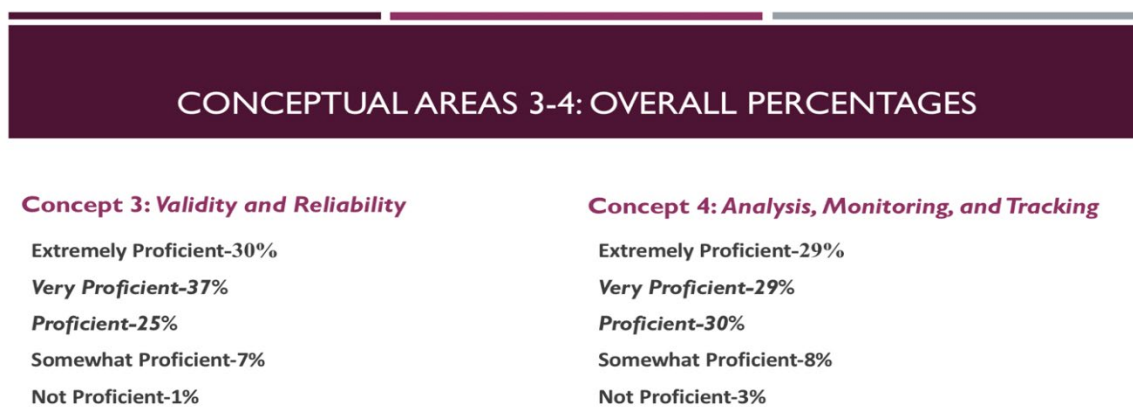
Within areas three and four, survey respondents still chose a rating of "Proficient" or higher most often. However, eight percent chose a rating of "Somewhat Proficient" or "Not Proficient" in conceptual area three and eleven percent chose "Somewhat Proficient" or "Not Proficient" in conceptual area four, much higher than other areas. The greatest response rate variety was found within conceptual area eight, which had the lowest ratings for "Very Proficient," or "Extremely Proficient," comparatively. Conceptual area eight also included the greatest number of "Somewhat Proficient" or "Not Proficient" ratings, sixteen percent.

Again, the supporting information within the open-ended questions for each section provided important understanding about how participant stakeholders in group three perceived their knowledge and skills. For example, within conceptual area four, where eleven percent chose "Somewhat Proficient" or "Not Proficient," responses to

question six reflected a lower confidence level among teacher-candidates. One respondent noted, "…I do not have the most effective skills in this area right now." Another respondent, referring to more formal analysis of classroom assessments, commented "…it's honestly hard to find the time when there are so many other things that I could and should be doing." This statement may provide a greater understanding of the ratings and may also reflect an area for potential program improvement, in terms of teaching candidates ways to quickly track and make sense of classroom assessment data. Responses to the open-ended question this conceptual area could also represent a gap in teacher-candidates understanding about the importance of gathering and using classroom assessment data (see Figure 34).

Figure 34

Conceptual Areas 5-6 (overall percentages)

CONCEPTUAL AREAS 5-6: OVERALL PERCENTAGES

Concept 5: *Instruction and Interventions*

Extremely Proficient-34%
Very Proficient-33%
Proficient-30%
Somewhat Proficient- 2%
Not Proficient-1%

6: *Equitable Assessment Practices*

Extremely Proficient-41%
Very Proficient-31%
Proficient-29%
Somewhat Proficient-2%
Not Proficient-0%
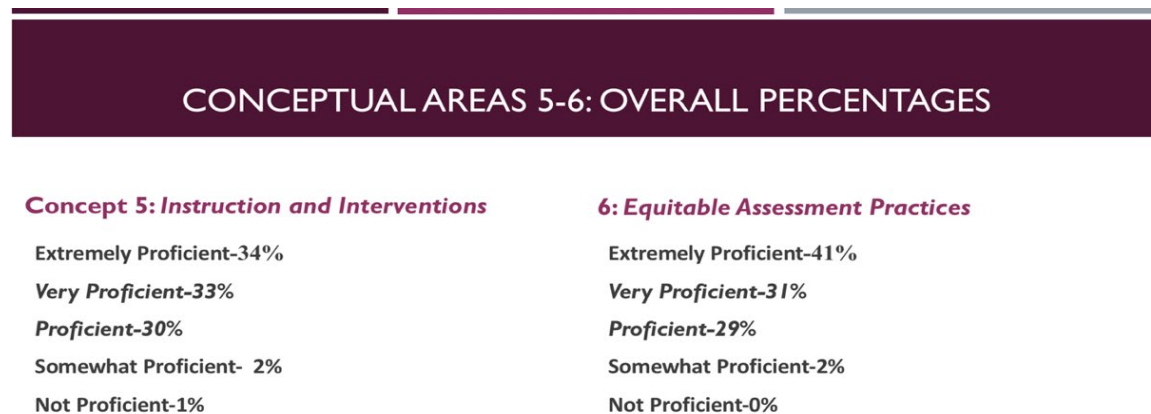
This phenomenon was also apparent in the open-ended responses for conceptual area eight which focused on classroom community and feedback provided, see Figure 35. One respondent commented, "This is an area that I could work on. I do try to comment on student work that is submitted, and if there is something they could change, I say so.

Students would sometimes make those changes and resubmit. But I could be better about

writing on rubrics and giving more constructive feedback."

Figure 35

Reponses to Open-Ended Questions for Concept 4

Concept 4: *Analysis, Monitoring, and Tracking*
- This is an area I am still trying to perfect and feel that I do not have the most effective skills in this area right now.
- I only created one chart to view student progress..
- I am happy to look at general trends and see how students are doing within the classroom. when it comes to graphs and charts, it's honestly hard to find the time when there are so many other things that I could and should be doing. I definitely tried to notice when a concept wasn't being mastered within an assessment, but I don't know that I would make a chart/graph outside of an assignment.
- I was not able to do this during my student teaching but hope to do so in the future. I understand the importance of analysis, I just wasn't able to do it.

This comment may indicate the TPP could improve in this area by providing more

examples, experience, and exposure to providing feedback to students, prior to student

teaching (see Figure 36).

Figure 36

Conceptual Areas 7-8 (overall percentages)

CONCEPTUAL AREAS 7-8: OVERALL PERCENTAGES

Concept 7: *Variety and Pacing of Assessments*

Extremely Proficient-45%

Very Proficient-31%

Proficient-21%

Somewhat Proficient- 2%

Not Proficient-1%

Concept 8:-*Community and Feedback*

Extremely Proficient-24%

Very Proficient-30%

Proficient-32%

Somewhat Proficient- 12%

Not Proficient-4%

Analysis of the Concept Mapping Process

In this section I will present findings from an analysis of the process gathered through semistructured interviews with participant stakeholders. After a preliminary discussion of the Integrated Concept Mapping Process and prior to taking part in the evaluation, stakeholder participants in groups one and two responded to the following questions in a digital meeting:

1. How could the process of integrated concept mapping contribute to the evaluation of our teacher preparation program in the area of classroom assessment outcomes?

2. How could the process of integrated concept mapping shift the focus of evaluation in our TPP toward a more democratic approach to accountability?

3. Are there any additional questions or concerns, relative to this process or what we hope to discover?

Following the evaluation, the same stakeholder participants responded to similar questions:

1. How did the process of integrated concept mapping contribute to the evaluation of our teacher preparation program in the area of classroom assessment outcomes?

2. How did the process of integrated concept mapping shift the focus of evaluation in our TPP toward a more democratic approach to accountability?

3. Are there any additional questions or concerns, relative to this process or what we hoped to discover?

The data collected from these sources were recorded and automatically transcribed through Microsoft Teams. I copied the transcript into a researcher journal with three columns. The first column contained transcript information, the second and third columns were used for the data analysis process. The final codes were generated through the

sorting of open-codes and concepts in multiple iterations (Saldana, 2021, p. 31). See the Delineation of the Coding to Theory Process in Appendix D and excerpts from the researcher journal in Appendix E.

As previously explained, I used grounded theory as a tool for analysis to uncover the common experience of participant stakeholders and to seek for a theoretical understanding of the process which was formatted as an Axial Coding (Creswell & Poth, 2018). During open coding, codes were assigned that symbolically represent the "…summative, salient, essence-capturing, and/or evocative attribute[s]" of the data (Saldana, 2021, p. 5). For example, I chose the following in vivo codes "More voices," "More inclusive of ideas," "Everyone has a say," "Feel more free to say what you really think" because the ideas were similar, but not the same. I did not want to reduce the ideas represented by summarizing the data within a "researcher-denoted code" exclusively (Saldana, 2021).

At other times, I chose to use "researcher-denoted codes," mainly when the data contained longer descriptions, or a variety of examples meant to express a similar attribute. For example, the following "researcher denoted codes" such as "Commitment to understanding the process," "Desire to take part in collaboration" and "Desire to participate in the process" included in the analysis are summaries of stakeholder descriptions of experiences, longer statements, and questions. These codes are meant to capture the "essence" of data collected, rather than specific ideas.

Grounded Theory-Emerging Concepts

Next, I grouped the open codes into similar concepts. I reviewed the open-codes

and concepts repeatedly until no new concepts could be identified and saturation was determined due to the density of each concept and after coming to the end of the data gathered through the interviews and other meetings. After the list of concepts was exhausted, I moved some of the open codes that were misplaced or fit better to different concepts. At this point, I created a second iteration of the concepts. Within the second iteration, some concepts were combined, and the names were changed. For example, within the first iteration, one concept was identified as "More Voices" and in the second iteration, it became "Being Included," which resulted in open codes within other concepts being moved again.

Causal Conditions

In the final iteration, this concept developed further into the category of "Desire to be Heard." The remaining categories and the "core phenomenon" were identified through this process as well as (Creswell & Poth, 2018, p. 85). The categories were then considered by condition, as suggested by Strass and Corbin (1990), both "causal conditions" and "intervening conditions" were identified.

The categories identified as "causal" were *Desire to be Heard* and *Strong Collaboration* because they help to answer the question of "Why" as it relates to the "core phenomenon" (Charmaz, 2009; Saldana, 2021). Both categories provide an explanation as to why commitment to the TPP is strengthened through democratic evaluation methods. For, example the open codes subsumed within the category *Desire to be Heard* were "a strong desire to participate," "a need to include more voices," and "awareness of power dynamics" within organizations. Likewise, the open codes

contributing to the identification of the category *Strong Collaboration* encompassed a desire to engage in meaningful work, to work closely with colleagues, and to participate in shared experiences within the evaluation process.

The ideas, values, and experiences uncovered through open coding demonstrate a strong commitment to the program. This commitment continued to be evident as stakeholders participated in the program evaluation and commented on the democratic processes that took place. The following comments from internal stakeholders are represented within the causal categories *Desire to be Heard* and *Strong Collaboration*.

Desire to be Heard

Faculty Member 1: You can be involved and have your voice be a part of things in, you know, a way that feels less overwhelming.

Faculty Member 5: When you're in a group of your colleagues and you're having these discussions it can be difficult to say things that maybe need to be [said]. [With this process] you're just more free to really say what you think and you don't have to…worry about [creating problems] with your colleagues.

Administrator 3: It's a valuable way to include more voices. There are protocols to include and represent stakeholders.

Strong Collaboration

Faculty Member 6: I see the element of [better understanding] colleagues who work in the program. I mean, this is the opportunity of sharing ideas and getting to know what colleagues are doing in their classroom. So the ideas are kind of organic in that sense…. We are not listening to what we're being asked to do from outside, but we're talking about what we are doing from within and I think that's an opportunity to maybe learn and grow from there.

Administrator 3: I'm thinking back to as a school principal, how I always wanted to have everybody own whatever decision we were making on [a] program or whatever. And certainly this…would certainly avail itself to do exactly that. Everybody has a say, we can all see the big picture, decisions are made together and then we go forward because now it's our program.

Faculty Member 1: [I] see the wide variety of expertise that students are coming in contact with…It really made me want to talk about it with everybody you know, like get together and talk about how this is showing up in our classes and where we could strengthen and where we were overlapping too much and like that kind of stuff.

Intervening Categories

The categories identified as "intervening" include *Taking Ownership*, *Seeking a Deeper Understanding,* and *Considering Impacts*. Each of these categories influences the actions and interactions of strategies impacting the "core phenomenon" (Charmaz, 2009; Saldana, 2021; Vollstedt & Rezat, 2019). For example, strategies directed at or toward the phenomenon included concern with getting it right (*Taking Ownership),* commitment to understanding the process (*Seeking a Deeper Understanding) and* considering the impacts of the evaluation (*Considering Impacts).* The following comments from internal stakeholders are represented within the intervening categories.

Taking Ownership

Administrator 1: "[This process] motivates people. We need to do things like this more often.

Administrator 2: "It is good to look at the program from an internal perspective and take ownership of data by those within the program."

Faculty Member 5: "[There is] better thinking about issues and better buy in-it's much more defensible."

Faculty Member 1: "We are more invested in what we can learn from the evaluation."

Seeking a Deeper Understanding

Faculty Member 9: "[This process] promotes greater participation. Current accreditation processes are labor-intensive, technical, detailed things. [The

perception is that] somebody else needs to take care of it because it is so difficult"
Administrator 2: "[This process] builds in a mechanism to make important conversation happen-to make sure it takes place. This [method] structures the time and space for it to happen."

Faculty Member 6: "This made us-it makes us slow down and think and do. At the school level decisions are often made with partial input- this has multiple points of review and reflection."

Considering Impacts

Faculty Member 1: "[The process could be] helpful to make future decisions about the program. Could this be combined with other methods of evaluation such as edTPA?"

Administrator 3: "This process may lead to better alignment within the program and impact the future direction of the program. It may help with future planning and identifying gaps. We may see difference in perspective across roles within the program as it relates to classroom assessment."

Administrator 1: "This tool could be valuable within other contexts -such as within K-12 schools."

Faculty Member 7: Maybe we include more stakeholders who are at different points in program-students. This could be used to help students have a better experience throughout the program. It could provide information earlier within the program."

Theoretical Understanding

The "core phenomenon," "causal conditions," and the "intervening conditions" were organized within an "axial coding paradigm" (Charmaz, 2009; Saldana, 2021). I used "selective coding" to develop the following theoretical understanding: *Stakeholder commitment to evaluation is strengthened through democratic accountability methods.* This theoretical understanding and a more complete description of the understanding is presented in the Axial Paradigm (see Figure 37; Creswell & Poth, 2018, p. 85).

Figure 37

Axial Coding Paradigm

Intervening Condition: Grass-roots effort by stakeholders	Intervening Condition: Voluntary participation	Categories identified:
Causal Condition: Evaluation of Program → Core Phenomenon: Stakeholder commitment to evaluation is strengthened through democratic accountability methods. →		1. Desire to be Heard (causal) 2. Taking Ownership (intervening) 3. Seeking a Deeper Understanding (intervening) 4. Strong Collaboration (causal) 5. Considering Impacts (intervening)
Context: Integrated concept mapping process	Context: Secondary TPP	

Charmaz (2009) explains that different understandings of what constitutes a theory is depends on the "epistemological underpinnings" requisite for establishing a theory. From a positivist standpoint, theory is treated as a "…statement of relationships between abstract concepts as variables or construct operational definitions of…concepts for hypothesis testing through accurate, replicable empirical measurement" (p. 126). From this perspective, theories often cross academic fields and are often ubiquitous within textbooks and academic studies. Positivist theories are used to explain and predict; they "emphasizes generality and universality" and may result in more narrow explanations that reduce the complexity of the phenomenon (p. 126).

Alternatively, considering the concept of theory from an "Interpretative" standpoint places emphasis on "understanding rather than explanation" (Charmaz, 2009, p. 126). This approach to understanding theory relies on the "… theorists' interpretation of the studied phenomenon" (p. 126). From this perspective, theories rely on seeking patterns and connections. An "Interpretive" approach to theory asks "…for imaginative

understanding of the studied phenomenon" and "assumes "…multiple realities;

indeterminacy; facts and values are inexorably linked; truth as provisional; and social life

as processional" (p. 127).

In this study, the proposed theoretical understanding seeks to interpret the

participant's contributions within the specific experience of evaluating a secondary

education program using a democratic method of evaluation. The following theoretical

understanding emerged as a result of Axial Coding, consideration of the supporting

categories, and further reflection on open codes. It is described as follows: Using

grounded theory, I found a *Desire to be Heard* (causal condition) and *Strong*

Collaboration (causal condition) among stakeholders within a Secondary TPP (faculty,

administration, and teacher-candidates) led to *Taking Ownership (causal condition)* of

the evaluation process and evaluation outcomes. The grass-roots method for evaluation

(integrated concept mapping), led stakeholders to *Seek for a Deeper Understanding*

(intervening condition) of both the evaluation topic (classroom assessment) and the

method of evaluation. Finally, taking part in the evaluation process led stakeholders to

C*onsider Impacts (intervening condition)* of the evaluation within the program and within

broader contexts.

CHAPTER V

DISCUSSION

This chapter presents a summary of the study, discussion of the results in light of published literature, limitations of the study, implications of the study, and recommendations.

Seeking out TPP evaluation methods that forward democratic methods of accountability was largely ignored throughout the "accountability era," which may have stalled TPP improvement. Furthermore, classroom assessment practices were often ignored because of an intense focus on annual summative assessments. Finally, policy makers have been too far removed from local education evaluation to understand areas of teacher preparation in need of improvement and the community the program serves (Cochran-Smith et al., 2016).

Primarily, TPP evaluation is still a removed process completed in conjunction with outside organizations. The focus of evaluation often includes a focus on annual student summative assessments, tied to the individual teachers and TPPs where they graduated (McDiarmid, 2019). These measures do not often promote or encourage democratic methods of evaluation. Likewise, a strong focus on community considerations and "strong equity" is largely absent within program evaluation, topics recommend by those calling for currently calling for reform (Apple & Beane, 2007; Cochran-Smith et al., 2018; Cochran-Smith & Reagan 2022).

Summary of the Study

This study took place within a TPP in the western region of the U.S. The purpose of the study was to investigate alternative program evaluation methods, which may impact current evaluation methods. I chose to situate this study within a Democratic Accountability framework (Cochran-Smith et al., 2018) and to focus on classroom assessment because of its potential impact on K-12 student learning (Ahsan, 2018; William, 2011).

Stakeholders, including faculty, administration, and teacher-candidates (students in their last semester of the program) within the TPP took part in a grass-roots program evaluation using Kane and Trochim's (2007) method for integrated concept mapping. This method is recommended when there is a close connection between the program and the evaluator, and it is most often used to identify criteria for evaluation and planning (Kane & Trochim, 2007). In this study, a survey with open-ended questions was developed as an evaluation tool, following stakeholder participation in the concept mapping process.

Discussion of Results and Published Research

The first research question for this study was: *How can integrated concept mapping contribute to the evaluation of a teacher preparation program in the area of classroom assessment outcomes?* The second research question in this study was: *How can integrated concept mapping shift the focus of TPP evaluation toward a democratic approach to accountability?* The results of this study are presented in connection to the

research questions as well as existing research.

Democratic Education

John Dewey's theory of education prioritizes democratic education practices, including democratic school governance (Dewey,1916; Kira, 2019). Dewey's conception of education was intended to forward principles of democracy within schools and, by extension, society which are viewed by Dewey and others as inexorably connected (Cochran-Smith et al., 2018; Kira, 2019). Within this concept of education and democratic school governance is the consideration of community and contextual factors as they relate to evaluation of schools and related programs. Newer proponents of democratic education have added the consideration of "strong equity" within democratic education (Cochran-Smith & Reagan, 2022). Strong equity is concerned with defining equitable factors within education programs and schools such as graduation rates, school participation, and equitable access to the curriculum, materials, courses (including advanced courses), clubs, and so forth. Strong equity also includes advocating for traditionally marginalized groups and individuals by identifying and naming factors relating to equity within schools and related educational entities, rather than talking about equity in abstract terms (Cochran-Smith & Reagan, 2022).

Community Considerations, Contextual
Factors, and Equity Considerations

The evaluation method used in this study made it possible to include community considerations, contextual factors, and equity considerations which should be part of program evaluation and may contribute to greater improvements within programs (Apple

& Beane, 2007. Cochran-Smith et al., 2018; Cochran-Smith & Reagan, 2022). Within

this study, two of the eight conceptual areas identified by stakeholders were connected to

equity and community factors (see Figure 38). Conceptual areas six and eight asked

teacher-candidates to consider their knowledge and skills in these areas and comment on

their experiences within the classroom during their student-teaching experience. The

identification of these conceptional areas as part of a grass-roots evaluation process

demonstrates how integrated concept mapping may shift evaluation toward a more

democratic approach to accountability (Cochran-Smith & Reagan, 2022).

Figure 38

Conceptual Areas Identified

Concept 1-General Assessment Knowledge
Concept 2-Designing and Choosing Assessments
Concept 3-Validity and Reliability
Concept 4-Analysis, Monitoring, and Tracking
Concept 5-Instruction and Interventions
Concept 6-Equitable Assessment Practices
Concept 7-Variety and Pacing of Assessments
Concept 8-Community and Feedback

Within Concept 6, Equitable Assessment Practices, teacher candidates perceptions

of their knowledge and skills were measured in the following areas: modifying

assessments, making-trauma informed decisions, using culturally relevant assessments,

and being aware of personal bias that may influence assessments (see Figures 39 and 40).

Furthermore, teacher candidates commented on how they addressed these issues of equity

Figure 39

Concept 6: Questions 1-4

CONCEPT 6: EQUITABLE ASSESSMENT PRACTICES

Concept 6: *Equitable Assessment Practices*

1. I modify assessments for students with an IEP or 504.

   Not Proficient-0

   Somewhat Proficient-0

   Proficient-10

   Very Proficient-15

   Extremely Proficient-18

Concept 6: *Equitable Assessment Practices*

2. I modify assessments for students learning English.

   Not Proficient-0

   Somewhat Proficient-2

   Proficient-9

   Very Proficient-17

   Extremely Proficient-15

3. I make trauma-informed decisions when considering classroom assessments.

   Not Proficient-2

   Somewhat Proficient-4

   Proficient-9

   Very Proficient-11

   Extremely Proficient-17

4. I design/choose assessments that are culturally relevant

   Not Proficient-2

   Somewhat Proficient-2

   Proficient-11

   Very Proficient-11

   Extremely Proficient-17

Figure 40

Concept 6: Question 5 with Sampling of Student Responses

CONCEPT 6: EQUITABLE ASSESSMENT PRACTICES

Concept 6: *Equitable Assessment Practices*

5. I consider personal biases that may influence classroom assessments.

   Not Proficient-1

   Somewhat Proficient-1

   Proficient-7

   Very Proficient-13

   Extremely Proficient-21

Concept 6: *Equitable Assessment Practices*

- I was given the opportunity to work with a variety of students during my student teaching, from AP proficient students to students that had English specific IEP's or other students with learning disabilities. This made me have to constantly assess my students progress and make adjustments for them based off of their skills.

- I try to use various levels of blooms taxonomy to help differentiate assessments for all of my students. Also as an example of trauma informed assessments, I had a student that struggled with an eating disorder. Because of this she really struggled with the nutrition unit. I had her write a 2 page paper on the importance of fitness and nutrition to allow her to demonstrate her understanding, but also avoid the subjects that have caused trauma.

- I worked especially hard with ELL students in my classes. I felt that I made personal connections and found at least some examples and material that were culturally relevant to them. I worked very hard to consider their unique learning needs. My literacy class was invaluable for helping learn how to work with these students.

within the open-ended question. One teacher-candidate said,

> …as an example of trauma-informed assessments, I had a student that struggled with an eating disorder. Because of this she really struggled with the nutrition unit. I had her write a…paper on the importance of fitness and nutrition to allow her to demonstrate her understanding, but also avoid the subject that caused trauma (Figure 40).

Another teacher-candidate explained:

> I worked especially hard with ELL students in my classes. I felt that I made personal connections and found…examples and materials that were culturally relevant to them. I…consider[ed] their unique learning needs. My literacy class was invaluable for helping learn how to work with these students (Figure 40).

Within Concept 8, Community and Feedback, teacher candidates' perceptions of their knowledge and skills were measured in the following area: assisting students and guardians to interpret classroom assessment data, providing corrective and supportive feedback to students, planning ways for students to respond to feedback, creating an environment where students view assessments as a positive way to check for progress, and collaborating with peers on common assessments.

This conceptual area demonstrates an awareness of the school community. It also links feedback, identified by research as a critical factor of successful classroom assessments, to TPP evaluation (Ahsan, 2018; Birenbaum et al., 2015; Garrison & Ehringhaus, 2011). Again, teacher candidates provided details within the open-ended question that may be helpful in understanding the current program and future changes that may improve the program. One student commented, "…Most of my students responded very well to written feedback and check-ins were done throughout the semester with those students that struggle" (see Figure 41). Another student explained, "This is an area that I could work on. I do try to comment on student work that is

Figure 41

Concept 8: Survey Questions

CONCEPT 8:-COMMUNITY AND FEEDBACK

Concept 8:-Community and Feedback

1. I assist students and/or guardians in interpreting classroom assessment data.

Not Proficient-3

Somewhat Proficient-0

Proficient-20

Very Proficient-10

Extremely Proficient-10

Concept 8:-Community and Feedback

2. I provide corrective and supportive feedback to students, in connection with classroom assessments.

Not Proficient-0

Somewhat Proficient-0

Proficient-13

Very Proficient-16

Extremely Proficient-14

3. I plan ways for students to respond to feedback provided on classroom assessments.

Not Proficient-

Somewhat Proficient-5

Proficient-8

Very Proficient-12

Extremely Proficient-18

4. I create an environment where students view assessments as a positive way to check for academic progress.

Not Proficient-0

Somewhat Proficient-0

Proficient-12

Very Proficient-15

Extremely Proficient-16

submitted, and if there is something they could change then I say so…I could be better about writing on rubrics and giving more constructive feedback" (see Figure 42).

Democratic Evaluation and Integrated Concept Mapping

While planning for program improvement is beyond the scope of this study, it is interesting to consider how the survey data collected through this evaluation may contribute to an examination of the curriculum within TPP. Course content could be considered and modified to improve courses within the program and provide more opportunities for students to practice the classroom assessment concepts identified.

Figure 42

Concept 8: Survey Questions with Student Responses

CONCEPT 8:-COMMUNITY AND FEEDBACK

Concept 8:-*Community and Feedback*

5. I collaborate with peers on common assessments.

   Not Proficient-1

   Somewhat Proficient-3

   Proficient-14

   Very Proficient-11

   Extremely Proficient-14

Concept 8:-*Community and Feedback*

- Most of my students responded very well to written feedback and check-in's were done throughout the semester with those students that struggle

- This is an area that I could work on. I do try to comment on student work that is submitted, and if they is something they could change then I say so. Students would sometimes make those changes and resubmit. But I could be better about writing on rubrics and giving more constructive feedback.

- My cooperating teacher used a grading conference to assess student learning, and this gave the students ample time to take into account feedback from the teacher as well as each other in the class and progress their learning in that regard up until the very end of the quarter.

By conducting this program evaluation using integrated concept mapping, I found it is possible for stakeholders to carry out a fairly complex grass-roots program evaluation with a high rate of participation and a strong commitment to the evaluation process. I also found the evaluation has the potential to impact the program through consideration of the data collected. For example, the findings could be used by instructors as a checklist to map out where exactly the concepts are taught within the program. This may take place in a formal setting such as a faculty meeting where the courses are matched with the concepts taught, along with the materials, experiences, and assignments connected to the concepts. This level of examination may extend the program evaluation and be used in concert with the survey results and the integrated concept mapping findings. Furthermore, the evaluation data and related findings could also be used to

consider the need for additional specific, assessment courses within the program, including courses on assessment design and adaptation.

Axial Coding Paradigm

Using grounded theory (Corbin & Strauss, 2007), I was able to uncover how democratic evaluation methods affect stakeholder commitment to the evaluation process. The Axial Coding Paradigm, presented in Ch IV demonstrates how stakeholder commitment to evaluation is strengthened through a democratic approach to TPP evaluation. Furthermore, I was able to demonstrate how these phenomena contribute to three intervening occurrences including: taking ownership of the process, seeking a deeper understanding, and (closer to the end of the process) considering the impacts. These categories were identified through open-coding and represent stakeholder actions and experiences during the integrated concept mapping process. Uncovering these phenomena explains why democratic methods of accountability may have the potential to contribute to impactful and relevant TPP evaluation. This method may also help to establish a unified understanding of how stakeholders work together when they share collective expertise. This study contributes to a greater understanding of the dynamics of democratic methods within the context of TPP evaluation, and possibly additional areas of education where the collective expertise of coworkers exists within independent spheres, and they are willing to collaborate for evaluation.

Classroom Assessment

While integrated concept mapping can be used to for complex evaluations that

include many areas of evaluation, I chose to limit the scope of this study and the evaluation to one area, classroom assessments. I chose this area for evaluation due to its potential impact on K-12 learning (Stiggins, 1999; Volante & Fazio, 2007) and because research suggests it is an area that needs improvement within TPPs (DeLuca & Bellara, 2013; Popham, 2011).

The topic of classroom assessment is not often included within current TPP evaluation and has not received enough attention to garner the needed improvements (Popham, 2011). Though this grass-roots evaluation process, the topic was identified and discussed by a few members of the TPP and presented to the remaining members of the program, who agreed this was an important focus and an area we do not know much about within the program. As a result of this dialogue, the initiative was taken up by the stakeholders. By employing this method, we found teacher candidates within the program had a high level of confidence in many areas of classroom assessment identified by the faculty and administration. We also found some areas for improvement. This topic is an example of a blind spot that may exist within a program that is not frequently addressed through current evaluation methods. Integrated concept mapping, with its inherently democratic processes, allowed for the identification and consideration of this area.

Themes Identified through Interpretivism

In this section, I will share thematic considerations resulting from of my experiences as a participant stakeholder, as the evaluation facilitator (a designated role within integrated concept mapping), and as a researcher for this study. I decided to

include this section after I read through the study results and notes, and after reflecting on my experiences facilitating the program evaluation. My goal was to better understand the evaluation process and how it played out.

Interpretivism is a research method that can be used to organize and make sense of data collected and associated experiences relevant to the phenomena. Many qualitative methods involve observation, interaction with subjects, and asking questions (Duffy et al., 2021). Within interpretivism, the researcher also considers how study participants responded and interacted, often after the study is completed. The researcher interprets the actions and intentions of the participants. While some of the themes identified through interpretivism may be part of the phenomena identified through the Axial Coding Paradigm and other qualitative methods, they may not be directly identified through the open-coding process. Instead, they may be themes identified throughout the study (Duffy et al., 2021).

Stakeholder Participation

The first research question for this study sought to uncover how integrated concept mapping may contribute to the evaluation of a TPP in the area of classroom assessment outcomes. Through the process of integrated concept mapping, the faculty and administration within this TPP were able to work together collaboratively to identify an area for evaluation, set parameters for the topic, come to a consensus about the areas of expected proficiency, create a tool to measure teacher-candidates perceptions of their skills and knowledge, and reflect on the data collected.

Stakeholders participated in the process and followed through with the evaluation,

despite being new to the process and working under a deadline. Teacher-candidates, who made up the third stakeholder group, were willing to participate at a rate high which allowed the program evaluation to take place. Without their participation the evaluation tool would have stalled. Teacher-candidates also provided optional examples and details about their experiences within the classroom, at a high rate.

In terms of how intergraded concept mapping may have the potential to shift TPP evaluation toward a more democratic approach to accountability, stakeholder participation throughout the process demonstrated a high level of interest in the topic and evaluation method.

Investment in the Process

Stakeholder participants not only participated but were invested in the Integrated Concept Mapping Process. The statements contributed by stakeholders in the initial stage in the process represented a variety of areas related to classroom assessment. While some statements were removed, because they were unrelated to the agreed-upon definition of classroom assessment, stakeholder participants within the advisory committee took the initiative to identify the "off-topic" statements. Lackluster participation or attitudes may have resulted in misrepresentation or underrepresentation of the range of classroom assessment concepts teacher-candidates encounter within the program.

Stakeholders continued to be invested in the process throughout the evaluation. For example, fourteen of fifteen stakeholder participants took part in the meetings and voted for the scenario that would be used for the survey development. In another example, the advisory committee was concerned with presenting the best possible

scenarios to the larger stakeholder group. Finally, the efforts of the larger group to choose the right scenario for the survey development, demonstrated a sustained commitment to the evaluation process and a growing interest in accurately representing the conceptual areas identified to the teacher candidates.

The process was essentially a grass-roots approach to evaluation, requiring a commitment to transparency and a level of trust between the facilitator, advisory committee, and the stakeholders. For example, the advisory committee wanted to make sure there was a complete understanding of the definition of classroom assessment,

> [The slide] you gave us…cleared some things up for me, because you could go all over the place with [classroom assessment]. We are trying to see what people think, and what classroom assessment is. I would give a definition, so they know.

The committee also wanted to make sure the process was clear for those taking part in the program evaluation, "How much are they going to know about the whole process? [I am] just thinking about that and how much is helpful for people to understand…" (Advisory Committee Member 2). Advisory Committee Member 3 was concerned with making the direction clear and specific "…will there be any kind of direction…so they know exactly what they are doing." Advisory Committee Member 1 agreed and summarized their discussion, "So [we should] be more explicit? I think [that is] what I'm hearing…. I appreciate the feedback and [we can] make sure that happens tomorrow."

These statements are an example of how the committee worked through each aspect of the evaluation within their purview. Regarding shifting evaluation toward a more democratic approach to accountability, each aspect of the evaluation was considered, designed, facilitated, and carried out solely by the faculty and administration

within the TPP.

User-Friendly Software

While it is possible to complete a program evaluation using the Integrated Concept Mapping method without the use of Concept Systems Inc. software, the "user-friendly" interface contributed greatly to the speed of the evaluation process, stakeholder motivation, and the ability of the stakeholders within this TPP to carry out the evaluation. I am not sure if there would have been the same level of commitment, without the software. Additionally, it would have been extremely difficult for me to carry out the evaluation without the software, as I am not an expert in the method. My experiences carrying out this study helped me to understand the impact of user-friendly software on research, and more specifically as part of the integrated concept mapping method.

A Community of Experts

The willingness of stakeholders to share their expertise throughout the evaluation process was a key factor in being able to carry out the evaluation of this program. If stakeholders had been unwilling to collaborate, this grass-roots program evaluation would have been difficult at best. This community of experts came together to discuss and debate the topics and came to consensus to define the parameters of the study and to choose the focus prompt, statements for ratings, and cluster scenarios.

In terms of shifting evaluation toward a more democratic approach to accountability, the collective expertise of stakeholder participants contributed to rich discussion, deliberation, and thoughtful questions, each of which were critical aspects of

this evaluation. Furthermore, it was the differences among stakeholders' expertise and roles within the program that provided the needed checks to keep the balance of power from shifting to any one group or individual throughout the evaluation process.

Pushing Back

There were a couple of concerns about democratic forms of evaluation voiced at the end of the process in response to the semistructured interview questions. I included the concerns within Category 3 of the Axial Coding Paradigm, Seeking a Deeper Understanding. However, there may be value in reflecting on the specific concerns within this section as well.

Two of the participant stakeholders within the administrator group briefly mentioned concerns about using more democratic methods of accountability. Their concerns related to members of an organization possibly becoming "frustrated" if different or opposing views were voiced. Administrator one stated, "…I could see how this could get potentially messy because…we call something a democratic process, but that doesn't mean that people can't be frustrated by the outcome." The same individual mentioned "…if you really had a split group…then you could sort of disenfranchising some individuals because the outcome doesn't match what their input was and therefore, they don't see themselves…. It's kind of like my vote didn't count." Administrator two followed this statement up with a question, "…what happens when we disagree" and "…what if someone's ideas are not included?"

These questions and concerns represented a very small number of stakeholder responses and were only voiced by two stakeholder administrators. The same

administrators also voiced many positive attributes of the process. For example,

administrator one stated the process was an

> …opportunity to identify any gaps where either professors had one idea about what they thought they should be teaching [and] students and administrators had another…. I thought that was really interesting and something that could contribute to, you know, our evaluation of our program.

> Administrator two stated, "It did feel…like there was some ownership there

because I can see where each of the things we put in…I wanted to see what my

colleagues had to say-we can become isolated" and "This process [brings] people

together that generally operate in isolation." Administrator three added,

> Decisions are often made with partial input…having a protocol process like this that has multiple points of review and reflection and has a formal way of involving people you get better thinking [and] you get better buy in and it's much more defensible.

One more consideration within this section was related to "outliers." One faculty member

asked questions about ideas that may not be included in the process or may have been

overlooked. Faculty Member three stated "…from what I understand of the process…it

does a good job of capturing consensus views; does the process bring out…unique views

and capture outliers?" Other participant stakeholders addressed to this question with the

following, "…I can see each of the things we put in...I can see where we are represented

in each area," stated Administrator two. Administrator three mentioned, "At the school

level decisions are often made with partial input, this [process] has multiple points of

review and reflection and…a formal way of involving people." From my perspective as

the evaluation "Facilitator," a member of the "advisory committee," and a member of the

faculty I can trace the process and see how unique views and outliers were included. If

ideas were consistent with the topic, and more specifically, the agreed upon definition of "classroom assessment," they seemed to be included. This issue was also addressed by Kane and Trochim (2007) who explain decisions about what is included within evaluation is up to the stakeholders, including the advisory committee. Additional studies could focus on this question of outliers, which may be answered in connection with questions about transparency throughout the process.

<div align="center">**Implications and Recommendations**</div>

Teacher Preparation Programs

Regarding the research questions for this study, the data collected suggests it is possible to use integrated concept mapping as a method for evaluating a TPP in the area of classroom assessment. Furthermore, the evidence collected and presented suggests it may help shift the focus of evaluation toward a more democratic approach to accountability. The integrated concept mapping data, the evaluation criteria and survey, the data derived from the survey, and the data resulting from the semistructured interviews all suggest a valid and reliable TPP evaluation can result from integrated concept mapping. Furthermore, this data also demonstrated integrated concept mapping is an evaluation tool that promotes and relies on democratic accountability methods.

This study also implies rich data about classroom assessment can be gathered through integrated concept mapping and the democratic methods inherent within its design. For example, the variety of assessment considerations within the eight conceptual areas identified demonstrated alignment with current research and reflect democratic

considerations of equity and community. Likewise, inclusion of teacher candidates as stakeholders in the study provided valuable data about teacher-candidates perceptions of their skills and knowledge, which was furthered through their responses to the open-ended questions within the survey. Expanded use of integrated concept mapping within TPPs evaluation, and possibly other areas of education, may strengthen understanding about the method and its use within the field.

K-12 Education

One of the major purposes of TPP evaluation is its potential to improve K-12 education. Research suggests improvements within TPPs translate into improvements within local schools where teachers graduating from programs find employment (McDiarmid, 2019). Program improvement is often stalled due to the continued use of "accountability era" evaluation methods which can still be found within the current TPP evaluation (Cochran-Smith, et al., 2018). Findings in this study may lead to more attempts at evaluation that forward democratic methods for accountability which may, in turn, lead to program improvement. Additional studies focusing on TPP evaluation and democratic accountability methods are needed to strengthen the findings of this study and more completely understand how democratic methods of evaluation may impact TPPs.

The Field of Education

Integrated concept mapping is currently used within the fields of social science and nursing for program evaluation and planning (Abrahams, 2004; Bedi, 2004). This method has been used within the field of education on a limited basis, mainly by course

instructors and teachers to evaluate modes of instruction within their classrooms (Davis, 2003; Edwards, 2002). It has also been used within higher education for that same purpose. Because integrated concept mapping has resulted in successful evaluation and planning in other fields and based on the results of this study, the method may be a good fit for the field of education on a larger scale.

Classroom Assessment Considerations

Future studies should consider spending more time presenting current research about classroom assessment to stakeholders. While the collective expertise of the stakeholders captured many details within the literature, there were some areas related to classroom assessment and program design that were not considered and which may be helpful when considering program improvement (DeLuca & Bellara, 2013; Popham, 2011; Stiggins, 1999; Volante & Fazio, 2007). For example, the faculty did not identify areas of classroom assessment related to educational psychology topics or specific assessment design considerations, although they did mention assessment design more generally. One reason this area may have resulted in general statements is the program does not currently have a specific assessment design course or instructor. Another reason may have been the recent departure of the tenured educational psychology teacher, a position that remained open throughout this study.

Comparing Results

Finally, it may be useful to compare the results of this evaluation method to other evaluation methods. For example, comparing classroom assessment data gathered

through performance assessments or evaluations by student-teacher supervisors to results from integrated concept mapping. This type of comparison may provide a meaningful understanding of teacher-candidates proficiency in the area of classroom assessment.

Limitations

The most important limitation of this study was its scope. While it was possible to identify areas where integrated concept mapping could contribute to TPP evaluation, the scope and breadth of the data collected were at times overwhelming. For example, a team of researchers may have been able to conduct an analysis of the open-ended survey questions and additional concept maps that may have provided additional evidence related to the purpose of the study and the research questions. However, this data may be useful within future studies with research questions.

Another limitation was the number of stakeholder groups included in the study. The stakeholders included in this study asked about the usefulness of expanding future program evaluation to include more varied groups. This evaluation was limited to three internal stakeholder groups due to time, the feasibility of the study, and concern about the scope of the study. Including external stakeholder groups may provide additional understanding about classroom assessment, an expanded understanding of democratic accountability within the evaluation, and additional data related to the usefulness of integrated concept mapping for TPP evaluation.

Still another limitation was the survey design. As mentioned in Chapter IV, there were additional changes the advisory committee would have made to the survey. Some

design changes that may have resulted in a better evaluation tool for the program. For example, organizing the questions according to Bloom's Taxonomy and additional editing of questions and directions would be considered (Armstrong, 2010). Additionally, the first conceptual area had similar questions and concepts that were covered in other conceptional areas, so it may have been possible to eliminate that section of the survey.

Finally, this study focused on one TPP, and it is difficult to generalize the significance beyond this specific program. Potentially, other TPPs could use the same method for program evaluation which may impact TPP evaluation, classroom assessment, and democratic methods of accountability on a larger scale.

Conclusion

Current evaluation methods used by TPP largely ignore or minimize community factors, "Strong" equity considerations, and stakeholder input (Cochran-Smith & Reagan 2022). Democratic methods for program evaluation may include more opportunities to focus on these issues. In this study, integrated concept mapping depended on a diverse group of stakeholders within the organization to work together to evaluate the program by identifying local factors. Using this method of evaluation, I found it was possible for members of a TPP to complete a meaningful democratic program evaluation This method of evaluation may hold the potential to further the ideals of democratic education put forth by John Dewey and others through a grass-roots approach to evaluation as well as more recent notions of democratic education which emphasize the need for more democratic methods for accountability (Cochran-Smith et al., 2018; Kira, 2019).

REFERENCES

Abrahams, D. (2004). *Technology adoption in higher education: A framework for identifying and prioritizing issues and barriers to adoption.* [Doctoral dissertation, Cornell University].

Ahsan, S. (2018), *Teacher education and professional development on classroom assessment in Bangladesh: Exploring policy and practice through vertical case study* [Doctoral dissertation, University of Massachusetts Amherst]. https://scholarworks.umass.edu/dissertations_2/1319

Apple, M. W., & Beane, J. A. (2007). *Democratic schools: Lessons in powerful education*. Heinemann.

Armstrong, P. (2010). *Bloom's taxonomy*. Vanderbilt University Center for Teaching. https://cft.vanderbilt.edu/guides-sub-pages/blooms-taxonomy/

Association for Advancing Quality in Educator Preparation (2022). *About us.* https://aaqep.org/about-us

Baker, E., Barton, P., Darling-Hammond, L., Haertel, E., Ladd, H., Linn, R., Ravich, R., Rothstein, R., Shavelson, R., L, Shepard, L., (2010). *Problems with the use of student test scores to evaluate teachers*. Washington, DC: The Economic Policy Institute.

Ball, D., & McDermid, G. (1990). The subject matter preparation of teachers. In R. Houston (Ed.), *Handbook of research on teacher education* (pp. 437–449). Macmillan.

Banta, T.W. (2007). Can assessment for accountability complement assessment for improvement? *Peer Review. Spring 2007*, 9-12.

Barber, B. R. (2009). *Strong democracy: Participatory politics for a new age*. University of California Press.

Bastian, K., Fortner, K., Chapman, A., Fleener, J., McIntyre, E., & Partiarcha, L. (2015). *Data sharing to drive the improvement of teacher preparation programs*. Chapel Hill, NC: Education Policy Initiative at Carolina.

Bastian, K., Henry, G., Pan, Y., & Lys, D. E. (2015). *Evaluating a pilot of the teacher performance assessment: The construct validity, reliability, and predictive validity of local scores*. Chapel Hill, NC: Education Policy Initiative at Carolina.

Bastian, K., Patterson, T., & Pan, Y. (2015). *Teacher preparation program effectiveness report*. Chapel Hill, NC: Education Policy Initiative at Carolina.

Becker, B. (n.d.). *Logic models and theories of action enhancing program performance and accountability*. Office of Public Charter School Financing and Support (OPCSFS). https://osse.dc.gov/.

Bedi, R. (2004). *Concept mapping the client's perspective on counseling alliance formation.*[Doctoral dissertation, University of British Colombia, Canada].

Bennett, R. E. (2011). Formative assessment: a critical review. *Assessment in Education: Principles, Policy & Practice, 18*(1), 5-25. DOI: 10.1080/0969594X.2010.513678

Birenbaum, M., DeLuca, C., Lorna, E., Heritage, M., Klenowski, V., Looney, A., Smith, K., Timperley, H., Volante, L., & Wyatt-Smith, C. (2015). Policy international trends in the implementation of assessment for learning: Implications for policy and practice. *Policy Futures in Education, 13*(1) 117–140.

Black, P., & Wiliam, D. (2004). The formative purpose: assessment must first promote learning. In M. Wilson (Ed.), *Towards coherence between classroom assessment and accountability* (pp. 20–49). University of Chicago Press.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, *21*(1), 5-31.

Brady, M. P., Miller, K. M., McCormick, J., & Heiser, L. A. (2018). A rational and manageable value-added model for teacher preparation programs. *Educational Policy*, *32*(5), 728–750.

Braun, H. (2005). *Using student progress to evaluate teachers: A primer on value-added models*. Educational testing Service.

Brookhart, S. M. (2001, March 21-23). *The standards and classroom assessment research* [Paper presentation]. The Symposium on Role of the Standards for Teacher Competence in Educational Assessment of Students in Developing Quality Teacher Assessors, Dallas, TX. https://archive.org/stream/ERIC_ED451189/ERIC_ED451189_djvu.txt

Brown, G. T. L. (2004). Teachers' conceptions of assessment: implications for policy and professional development. *Assessment in Education, 11*(3), 301-318, DOI: 10.1080/0969594042000304609

Brown, J., & Duguid, P. (2000). *The social life of information*. Harvard Business School Press.

Brownson, J. (2018). *Teaching and learning in the co-teaching model: Analyzing the cooperating teacher/teacher candidate co-planning dialogue* [Doctoral Dissertation, University of Wisconsin, Milwaukee].

Campbell, C., & Evans, J. A. (2010). Investigation of pre-service teachers' classroom assessment practices during student teaching. *The Journal of Educational Research. 93*(6), 350-355.

Campbell, C., Murphy, J. A., & Holt, J. K. (2002, October). *Psychometric analysis of an assessment literacy instrument: Applicability to preservice teachers* [Paper presentation]. Mid-Western Educational Research Association, Columbus, OH.

Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes? A cross-state analysis. *Educational Evaluation and Policy Analysis, 24,* 305–331.

Charmaz, K. (2009). *Constructing grounded theory: A practical guide through qualitative analysis.* Sage Publishing.

Chen, P. (2005). Teacher candidates' assessment literacy. *Academic Exchange Quarterly, 62*, 62-66.

Chen, H. T., Morosanu, L., Turner, N., & Pan, H.-L. W. (2018). Using logic models and the action model/change model schema in planning the learning community program: A comparative case study. *Canadian Journal of Program Evaluation*, *33*(1), 49-68–68. https://doi-org.dist.lib.usu.edu/10.3138/cjpe.42116

Chizhik, E.W., Chizhik, A.W., Close, C., & Gallego, M. (2018), "Developing student teachers' teaching self-efficacy through Shared Mentoring in Learning Environments (SMILE)", International Journal of Mentoring and Coaching in Education, Vol. 7 No. 1, pp. 35-53. https://doi.org/10.1108/IJMCE-02-2017-0014

Ciccone, J. K. (2019, June 27). *Education deans for justice and equity convene at Ohio State*. https://ehe.osu.edu/news/listing/education-deans-justice-and-equity-convene-ohio-state

Clandinin, D. & Connelly M., (1998). Stories to live by: Narrative understandings of school reform. *Curriculum Inquiry, 28*(2), 149-164, DOI: 10.1111/0362-6784.00082

Cochran-Smith, M., Baker, M., Burton, S., Chang, W.-C., Cummings Carney, M., Fernández, M. B., Stringer Keefe, E., Miller, A. F., & Sánchez, J. G. (2017). The accountability era in US teacher education: looking back, looking forward. European Journal of Teacher Education, 40(5), 572–588. https://doi-org.dist.lib.usu.edu/10.1080/02619768.2017.1385061

Cochran-Smith, M., & Boston College Evidence Team. (2009). "Re-culturing" teacher education: Inquiry, evidence, and action. *Journal of Teacher Education*, *60*, 458–468.

Cochran-Smith, M., Carney, M. C., Keefe, E. S., Burton, S., Chang, W.-C., Fernández, M. B., Miller, A. F., Sánchez, J. G., & Baker, M. (2018). *Reclaiming accountability in teacher education*. Teachers College Press.

Cochran-Smith, M., & Reagan, E. M. (2022). Centering Equity in Teacher Education Evaluation: From Principles to Transformative Enactment. *Journal of Teacher Education*, *73*(5), 449. https://doi-org.dist.lib.usu.edu/10.1177/00224871221123728

Cochran-Smith, M., Stern, R., Sánchez, J. G., Miller, A., Keefe, E. S., Fernández, M. B., Chang, W.-C., Carney, M. C., Burton, S., Baker, M., & University of Colorado at Boulder, N. E. P. C. (2016). *Holding teacher preparation accountable: A review of claims and evidence*. National Education Policy Center.

Coggshall, J. G., Bivona, L., Reschly, D. J., & National Comprehensive Center for Teacher Quality. (2012). Evaluating the Effectiveness of Teacher Preparation Programs for Support and Accountability. Research & Policy Brief. In *National Comprehensive Center for Teacher Quality*. National Comprehensive Center for Teacher Quality.

Concept Systems Inc. (2022, March 21). *Group knowledge is power*. https://www.conceptsystems.com/home

Creswell, J. W. (2008). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage Publications. doi: 10.1016/j.math.2010.09.003

Creswell, J. W., & Plano Clark, V. L. (2018) *Designing and conducting mixed method research.* Sage Publications.

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

Cuban, L. (2004). Looking through the rearview mirror at school accountability. In K. Sirotnik (Ed.), *Holding accountability accountable* (pp. 18–34). Teachers College Press.

Darling-Hammond, L. (2000). *Teacher quality and student achievement: A review of state policy evidence*. http://epaa.asu.edu/epaa/v8n1/

Darling-Hammond, L., Amrien-Beardsley, A., Haertel, E., & Rothstein, J. (2011). *Getting teacher evaluation right: A background paper for policy makers*. http://www.aera.net/Portals/38/docs/New%20Logo%20Research%20on%20Teacher%20Evaluation%20AERA-NAE%20Briefing.pdf

Davis, T. (2003). *Validity of concept mapping for assessing cultural competence in children's mental health systems of care: A comparison of theoretical and community conceptualizations.* [Doctoral dissertation, University of Texas, Austin].

DeLuca, C., & Bellara, A. (2013). The current state of assessment education: aligning policy, standards, and teacher education curriculum. Journal of Teacher Education, 64(4), 356.

DeLuca, C., Chavez, T., Bellara, A., & Cao, C. (2013). Pedagogies for preservice assessment education: Supporting teacher candidates' assessment literacy development. *Teacher Educator, 48*, 128-142.

DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice, 17*, 419-438.

DeLuca, C., Klinger, D. A., Searle, M., & Shulha, L. M. (2010). Developing a curriculum for assessment education. *Assessment Matters, 2*, 20-42.

DeLuca, C., & Volante, L. (2016). Assessment for learning in teacher education programs: Navigating the juxtaposition of theory and praxis. *Journal of the International Society for Teacher Education, 20*(1), 19–31.

Denzin, N., & Giardina, M. (2018). *Qualitative inquiry in the public sphere*. Taylor and Francis.

Dewey, J. (1916). *Democracy and education: An introduction to the philosophy of education*. MacMillan.

Duffy, L. N., Fernandez, M., & Sène-Harper, A. (2021). Digging Deeper: Engaging in Reflexivity in Interpretivist-Constructivist and Critical Leisure Research. Leisure Sciences, 43(3/4), 448–466.

Earley, P. (2001). *Title II requirements for schools, colleges, and departments of education*. http://www.ericdigests.org/2002-3/title.htm.

Edwards, A. (2002*). Mobilizing the village: Collaborating with parents and the community to increase parental involvement in a high school of the arts.* [Doctoral dissertation, Georgia State University].

Ellis, S., & Smith, V. (2017). Assessment, teacher education and the emergence of professional expertise. *Literacy, 51*(2), 84–93.

Engel, M. (2000). *The struggle for control of public education: Market ideology vs. Democratic values*. Temple UP.

Eurich, N., & Wade, B. (1986). *Corporate classrooms: The learning business* (A Carnegie Foundation special report). San Francisco, CA: Jossey-Bass.

Feuer, M., Floden, R., Chudowsky, N., & Ahn, J. (2013). *Evaluation of teacher preparation programs: Purposes, methods, and policy Options*. Washington, DC: National Academy of Education.

Finnigan, K., & Gross, B. (2007). Do accountability policy sanctions influence teacher motivation? Lessons from Chicago's low-performing schools. *American Educational Research Journal*, *44*, 594–629.

Fleener, J., & Exner, P. (2011). Dimensions of teacher education accountability: A Louisiana perspective of value-added. In P. D. Imig & N. Michelli Earley (Eds.), *Teacher education policy in the United States* (pp. 26–43). Routledge.

Fraser, J. W. (2007). *Preparing America's teachers: A history.* Teachers College Press.

Garrison, C., & Ehringhaus, M. (2011). *Formative and summative assessments in the classroom*. http://ccti.colfinder.org/sites/default/files/formative_and_summative_ assessment_ in_the_classroom.pdf

Graham, P. (2005). Classroom-based assessment: Changing knowledge and practice through preservice teacher education, *Teaching and Teacher Education, 21* (6), 607-621.

Greenberg, J., & Walsh, K. (2010). Ed school essentials: Evaluating the fundamentals of teacher training programs in Texas. http://www.nctq.org/dmsView/Ed_School_ Essentials_Training_Programs_Texas_Executive_Summary_pdf.

Greenberg, J., & Walsh, K. (2012). *What teacher preparation programs teach about K-12 assessment: A review*. Washington DC: National Council on Teacher Quality.

Groupwisdom. (2022, August 15). *Analysis and features*. https://groupwisdom.com/ groupwisdom?gclid=EAIaIQobChMIzamF5oGz-wIVARXUAR2RdArLEA AYASAAEgIyFfD_BwE

Guskey, T. R. (2003). How classroom assessments improve learning. *Educational Leadership, 60*(5), 6-11.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge. doi: 10.1007/s11159-011-9198-8

Hattie, J., Brown, G. T., & Keegan, P. (2005). A national teacher-managed, curriculum-based assessment system: Assessment tools for teaching & learning. *International Journal of Learning,10*, 770–778.

Hattie, J., & Yates, G. (2013). *Understanding learning: Lessons for learning, teaching and research*. https://research.acer.edu.au/cgi/viewcontent.cgi?article=1207&context=research_conference

Hawk, P. P., & Schmidt, M. W. (1989). Teacher preparation: A comparison of traditional and alternative programs. *Journal of Teacher Education, 40*(5), 53-58.

Horne, C. (2021, November 1). *Teacher data literacy: It's about time*. Data Quality Campaign. https://dataqualitycampaign.org/resource/teacher-data-literacy-time/

Imig, D., & Imig, S. (2008). From traditional certification to competitive certification: A twenty-five year retrospective. In M. Cochran-Smith, S. Feiman-Nemser, D. McIntyre, & K. Demers (Eds.), *Handbook of research on teacher education* (3rd ed., pp. 886–907). Routledge.

Ivankova, N. (2002*), Students' persistence in the university of Nebraska-Lincoln distribuited doctoral program in educational administration: A mixed methods study* [Doctoral dissertation, University of Nebraska]. https://researcher20.com/wp-content/uploads/2017/11/Sample-MIxed-Methods-Proposal-Sage-Pub-compressed.pdf

Johnson, R. L., & Morgan, G. B. (2016). *Survey scales: A guide to development, analysis, and reporting.* Gilford Press.

Kamenetz, A. (2018, April 29). *What 'A nation at risk' got wrong, and right, about U.S. schools*. https://www.npr.org/sections/ed/2018/04/29/604986823/what-a-nation-at-risk-got-wrong-and-right-about-u-s-schools.

Kane, M., & Trochim, W. (2007). *Concept mapping for planning and evaluation.* Sage Publications.

Kamens, M. W. (2007). Learning about co-teaching: A collaborative student teaching experience for preservice teachers. *Teacher Education & Special Education, 30*(3), 155–166. https://doi-org.dist.lib.usu.edu/10.1177/088840640703000304

Kira, N. (2019). Dewey's democratic conception in education and democratic schooling: Lessons from the United States for Japan in a time of democracy in crisis. *Educational Studies in Japan: International Yearbook, 13,* 55-66.

Knoester, M. (2012). Democratic Education in Practice: Inside Mission Hill School. New York: Teachers College Press.

Koh, K. H. (2011). Improving teachers' assessment literacy through professional development. *Teaching Education, 22*(3), 255-276, DOI: 10.1080/10476210. 2011.593164

Kumashiro, K. (2015). Review of Proposed 2015 Federal Teacher Preparation Regulations. Boulder, CO: National Education Policy Center. Retrieved [November 2022] from http://nepc.colorado.edu/thinktank/review-proposed-teacher-preparation.

Loeb, S., Valant, J., & Kasman, M. (2011). Increasing choice in the market for schools: Recent reforms and their effects on student achievement. *National Tax Journal*, 64(1), 141-164.

Maclellan, E. (2004). Initial knowledge states about assessment. *Teaching and Teacher Education, 20*(5), 523-535. DOI: 10.1016/j.tate.2004.04.008

Mathison, S., & Rogers, P. (Eds.). (2007). Theory-based evaluation: Reflections ten years on. In *Enduring issues in evaluation: The 20th anniversary of the collaboration between NDE and Aea* (pp. 63–67). Wiley.

Marshall, C., & Rossman, G. B. (2011). *Designing qualitative research*. Thousand Oaks, CA: Sage Publications. doi: 10.4135/9781849208826

McDiarmid, G. W. (2019). Competing theories for improving teacher preparation programs: The case of North Carolina. *ECNU Review of Education, 2*(2), 117–136.

McDiarmid, G., & Caprino, K. (2017). *Lessons from the Teachers for a New Era project: Evidence and accountability in teacher education*. Routledge.

Meens, D. E., & Howe, K. R. (2015). NCLB and its wake: Bad news for democracy. *Teachers College Records, 117*(6), 1-44.

Meier, D. (2003). *In schools we trust: Creating communities of learning in an era of testing and standardization*. Beacon Press.

Mertens, D. M., & Wilson, A. T. (2019). *Program evaluation theory and practice: A comprehensive guide*. Guilford Press.

Mertler, C. A., & Campbell, C. S. (2004). *Assessing those who assess: Development of an instrument to measure teachers' assessment literacy* [Paper presentation]. Mid-Western Educational Research Association, Columbus, OH.

Noell, G., & Burns, J. (2006). Value added assessment of teacher preparation: An illustration of emerging technology. *Journal of Teacher Education*, *57*, 37–50.

Novak, J. D., & Caas, A. J. (2008). *The theory underlying concept maps and how to construct and use them*. http://cmap.ihmc.us/Publications/ResearchPapers/Theory UnderlyingConceptMaps.pdf

Peck, C., & McDonald, M. (2013). Creating "cultures of evidence" in teacher education: Context, policy, and practice in three high-data-use programs. *New Educator*, *9*, 12–28.

Peck, C., & McDonald, M. (2014). *What is a culture of evidence? How do you get one? And... Should you want one?* Teachers College Record. https://www.tcrecord.org

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice, 48*(1), 4-11, doi: 10.1080/00405840802577536 -

Popham, W. J. (2011). Assessment literacy overlooked: A teacher educator's confession. *The Teacher Educator, 46,* 265-273. doi: 10.1080/08878730.2011.605048

Reagan, E. M., Schram, T., McCurdy, K., Chang, T. H., & Evans, C. M. (2016). Politics of policy: Assessing the implementation, impact, and evolution of the Performance Assessment for California Teachers (PACT) and edTPA. *Education policy analysis archives*, *24*, 9-9.

Rickenbrode, R., Drake, G., Pomerance, L., Walsh, K., & National Council on Teacher Quality. (2018). 2018 Teacher Prep Review. In National Council on Teacher Quality. National Council on Teacher Quality.

Rossman, G. B., & Rallis, S. F. (2012*). Learning in the field: An introduction to qualitative research*. Sage Publications.

Saldana, J. (2021). *The coding manual for qualitative researchers.* Sage Publications.

Sanders, W. L., & Horn, S. P. (1994). The Tennessee Value-Added Assessment System (TVAAS): Mixed model methodology in educational assessment. *Journal of Personnel Evaluation in Education, 8,* 299–311.

Sanders, W. L., & Rivers, J. C. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. Knoxville, TN: University of Tennessee Value-Added Research and Assessment Center.

Schiller, E., Hayes, S., Nagle, K., IDEA Data Center, & Westat. (2020). *Using a theory of action to develop performance indicators to measure progress towards a SIMR*. IDEA Data Center.

Shepard, L., Hannaway, J., Baker, E., & National Academy of Education. (2009). *Standards, assessments, and accountability* [Education Policy White Paper]. National Academy of Education.

Slade, M. L., Burnham, T. J., Catalana, S. M., & Waters, T. (2019). The impact of reflective practice on teacher candidates' learning. *International Journal for the Scholarship of Teaching and Learni*ng, 13(2), 1-8.

SPSS, Inc.(n.d.). *Statistical package for the social sciences*. http://www.spss.com.hk/corpinfo/history.htm

State Collaborative on Reforming Education (SCORE). (2017). Policy Brief: Understanding TVAAS: Student growth data and Tennessee's Teacher Evaluation System. https://tnscore.org/wp-content/uploads/2018/10/TVAAS_PolicymakerBrief2017.pdf

Stewart, A. R., Scalzo, J. N., Merino, N., & Nilsen, K. (2015). Beyond the criteria: Evidence of teacher learning in a performance assessment. *Teacher Education Quarterly*, *42*(3), 33–58.

Stiggins, R. J. (1999). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice*, 23-27. doi: 10.1111/j.1745-3992.1999.tb00004.x

Stiggins, R., & Chappuis, J. (2005) Using student-involved classroom assessment to close achievement gaps. *Theory Into Practice, 44*(1), 11-18, DOI: 10.1207/s15430421tip4401_3.

Strauss, A., & Corbin, J. M. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Sage Publications, Inc.

Strauss, V. (2021, November 30). Analysis: 'A nation at risk' demanded education reform 35 years ago. here's how it's been bungled ever since. *The Washington Post.* https://www.washingtonpost.com/news/answer-sheet/wp/2018/04/26/the-landmark-a-nation-at-risk-called-for-education-reform-35-years-ago-heres-how-it-was-bungled/.

Sykes, G., Schneider, B., Plank, D. N., & American Educational Research Association. (2009). *Handbook of education policy research*. Routledge, Taylor & Francis Group.

Trochim, W. M. K., & Linton, R. (1986). Conceptualization for planning and evaluation. Evaluation and Program Planning, 9(4), 289-308–308. https://doi-org.dist.lib.usu.edu/10.1016/0149-7189(86)90044-3

Trotter, S. (2019). *Enrollment numbers show UVU serves a diverse array of learners*. https://www.uvu.edu/news/2019/10/10022019_enrollment.html

Volante, L., & Fazio, X. (2007). Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education, 30*(3), 749-770.

Vollstedt, M., & Rezat, S. (2019). An introduction to grounded theory with a special focus on axial coding and the coding paradigm. In G. Kaiser & N. Presmeg (Eds.) *Compendium for early career researchers in mathematics education* (pp. 81-100). Springer. https://doi.org/10.1007/978-3-030-15636-7_4.

Walsh, K. (2006). *Teacher preparation: Coming up empty*. https://edexcellence.net/publications/teachered.html

Walsh, K. (2013). 21st century teacher education. *Education Next*, *13*(3), 18–24.

Wenglinsky, H. (2002). How schools matter: The link between teacher classroom practices and students academic performance. Education Policy Analysis Archives, 10, 29p.

Wheeldon, J. (2010). Mapping mixed methods research: Methods, measures, and meaning. *Journal of Mixed Methods Research*, *4*(2), 87–102.

William, D. (2011). What is assessment for learning? *Studies in Educational Evaluation. 37,* 3–14.

Wininger, S. R. (2005). Using your tests to teach: Formative summative assessment. *Teaching of Psychology, 32*, 164–166.

Worrell, F. C., Brabeck, M. M., Dwyer, C. A., Geisinger, K. F., Marx, R. W., Noell, G. H., & Pianta, R. C. (2014). *Assessing and evaluating teacher education programs.* Washington, DC: American Psychological Association.

Wylie, E. C. (2017). Winsight™ Assessment System: Preliminary theory of action (Research Report No. RR-17-26). Princeton, NJ: Educational Testing Service. https://doi.org/10.1002/ets2.12155

Zeichner, K. (2011). Assessing state and federal policies to evaluate the quality of teacher preparation programs. In P. Earley, D. Imig, & N. Michelli (Eds.), *Teacher education policy in the United States: Issues and tensions in an era of evolving expectations* (pp. 75–105). Routledge.

APPENDICES

Appendix A

IRB Letter

Institutional Review Board

Exemption #2
Certificate of Exemption

From: Melanie Domenech Rodriguez, IRB Chair

            Nicole Vouvalis, IRB Director

To: *Marla Robertson*

Date: *February 15, 2022*

Protocol #:   *12510*

Title: *INTEGRATED CONCEPT MAPPING AS A METHOD FOR DEMOCRATICALLY EVALUATING A TEACHER PREPARATION PROGRAM IN THE AREA OF CLASSROOM ASSESSMENT PROFICIENCY*

The Institutional Review Board has determined that the above-referenced study is exempt from review under federal guidelines 45 CFR Part 46.104(d) category #2:

> *Research that only includes interactions involving educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior (including visual or auditory recording) if at least one of the following criteria is met: (i) The information obtained is recorded in such a manner that the identity of the human subjects cannot readily be ascertained, directly or through identifiers linked to the subject; (ii) Any disclosure of the responses outside the research would not reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, educational advancement, or reputation, or (iii) the information obtained is recorded by the investigator in such a manner that the identity of the human subjects can readily be ascertained, directly or through identifiers linked to the subjects, and the IRB conducts a limited IRB review to make required determinations.*

This exemption is valid for five years from the date of this correspondence, after which the study will be closed. If the research will extend beyond five years, it is your responsibility as the Principal Investigator to notify the IRB **before** the study's expiration date and submit a new application to continue the research. Research activities that continue beyond the expiration date without new certification of exempt status will be in violation of those federal guidelines which permit the exempt status.

If this project involves Non-USU personnel, they may not begin work on it (regardless of the approval status at USU) until a Reliance Agreement, External Research Agreement, or separate protocol review has been completed with the appropriate external entity. Many schools will not engage in a Reliance Agreement for Exempt protocols, so the research team must determine what the appropriate approval mechanism is for their Non-USU colleagues. As part of the IRB's quality assurance procedures, this research may be randomly selected for audit during the five-year period of exemption. If so, you will receive a request for completion of an Audit Report form during the month of the anniversary date of this certification.

In all cases, it is your responsibility to notify the IRB **prior** to making any changes to the study by submitting an Amendment request. This will document whether or not the study still meets the requirements for exempt status under federal regulations.

Upon receipt of this memo, you may begin your research. If you have questions, please call the IRB office at (435) 797-1821 or email to irb@usu.edu.

The IRB wishes you success with your research.

Appendix B

(Pre-study) Semistructured Interview Protocol

Advisory Group Semistructured Interview Protocols

1. In a synchronous Microsoft Teams Meeting with the advisory group, the process and topic is introduced, and examples are presented. I will present the following semistructured group interview questions:

    a. How could the process of integrated concept mapping contribute to the evaluation of our teacher preparation program in the area of classroom assessment outcomes?

    b. How could the process of integrated concept mapping shift the focus of evaluation in our TPP toward a more democratic approach to accountability?

    c. Are there any additional questions or concerns, relative to this process or what we hope to discover?

    d. Creation of Open-ended statement for Meeting 1

2. An audio recording of the meeting was saved as a transcript. I listened to the recording and read the transcript. I made notes in the three-column transcript journal. In one column I used open-coding and in the third column I wrote final concepts and categories as they emerged.

Appendix C

(Post-Study) Semistructured Interview Protocol

Advisory Group Semistructured Interview Protocols

1. In a synchronous Microsoft Teams Meeting with the advisory group, the process and topic was reviewed briefly. I will present the following semistructured group interview questions:

   a. How did the process of integrated concept mapping contribute to the evaluation of our teacher preparation program in the area of classroom assessment outcomes?

   b. How did the process of integrated concept mapping shift the focus of evaluation in our TPP toward a more democratic approach to accountability?

   c. Are there any additional questions or concerns, relative to this process or what we may have discovered?

2. An audio recording of the meeting will be saved as a transcript. I listened to the recording and read the transcript. I made notes in a three-column transcript journal. In one column I will used open-coding and in the third column I will write final concepts and categories as they emerged.

Appendix D

Delineation of the Coding to Theory Process

Delineation of the Coding to Theory Process

Open coding-prior to the evaluation

In vivo codes
- The process promotes stakeholder buy-in.
- Appreciation of working with a group of colleagues.
- It can be difficult to say things in a group-more participation if it is anonymous
- Everyone has a say
- We can all see the big picture.
- It becomes our program
- You're more free to really say what you think.
- Not embarrassed to have to speak in public.
- Not worried to say what you really think.
- Might help with blind spots-areas not noticed if just a few people take part in evaluation or planning.
- Not outside in-but what is actually happening in the program from within
- Opportunity to learn and grow
- May remove power dynamics and individuals that overpower others in face to face situations.
- More voices
- More inclusive of ideas.

Researcher denoted codes
- Current accreditation processes problematic
    - "really labor intensive, technical, detailed things that somebody else needs to take care of because it's so difficult"
- Commitment to understand the process
- Desire to take part in collaboration
- Organic
- Desire to participate in the process
- Time commitment is a frequent consideration
- Seeking to understand the topic and process
- Genuinely concerned with being able to give this the time needed
- Desire to work closely with colleagues
- Congeniality-enjoyment of the process and working together for a common goal.
- Concern with getting it right
- Seeking for clarification and to understand the topic.
- Expertise demonstrated
- Investment in the process
- Evidence of working through the process together

Open-coding-**after the evaluation**

In vivo codes
- Liked seeing everyone's ideas-within the process.
- Collaborative and transparent process
- Expand[ed] collective understanding and thinking
- Common dialogue
- Connection with each other
- May lead to better alignment within the program
- May impact the future direction of the program
- Importance of including student voices
- Desire to know even more and include students more.
- May help with future planning and identifying gaps
- May help see difference in perspective across roles within the program as it relate to classroom assessment.
- More invested in what we can learn from the evaluation.
- Can this help with the other method we are currently required to use for evaluation.
- More useful process-streamlined-faster-more inclusive and transparent.
- Helpful to make future decisions about the program.
- Motivates people
- Gets a conversation going
- Builds in a mechanism to make important conversation happen-to make sure it takes place.
- This structures the time and space for it to happen
- More ideas included
- Makes us slow down and think and do
- We need to do things like this more often.
- It brings us together
- We got to see the bigger picture
- We got to see the context and specifics
- More voices-more participation-inclusive
- More involved
- I liked working independently and collaboratively
- Could this be combined with other method of evaluation
- Concern about voicing different views and some people getting frustrated.
- Could this lead to disenfranchisement within an organization?
- People can get frustrated within democratic processes
- What happened when we disagree?
- What if someone's ideas are not included?
- Good to get people together
- Good to look at program from an internal perspective

- Was not a painful process
- It did take some time
- Ownership of data by those within the program
- This tool could be valuable within other contexts -such as within K-12 schools
- Not overly labor intensive
- I could see where we each contributed to the process
- I can see where we are represented in each area
- At the school level decisions are often made with partial input- this has multiple points of review and reflection and has a formal way of involving people
- Better thinking about issues and better buy in- it's much more defensible
- Including more stakeholders is a good idea
- Maybe we include more stakeholders who are at different points in program-students.
- Could be used to help students have a better experience throughout the program
- Could provide information earlier
- Process does a good job of gathering consensus views-but what about unique views?
- Were outliers included? Does the process tell us this?
- I wanted to see what my colleagues had to say-we can become isolated.
- extremely valuable process
- See the way other folks were thinking
- Interesting process
- Helps to understand how others are thinking about assessment.
- Does the process capture outliers?
- Is validity an issue in this study and within the survey.
- We need for more valid and reliable assessments.

Researcher denoted codes

- Working through things together
- Engagement
- Excitement
- Democratic
- More invested
- More impactful
- More voices
- Time concern

First Iteration: Concepts prior to and following the evaluation

<mark>**Researcher denoted codes highlighted in yellow**</mark>

1. More Voices (Concept)
Prior to evaluation
- More voices
- More inclusive of ideas.
- Everyone has a say
- Feel more free to say what you really think
- It can be difficult to say things in a group-more participation if it is anonymous [like this study]
- You're more free to really say what you think.
- Not embarrassed to have to speak in public.
- Not worried to say what you really think.

After evaluation
- Importance of including student voices
- More ideas included
- More voices-more participation-inclusive
- What if someone's ideas are not included?
- Including more stakeholders is a good idea
- See the way other folks were thinking
- I wanted to see what my colleagues had to say-we can become isolated.
- <mark>More voices</mark>

2. Stakeholder buy-in (Concept)
- Process promotes stakeholder buy in
- Not outside in-but what is actually happening in the program from within
- <mark>Commitment to understand the process</mark>
- <mark>Desire to take part in collaboration</mark>
- <mark>Desire to participate in the process</mark>

After study
- Motivates people
- We need to do things like this more often
- Good to look at program from an internal perspective
- Ownership of data by those within the program
- Was not a painful process
- Better thinking about issues and better buy in-it's much more defensible
- <mark>More invested</mark>

4. Commitment to Process (Concept)
- <mark>Seeking to understand the topic and process</mark>

- Genuinely concerned with being able to give this the time needed
- Concern with getting it right
- Seeking for clarification and to understand the topic.
- Investment in the process

After study
- More invested in what we can learn from the evaluation.
- It did take some time
- Not overly labor intensive
- Is validity an issue in this study and within the survey.

5. Provides a structure for democratic evaluation (Concept)
- May help to identify blind spots
- Promotes greater participation
- Current accreditation processes are "really labor intensive, technical, detailed things that somebody else needs to take care of because it's so difficult"
- We can all see the big picture.
- Time commitment is a frequent consideration and topic
- Organic

After study
- Can this help with the other method we are currently required to use for evaluation.
- Builds in a mechanism to make important conversation happen-to make sure it takes place
- More people get to be a part of this process versus current evaluation methods
- This structures the time and space for it to happen
- Makes us slow down and think and do
- I could see where we each contributed to the process
- I can see where we are represented in each area
- Interesting process.
- extremely valuable process
- Does the process capture outliers?
- At the school level decisions are often made with partial input- this has multiple points of review and reflection and has a formal way of involving people
- More people get to be a part of this process versus current evaluation methods.=
- Democratic
- Time concern
- May remove power dynamics and individuals that overpower others in face to face situations.

After study
- Process does a good job of gathering consensus views-but what about unique views?
- Were outliers included? Does the process tell us this?

6. Collaborative effort (Concept)
- Opportunity to learn and grow.
- Appreciate working together
- Collaboration
- Appreciation of working with a group of colleagues.
- Desire to work closely with colleagues
- Congeniality-enjoyment of the process and working together for a common goal.
- Expertise demonstrated
- Evidence of working through the process together

After study
- Helps to understand how others are thinking about assessment
- Liked seeing everyone's ideas-within the process.
- Collaborative and transparent process
- Connection with each other
- Gets a conversation going
- It brings us together
- More involved
- I liked working independently and collaboratively
- What happened when we disagree?
- Working through things together
- Engagement
- Excitement

7. Expanded collective understanding (Concept)
After study
- Expand[ed] collective understanding and thinking
- Common dialogue
- We got to see the bigger picture
- We got to see the context and specifics
- More impactful
- It becomes our program

8. Future Impact (Concept)
After study
- May lead to better alignment within the program
- May impact the future direction of the program
- May help with future planning and identifying gaps
- May help see difference in perspective across roles within the program as it relate to classroom assessment.
- Helpful to make future decisions about the program.
- Could this be combined with other methods of evaluation
- This tool could be valuable within other contexts -such as within K-12 schools

- Maybe we include more stakeholders who are at different points in program-students.
- Could be used to help students have a better experience throughout the program
- Could provide information earlier within program
- Desire to know even more and include more students (prior to student teaching).
- Can this help with the other methods we are currently required to use for evaluation.

Second Iteration:
Categories (prior and following the evaluation):

Researcher denoted codes highlighted in yellow

1. Being included (phenomenon)
   - More voices
   - More inclusive of ideas.
   - Everyone has a say
   - Feel more free to say what you really think
   - It can be difficult to say things in a group-more participation if it is anonymous [like this study]
   - You're more free to really say what you think.
   - Not embarrassed to have to speak in public.
   - Not worried to say what you really think.
   - May help to identify blind spots

After study
   - Importance of including student voices
   - More ideas included
   - More voices-more participation-inclusive
   - What if someone's ideas are not included?
   - Including more stakeholders is a good idea
   - See the way other folks were thinking
   - I wanted to see what my colleagues had to say-we can become isolated.
   - More people get to be a part of this process versus current evaluation methods
   - I could see where we each contributed to the process
   - I can see where we are represented in each area
   - More people get to be a part of this process versus current evaluation methods.
   - May remove power dynamics and individuals that overpower others in face to face situations.
   - Liked seeing everyone's ideas-within the process.
   - Process does a good job of gathering consensus views-but what about unique views?

- Were outliers included? Does the process tell us this?
- Does the process capture outliers?
- More voices

2. Taking Ownership (phenomenon)
- Process promotes stakeholder buy in
- Not outside in-but what is actually happening in the program from within
- Commitment to understand the process
- Desire to take part in collaboration
- Desire to participate in the process
- Seeking to understand the topic and process
- Genuinely concerned with being able to give this the time needed
- Concern with getting it right
- Seeking for clarification and to understand the topic.
- Investment in the process

After study
- Motivates people
- We need to do things like this more often
- Good to look at program from an internal perspective
- Ownership of data by those within the program
- Was not a painful process
- Better thinking about issues and better buy in-it's much more defensible
- More invested in what we can learn from the evaluation.
- It did take some time
- Not overly labor intensive
- Is validity an issue in this study and within the survey.
- It becomes our program
- More impactful
- We got to see the bigger picture
- Interesting process
- Extremely valuable process
- More invested

3. Understanding the Structure (phenomenon)
- Promotes greater participation
- Current accreditation processes are- "really labor intensive, technical, detailed things that *somebody else needs to take care* of because it's so difficult"
- We can all see the big picture.
- Time commitment is a frequent consideration and topic
- Organic

After study
- Builds in a mechanism to make important conversation happen-to make sure it

takes place
- This structures the time and space for it to happen
- Makes us slow down and think and do
- At the school level decisions are often made with partial input- this has multiple points of review and reflection and has a formal way of involving people
- Democratic
- Time concern

4. Collaborating (phenomenon)
- Opportunity to learn and grow.
- Appreciate working together
- Collaboration
- Appreciation of working with a group of colleagues.
- Desire to work closely with colleagues
- Congeniality-enjoyment of the process and working together for a common goal.
- Expertise demonstrated
- Evidence of working through the process together

After study
- Helps to understand how others are thinking about assessment
- Collaborative and transparent process
- Connection with each other
- Gets a conversation going
- It brings us together
- [we were] more involved
- I liked working independently and collaboratively
- Expand[ed] collective understanding and thinking
- Common dialogue
- Expanded collective understanding
- What happens when we disagree?
- We got to see the context and specifics
- Working through things together
- Engagement
- Excitement

5. Considering Impacts
After study
- May lead to better alignment within the program
- May impact the future direction of the program
- May help with future planning and identifying gaps
- May help see difference in perspective across roles within the program as it relate to classroom assessment.
- Helpful to make future decisions about the program.

- Could this be combined with other methods of evaluation
- This tool could be valuable within other contexts -such as within K-12 schools
- Maybe we include more stakeholders who are at different points in program-students.
- Could be used to help students have a better experience throughout the program
- Could provide information earlier within program
- Desire to know even more and include more students (prior to student teaching).
- Can this help with the other methods we are currently required to use for evaluation.

Categories:
- Desire to be heard (phenomenon)
- Seeking a Deeper Understanding (phenomenon)
- Strong Collaboration (phenomenon)
- Taking Ownership (phenomenon
- Considering Impacts (phenomenon)

Axial Coding Paradigm: *Stakeholder commitment to evaluation is strengthened through democratic accountability methods.*
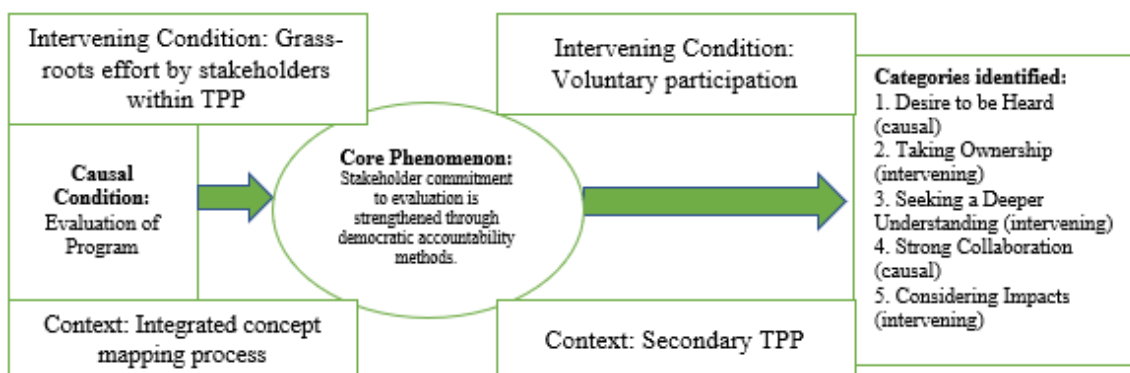
Core Phenomenon

The process of analyzing data collected for this study led me to identify a "Core Phenomenon" within this study: *Stakeholder commitment to evaluation is strengthened through democratic accountability methods.*

Core Phenomenon-Details

Using grounded theory, I found a *Desire to be Heard* (causal condition) and

Strong Collaboration (causal condition) among stakeholders within a Secondary TPP (faculty, administration, and teacher-candidates) led to *Taking Ownership (causal condition)* of the evaluation process and evaluation outcomes. The grass-roots method for evaluation (integrated concept mapping), led stakeholders to *Seek for a Deeper Understanding (intervening condition)* of both the evaluation topic (classroom assessment) and the method of evaluation, Integrated Concept Mapping (ICM). Finally, taking part in the evaluation led stakeholders to C*onsider Possible Impacts (intervening condition)* of the ICM process for evaluation and the evaluation outcomes in broader contexts

Appendix E

Coding Excerpt

00:15:10.860 --> 00:15:17.860	Liked seeing everyone's ideas- within the process.	Strong Collaboration
Faculty Member 1		
Yeah. No, I really liked being able to see what other people wrote, you know, with all the statements. Like, that was really.		
00:15:18.780 --> 00:15:21.180		
Faculty Member 1		
Helpful and like help me, Clara.		
00:15:21.260 --> 00:15:24.980	The collaborative and transparent process expanding the collective understanding and thinking	Strong Collaboration
Faculty Member 1		
Five, you know things and expand my thinking and.		
00:15:26.880 --> 00:15:34.390		
Faculty Member 1		
Kind of. See the wide variety of expertise that students were coming in contact with. They made me wanna know. It really made me want to.	Engagement	Seeking a Deeper Understanding
00:15:36.310 --> 00:15:47.330	Common dialoged	
Faculty Member 1	Excitement	
Talk about it with everybody you know, like get together and talk about how this is showing up in our classes and where we could strengthen and where we could, where we were overlapping too much and like that kind of stuff.	Connection with each other	Strong Collaboration
00:15:48.450 --> 00:16:00.930		Taking Ownership Considering Impacts
Faculty Member 1		
So it raised a lot of questions for me in that way. But really good, I think helpful questions that like could lead to better alignment and cohesion and expansion moving forward, so.	Helpful	
00:16:01.890 --> 00:16:04.460	May lead to better alignment within the program	
Faculty Member 1		
I really liked that aspect of it and.		Taking ownership
00:16:06.470 --> 00:16:09.680		
Faculty Member 1	May impact the future direction of the program	
To bring in the student voices. I mean, that's just.		
00:16:09.490 --> 00:16:09.930		Considering impacts
Faculty Member 2		
It's just.		
00:16:10.480 --> 00:16:14.900		
Faculty Member 1		

Such a democratic approach, and like just, you know, to know what they think.
00:16:16.390 --> 00:16:18.660
Faculty Member 1
It makes me want to know more about like.
00:16:20.260 --> 00:16:29.950
Faculty Member 1
Their understanding of those concepts, too, and how then they're assessing themselves based on those, their understanding of those concepts. But like to hear from them.
00:16:31.970 --> 00:16:37.060
Faculty Member 1
Feels like a really helpful piece of the puzzle and thinking about how we're doing.
00:16:38.130 --> 00:17:04.900
Faculty Member 2
Yeah, that's one thing I thought about, too, is, you know, one thing we talked about a lot in, in data is like triangulations. So, could this be part of a larger, you know, gathering of data points? So, we have their Ed TPA data point for task three that's related to classroom assessment practices. You know, we would maybe have this data point and see how they align and then I don't know if they're.
00:16:38.210 --> 00:16:38.570
Faculty Member 1
So.
00:17:05.070 --> 00:17:36.540
Faculty Member 1
Is an evaluation for student teachers by their supervisors and cooperating teachers in the area of classroom assessment, but maybe to look at how would all three of those things correlate that would you know, maybe strengthen that a little bit in terms of the data? Because I think maybe, you know that was the first thing when I looked at it, I

Importance of including student voices	Desire to be heard
Forwards a democratic approach	Taking Ownership Considering Impacts
Desire to know even more and include student more.	Considering Impacts
Helpful to consider with other evaluations	Considering Impacts
Considering-teacher candidates knowledge	Considering Impacts
Helpful to consider with other evaluations	

thought, wow, they're really confident. But then as I started to read their examples, you know, I was like, OK, they have examples of that.
00:17:36.710 --> 00:17:49.620

Faculty Member 2

And obviously I've had enough time to look at all of that and go through that data. But I think it will be interesting to see if their qualitative statements kind of match the ratings they provided.
00:17:50.580 --> 00:17:58.770

Faculty Member 1

Right. And I like that there's that piece so that we know not only their confidence but their understanding like that's really helpful.
00:18:01.410 --> 00:18:02.430

Faculty Member 1

Yeah, part of it.
00:18:02.620 --> 00:18:07.090

Carrie Ashcraft

Appreciate it. Anyone else have thoughts or you know on either of those questions?
00:18:10.040 --> 00:18:23.330

Administrator 1

You know, one of the things, uh, Carrie, that I was interested in here and I think is something that we can certainly use at the school level is the difference in expectations or you know.
00:18:23.420 --> 00:18:53.680

Administrator 1

The congruence across the expectations with different levels. I mean, I think that was one of the things that was most interesting to me was sort of seeing what you know, department heads or administrators of some sort versus faculty versus students in terms of what they thought, what we thought you know, our outcome should be for a program. And so, I think that's really interesting and that.

Open-ended question data was an important aspect of the survey.	Seeking a Deeper Understanding
Interested in understanding the data Helpful	Seeking a Deeper Understanding

00:18:54.320 --> 00:18:54.700
Administrator 1
You know.
00:18:55.380 --> 00:19:27.050
Administrator 1
I hope that in the end we were pretty close on all of that, but I think also that, that that would be an opportunity to sort of identify any gaps where you know either, you know, professors had one idea and instructors had one idea about what they thought they should be teaching students and, you know, administrators had another. And so, I thought that was really interesting and something that could contribute to, you know, our evaluation of our program.
00:19:27.190 --> 00:19:44.120
Administrator 1
It is, we should have uh, you know, parity across all levels. But if we don't, then that's an issue that needs to be addressed, you know, to make sure that expectations are clear so that we get the best for our students.
00:19:45.560 --> 00:20:16.210
Faculty Member 2
Yeah, thank you, Umm, one thing that was brought up in I did this meeting at 10:00 o'clock with some other and faculty and administrators and they meant somebody mentioned and I agree expanding that stakeholders to include cooperating teachers and administrators in the school and that also could be helpful especially in the area of classroom assessment to first of all gather their statements and then have them rate those as well.
00:20:16.290 --> 00:20:35.020

our student teachers and

May help with future planning and identifying gaps	Considering Impacts
May help see difference in perspective across roles within the program as it relate to classroom assessment.	Considering Impacts
Invested in what we can learn from the evaluation.	Taking Ownership

Administrator 2 Alright, can carry, I think anytime you can get the. 00:10:31.010 --> 00:10:34.960	Good to get people together	Strong Collaboration
Administrator 2 They're actually looking at the, you know, looking at a program internally. 00:10:36.310 --> 00:10:40.610	Good to look at program from am internal perspective	Taking Ownership
Administrator 2 Wasn't a painful process, it was. It was time, time consuming but. 00:10:41.340 --> 00:10:46.520	Was not a painful process	Taking Ownership
Administrator 2 It gives ownership to the to the way you start to collect data within the organization. 00:10:47.660 --> 00:10:54.780	It did take some time Time concern Ownership of data by those within the program	Taking Ownership
Administrator 2 I'm looking back from my perspective as a principal over some of the grants that I supervised and how those grants got evaluated and. 00:10:56.290 --> 00:10:59.460		Considering Impacts Desire to be heard
Administrator 2 This in terms of school improvement grants. 00:11:00.230 --> 00:11:08.090		Seeking a Deeper Understanding
Administrator 2 Would have been a valuable way to incorporate stakeholder input into the evaluation process. So as far as process goes. 00:11:09.310 --> 00:11:13.540	This tool could be valuable within other contexts -such as within K-12 schools	Considering Impacts
Administrator 2 I didn't feel like it was overly in labor intensive. 00:11:15.210 --> 00:11:21.640	Not overly labor intensive	Seeking a Deeper Understanding
Administrator 2 But it did. It did feel a little bit like there was some ownership there because I could. I mean I can see where each of the things we put in. 00:11:22.590 --> 00:11:25.730	I could see where we each contributed to the process	Strong Collaboration

Administrator 2
Are represented in those in those eight silos.
00:11:27.890 --> 00:11:58.380
Faculty Member 2
Uh, thank you, Joey. Appreciate your comments. I maybe wanna talk about a couple things you brought up, which is ownership. I know a couple of people brought this up originally. I think one person individually in said something about like you know, when you're trying to get people on board for a program and one thought that I had about that is there's a difference between trying to get people on board and having all the stakeholders come together to create something.
00:11:58.520 --> 00:12:05.570
Faculty Member 2
And that may be just something that maybe will wanna to think about or if we wanna take it in a different direction, that's great too.
00:12:08.720 --> 00:12:16.640
Faculty Member 3
Here just a quick thought along the same lines that a strength I saw in in the in participating one it was.
00:12:18.740 --> 00:12:20.870
Faculty Member 3
It seems, at least at the school level.
00:12:21.890 --> 00:12:40.880
Faculty Member 3
Decisions are often made with partial input worth one off input. This and having a protocol process like this that has multiple points of review and reflection and has a formal way of involving people, I think it's really noted is a great way to.
00:12:41.940 --> 00:12:50.810
Faculty Member 3
Do you get better thinking that way and you get better buy in and it's much more

I can see where we are represented in each area	Desire to be heard
At the school level decisions are often made with partial input- this has multiple points of review and reflection and has a formal way of involving people	Strong Collaboration Seeking a Deeper Understanding

defensible. So anyway, but that can't echoes what Joey had to say, but that's what I thought.
00:12:52.010 --> 00:13:22.320
Faculty Member 2
Thank you. And one thing I've been thinking about too is, you know, we were able to get some information from students as stakeholders, from faculty, from administration, and this is beyond the scope of this study. But how much more it maybe would have meant if we were able to gather data from cooperating teachers, administrators in the schools where students are working and gather their thoughts on classroom assessment practices too, and kind of incorporate that as well.
00:13:22.420 --> 00:13:34.670
Faculty Member 3
And so you know, that's just something to think about is, you know, could more stakeholders be involved and how would that contribute to, you know, this process and what comes out of it as well?
00:13:38.080 --> 00:13:39.890
Faculty Member 4
Had a couple thoughts also.
00:13:41.550 --> 00:13:47.860
Faculty Member 4
Just thinking about, so I like that idea of more stakeholders. I'm also wondering about.
00:13:50.720 --> 00:14:00.290
Faculty Member 4
Possibly following a similar process at different points during the program. So, the advantage of having it at the end is they've kind of experienced everything that the program has to offer.
00:14:01.040 --> 00:14:01.570
Faculty Member 4
Umm.

Better thinking about issues and better buy in- it's much more defensible

Including more stakeholders is a good idea

Maybe we include more stakeholders who are at different points in program- students.

Seeking a Deeper Understanding

Considering Impacts

Considering Impacts

00:14:03.270 --> 00:14:05.080 Faculty Member 4 But I think having it earlier on. 00:14:06.330 --> 00:14:14.020 Faculty Member 4 In terms of developing buy in can helpful can be helpful in for students having a better experience in the program as well as giving us information. 00:14:14.960 --> 00:14:20.430 Faculty Member 4 Umm that can be used to benefit them while they're still here, if that makes sense. 00:14:21.000 --> 00:14:37.250 Faculty Member 2 Yeah. So, I just want to clarify, do you mean maybe like ongoing feedback when it comes to this? And obviously we're just looking at one area of the program, but do you mean maybe gathering data at different points from the students as stakeholders as well as the faculty? 00:14:37.670 --> 00:14:38.770 Faculty Member 4 Yeah, yeah. 00:14:40.400 --> 00:14:42.950 Faculty Member 4 And one thing I'm wondering about and. 00:14:43.910 --> 00:14:44.450 Faculty Member 4 Umm. 00:14:45.780 --> 00:14:55.140 Faculty Member 4 It seems like from what I understand of the process that it's it does a good job of kind of capturing consensus views. 00:14:55.990 --> 00:14:57.590 Faculty Member 4 And I'm wondering. 00:14:58.410 --> 00:15:02.810 Faculty Member 4 How it can it does or could UM?	Could be used to help students have a better experience throughout the program Could provide information earlier Process does a good job of gathering consensus views-but what about unique views?	Considering Impacts Seeking a Deeper Understanding

Appendix F

Survey Administered to Group 3

Default Question Block

Please review the following IRB letter of consent before proceeding. You may download the document for your records by clicking on the link.

12510 robertson loi student final

 If you consent to participate in the survey, please choose "I consent" and click the arrow at the bottom of the page to continue on to the survey.

If you do not consent to participate, choose "I do not consent" and click the arrow at the bottom of the page to be finished.

○ I consent.   ○ I do not consent.

Classroom Assessment Proficiency Survey

This survey is a self-assessment of teacher-candidate's classroom assessment proficiencies. The purpose for this survey is to better understand the knowledge and skills teacher-candidates possess after completing their course work and clinical experience (student teaching or an internship). This information will be used to evaluate the secondary education program in the area of classroom assessment preparation.

Explanation of Terms:

Classroom assessments include a large variety of activities that take place within classrooms such as verbal questioning, portfolios, quizzes, performance demonstrations, and more. Data collected through classroom assessment should be used to inform ongoing instruction and can be defined as the "...formal and informal procedures that teachers employ in an effort to make accurate inferences about what their students know and can do" (Popham, 2009, p. 6).

Validity, in the context of classroom assessments, refers to the accuracy of an assessment, "...whether or not the assessment measures what it is supposed to measure" (Florida Center for Instructional Technology).

Reliability, in this context, is the ability of the assessment to produce accurate results; typically an assessment is reliable if the results can be replicated.

Quantitative analysis is associated with the numerical analysis of data. The goal of quantitative analysis is to understand the occurrence of events and describe them through statistical methods.

Qualitative analysis involves identification, examination, and interpretation of patterns and themes within the data collected. It seeks to understand the meaning within the patterns and themes identified.

Survey Instructions:

1. Choose the content area that is most closely associated with your education major.

2. Rate your level of classroom assessment proficiency on the following scale: Not Proficient, Somewhat Proficient, Proficient, Very Proficient, Extremely Proficient

3. At the end of each section of the survey, review the concept and the associated questions. If possible, provide examples and/or explanation of your experiences related to classroom assessment in connection with the concept and questions in the area provided.

Identify your content area (choose one):

○ English Language Arts
○ History/Social Studies
○ Math
○ Science
○ Business/Engineering
○ Dance
○ Theater
○ Visual Art
○ Music
○ World Languages
○ Physical Education
○ Health
○ SPED

Concept 1: Classroom Assessment-General Assessment Knowledge

	Not Proficient	Somewhat Proficient	Proficient	Very Proficient	Extremely Proficient
1. I know how to include classroom assessments in lesson planning.	○	○	○	○	○
2. I know how to use classroom assessment as part of an ongoing cycle of instruction and feedback.	○	○	○	○	○
3. I know the difference between formative and summative assessments.	○	○	○	○	○
4. I know how to align classroom assessments with state standards.	○	○	○	○	○
5. I know how to use classroom assessments to measure student progress against state standards.	○	○	○	○	○

6. Review this section (the concept and associated questions). If possible, provide examples and/or explanation of your experiences during student teaching or an internship as they relate to the concept, questions, and your proficiency ratings.

Concept 2: Classroom Assessment-Designing and Choosing Assessments

	Not Proficient	Somewhat Proficient	Proficient	Very Proficient	Extremely Proficient
1. I design and/or choose a variety of classroom assessments that allow students to demonstrate what they know and can do.	○	○	○	○	○
2. I design and/or choose assessments that approach a higher level on Bloom's Taxonomy and/or a deeper level on Webb's Depth of Knowledge.	○	○	○	○	○
3. I design and/or choose rubrics/grading criteria that are aligned to the assessment.	○	○	○	○	○
4. I design and/or choose rubrics/grading criteria that are easily interpreted by students.	○	○	○	○	○
5. I design and/or choose differentiated classroom assessments.	○	○	○	○	○

6. Review this section (the concept and associated questions). If possible, provide examples and/or explanation of your experiences during student teaching or an internship as they relate to the concept, questions, and your proficiency ratings.

Concept 3: Classroom Assessments- Validity and Reliability

	Not Proficient	Somewhat Proficient	Proficient	Very Proficient	Extremely Proficient
1. I align classroom assessments, rubrics/grading criteria, and state standards.	○	○	○	○	○
2. I use researched-based classroom assessments methods.	○	○	○	○	○
3. I evaluate the reliability of classroom assessments.	○	○	○	○	○
4. I evaluate the validity of classroom assessments.	○	○	○	○	○
5. I evaluate common assessments with colleagues.	○	○	○	○	○

6. Review this section (the concept and associated questions). If possible, provide examples and/or explanation of your experiences during student teaching or an internship as they relate to the concept, questions, and your proficiency ratings.

Concept 4: Classroom Assessment: Analysis, Monitoring, and Tracking

	Not Proficient	Somewhat Proficient	Proficient	Very Proficient	Extremely Proficient
1. I analyze classroom assessment data quantitatively.	○	○	○	○	○
2. I analyze classroom assessment data qualitatively.	○	○	○	○	○
3. I formally track student learning, in connection with classroom assessments.	○	○	○	○	○
4. I create progress charts or graphs based on classroom assessment data.	○	○	○	○	○
5. I interpret progress charts or graphs that reflect classroom assessment data.	○	○	○	○	○

6. Review this section (the concept and associated questions). If possible, provide examples and/or explanation of your experiences during student teaching or an internship as they relate to the concept, questions, and your proficiency ratings.

Concept 5: Classroom Assessment: Instruction and Interventions

	Not Proficient	Somewhat Proficient	Proficient	Very Proficient	Extremely Proficient
1. I adjust lesson plans based on classroom assessment data.	○	○	○	○	○
2. I use classroom assessments to inform on-going decisions about differentiation.	○	○	○	○	○
3. I use data from classroom assessments to provide individual instruction.	○	○	○	○	○
4. I adapt interventions based on classroom assessment data for students with an IEP or 504.	○	○	○	○	○
5. I design common assessments with colleagues.	○	○	○	○	○

6. Review this section (the concept and associated questions). If possible, provide examples and/or explanation of your experiences during student teaching or an internship as they relate to the concept, questions, and your proficiency ratings.

Concept 6: Classroom Assessments-Equitable Assessment Practices

	Not Proficient	Somewhat Proficient	Proficient	Very Proficient	Extremely Proficient
1. I modify assessments for students with an IEP or 504.	○	○	○	○	○
2. I modify assessments for students learning English.	○	○	○	○	○
3. I make trauma-informed decisions when considering classroom assessments.	○	○	○	○	○
4. I design/choose assessments that are culturally relevant.	○	○	○	○	○
5. I consider personal biases that may influence classroom assessments.	○	○	○	○	○

6. Review this section (the concept and associated questions). If possible, provide examples and/or explanation of your experiences during student teaching or an internship as they relate to the concept, questions, and your proficiency ratings.

Concept 7: Classroom Assessment- Variety and Pacing of Assessments

	Not Proficient	Somewhat Proficient	Proficient	Very Proficient	Extremely Proficient
1. I know when to use formative and summative classroom assessments.	○	○	○	○	○
2. I regularly carry-out a variety of informal classroom assessments.	○	○	○	○	○
3. I provide multiple opportunities for students to demonstrate proficiency in connection with learning goals.	○	○	○	○	○
4. I include many low-stakes assessments within daily classroom instruction.	○	○	○	○	○
5. I adapt the pace of classroom assessments for students with different needs and abilities.	○	○	○	○	○

6. Review this section (the concept and associated questions). If possible, provide examples and/or explanation of your experiences during student teaching or an internship as they relate to the concept, questions, and your proficiency ratings.

Concept 8: Classroom Assessment–Community and Feedback

	Not Proficient	Somewhat Proficient	Proficient	Very Proficient	Extremely Proficient
1. I assist students and/or guardians in interpreting classroom assessment data.	○	○	○	○	○
2. I provide corrective and supportive feedback to students, in connection with classroom assessments.	○	○	○	○	○
3. I plan ways for students to respond to feedback provided on classroom assessments.	○	○	○	○	○
4. I create an environment where students view assessments as a positive way to check for academic progress.	○	○	○	○	○
5. I collaborate with peers on common assessments.	○	○	○	○	○

6. Review this section (the concept and associated questions). If possible, provide examples and/or explanation of your experiences during student teaching or an internship as they relate to the concept, questions, and your proficiency ratings.

Appendix G

Edited Statement List

Statement number, ENGLISH						
1,help their students be aware of what assessment data is communicating about what they know and can do.						
2,understand and apply research-based practices in the area of assessment.						
3,provide multiple opportunities for students to demonstrate proficiency in connection with learning targets.						
4,modify assessments for students with an IEP or 504.						
5,modify assessments for students learning English.						
6,include many low stakes assessments within lesson planning.						
7,understand the role of assessment in the classroom.						
8,know the difference between a formative and summative assessment.						
9,know the difference between a formal and informal assessment.						
10,analyze assessment data for students, both qualitatively and quantitatively						
11,think critically about assessments they create and what it tells then about where their students are in connection with the standards or learning goals.						
12,create alternative assessments formats that allow students to demonstrate what they know and can do in different ways.						
13,consider ways that students may respond to feedback provided by the teacher.						
14,provide feedback to students that is both corrective and supportive						
15,create assessments that approach a higher level on Bloom's Taxonomy or Webb's Depth of Knowledge						
16, align the assessment, rubrics/grading criteria, and state standards.						
17,use multiple types of assessment information to make decisions about students with disabilities						
18,develop authentic performance assessments.						
19,knowledge of interpreting assessment results to guide educational placement decisions for students with disabilities						
20,use appropriate informal and formal assessments throughout instruction.						
21,Identify appropriate evidence of learning for learning goals						
22,create a progress graph and read visual graphs on student progress.						
23,adapt interventions used based on students weekly data summary.						
24,use assessments to determine treatment plans for individuals.						
25,create and use data sheets to track student learning.						
26,Identify researched-based assessments.						
27,set up a progress monitoring schedule for students.						
28,create data sheets and use them to track progress towards IEP goals.						
29,create their own curriculum-based measurements.						
30,"know how to administer academic, behavior and adaptive testing. "						
31,interpret data to adjust interventions based on the data to meet students IEP goals.						

32,select formal assessments that are evidence-based.					
33,use multiple assessments to make educational decisions.					
34,impliment low-stakes/ low-risk assessments.					
35,co-create learning experiences based on student needs and interests.					
36,adjust lesson planning based on assessment.					
37,make trauma-informed decisions for assessment and student interventions.					
38,use assessment to inform differentiation.					
39,create rubrics to aid in giving feedback to students based on assessments.					
40,create authentic assessments.					
41,create assessments that are culturally relevant.					
42,provide students with a variety of ways to show what they know or can do.					
43,make adjustments to their teaching based on the results of their assessments.					
44,use formative assessments throughout a class.					
45,plan for student assessment in the planning stage of lesson design.					
46,use data from student assessment to differentiate instruction.					
47,understand how to design carefully thought-out test items for multiple-choice, short answer, essay, and so forth.	et" "c. """""				
48,understand specific informal assessment strategies and where to implement them.					
49,use data from student assessment to provide individual instruction and support based on student' s needs .					
50,understand the role of careful student observation as a form of classroom assessment.					
51,know and be able to use a variety of informal classroom-based assessment strategies.					
52,understand when to use of different forms of assessment (such as formal and informal).					
53,assess the reliability of assessments .					
54,check assessments for validity.					
55,explain rubrics to parents and students.					
56,create rubrics that are easy to understand.					
57,understand that a summative assessment is the final assessment.					
58,understand the purpose of a summative assessment.					
59,design assessment practices that encourage students to use knowledge and skills in motivating contexts.					
60,design assessment practices that encourage growth over time.					
61,design assessments where the skills and knowledge being assessed is clearly defined and understood by students.					
62,understand assessment as a complex process involving cycles of feedback and					

growth.					
63,design assessments with colleagues.					
64,design assessments with students.					
65,connect assessment data to instructional decision-making for future planning.					
66,understand the value of equitable assessment practices.					
67,measure how their own teaching is being understood by students.					
68,use assessments to track students' progress against state standards.					
69,use assessments to identify areas for interventions.					
70,create a culture where assessments are seen as a positive way to check progress, not as a punitive measure.					
71,use assessments to better differentiate instruction to meet the needs of all learners.					
72,design simple checks for understanding to assess student progress within lessons.					
73,help students self-assess to better understand their own misconceptions.					
74,design assessments that are culturally responsive.					
75,locate assessments that accurately reflect what students are learning and guide future instruction.					
76,use statewide testing to monitor group and individual student progress.					
77,understand how both formative and summative assessments can be used to support student learning.					
78,provide multiple ways for students to demonstrate learning and understanding.					
79,develop a system for recording and analyzing data that drives instruction.					
80,assess students informally on a daily basis.					
81,interpret assessment data to guide instruction.					
82,prepare students for high stakes assessments.					
83,candidates should be able to align formative assessment with summative assessment and state standards.					
84,candidates should be able to assess without bias.					
85,design formal and informal assessments that allow for differentiation within instruction.					
86,assess in equitable ways.					
87,adapt assessments to individual students needs and abilities.					
88,collaborate with peers on common assessments.					
89,assess student performance with a broad range of formative assessments.					
90,create a variety of assessment tools that provide valid data on student progress.					
91,track and report student progress in a timely manner.					
92,pace assessments to meet the needs of different learner ability levels.					
93,differentiate instruction to meet unique individual needs of students in classroom					

population.				
94,construct assessments that are adaptable to diverse student populations.				
95,locate demographic data on the student population assigned to their courses and school.				
96,connect assessment data to instructional decision-making informally for immediate use during class.				
97,carry-out equitable assessment of students.				
98,understand student needs in connection with assessments.				
99,be aware of personal biases that may influence assessment.				
100,utilize a variety of assessment tools that provide valid data on student progress.				
101,track and report student progress in an efficient manner.				
102,select informal assessments that are evidence-based.				
103,select informal assessments that minimize bias.				
104,select formal assessments that minimize bias.				
105,understand why to use of different forms of assessment (such as formal and informal).				

CURRICULUM VITAE

CARRIE ELIZABETH ASHCRAFT

Utah Valley University
School of Education
Carrie.ashcraft@uvu.edu
Carrie.ashcraft23@gmail.com
801-471-1738 (cell)

EDUCATION

University/College	Degree	Year
Utah State University, Doctor of Education-Literacy	Ph.D.	2023

Dissertation: *Integrated concept mapping as a method for democratically evaluating a teacher preparation program in the area of classroom assessment proficiency*

Chair*:* Marla Robertson, Ph.D.

Westminster College, Master of Education	M.Ed.	2012

English Language Learner (ELL) Concentration
Master's Thesis: *Teacher Perspectives: Working in a Pay for Performance* School

Boise State University, Bachelor of Arts-English Literature	B. A.	1998

PROFESSIONAL WORK EXPERIENCE

Institution	Position	Year
Utah Valley University	Assistant Professor	2022
Utah Valley University	Lecturer	2017- 2022
Utah Valley University	Adjunct Instructor	2015-2017
Nebo School District	Teacher; Current Enrollment	2014-2017
Canyons School District	Teacher; Team Lead	2013-2014
Nebo School District	Teacher; Administrative Team	2005-2013
Aspen Academy of Learning	Teacher; Grant Writer	2000-2005

COURSE INSTRUCTION

Course Name	*Course Number*	*Year*
Content Area Literacy	4440	2022
Flexible Learning-Preparation for Instructors	N/A	2020-2021
Educator Performance Assessment	4990	2019-Present
Foundations of American Education	3050	2017-2022
Classroom Management I	4200	2017-Present
Classroom Management II	4250	2017-Present
Introduction to Research Writing	2010	2016-2018
Introduction to College Writing	1010	2015-2018

COURSE DEVELOPMENT AND DESIGN

Course Name	*Course Number*	*Year*
Classroom Management I (Online and In-person)	4200	2022
Classroom Management II (Online Course)	4250	2020-2021
Flexible Learning-Preparation for Instructors	N/A	2020
Teacher Performance Assessment	4990	2018

PROFESSIONAL LICENSING	*Year*
Level II Secondary Education	2022
Level II Elementary Education	2022

ADDITIONAL PROFESSIONAL COURSEWORK	*Year*
Online Teaching Certification-UVU Office of Teaching and Learning	2020-2021
Service-Learning Certification-UVU Office of Service Learning	2019

RESEARCH INTERESTS

My research interests center on literacy, program evaluation, classroom assessment, equitable access to an advanced curriculum, democratic accountability, and teacher education. I am interested in better understanding how a democratic approach to accountability may transform the field of education; specifically, how it may contribute to the dismantling of "accountability era" evaluation structures and the language associated with such structures.

PUBLICATIONS

Ashcraft, Chou & Jones (2020). Desegregating classrooms in the 21st century: Recognizing and removing barriers to advanced courses and a rigorous curriculum for all students. Teachers College Record.

RESEARCH EXPERINCE	*Year*
Program Evaluation: Evaluating a Secondary Education Program through Teacher-Candidates' Knowledge and Skills of Assessment with Dr. Marla Robertson of Utah State University.	2021
Classroom Assessment Practices of Student Teachers with Dr. Marla Robertson of Utah State University	2020
Review of articles with an equity focus for Utah Academy of Science and Letters	2020
Literature review of Multimodal literacies with Dr. Kortney Sherbine of Utah State University in preparation for a national conference.	2020
Evidence and perceptions of equity and inclusivity within teacher performance assessments with Dr. Makenzie Sellend of Utah Valley University.	2019-2020

CONFERENCE PRESENTATIONS

Institution	**Topic**	*Year*
AAQEP	Democratic Accountability: What about Integrated Concept Mapping?	2023
Utah Valley University	Democracy and Ethics in Education	2022

AAQEP	Teacher Performance Assessments and Accreditation Panel 2022	
AACTE	Equity, Inclusivity, and the edTPA: Pre-service teacher perceptions of secondary education program goals and the edTPA experience.	2020
For the Love of Reading	Using Film and Podcast in the Secondary Classroom as a Gateway, Support, and Extension to Literature	2019
Nebo School District	Providing Corrective and Supportive Feedback	2017
Utah School Districts	Planning for Engaged Learning	2015
Nebo School District	Curriculum Planning-Scope and Sequence	2014
Boise State University	Belief and Doubt in American Fiction	1998

HONORS AND AWARDS *Year*

UVU Student Advocate Award-Outstanding Faculty Advocate of the Year	2020
Recognition for Concurrent Enrollment Education Expansion	2016
Teacher of the Year (School Level-8th Grade Team Award)	2011
Teacher of the Year (School Level-Individual Award)	2010
"Governor's Invitation" students invited to decorate Governor's Mansion for Christmas	2009
Student-led "Town Hall Meeting" with Governor, Attorney General, and Senators	2008-2010

PROFESSIONAL COMMITTEES AND SERVICE *Year*

Utah Valley University, Scholarly Creative Activities Committee	2022
Utah Valley University, Writing Enrichment Committee	2022
Utah Valley University, School of Education Accreditation Team	2022

Utah Valley University, Service-Learning Faculty	2019-Current
Nebo School District, *New Teacher Mentorship (English)*	2014-2017
Nebo School District, *Teacher Growth Plan Evaluation Committee*	2014-2017
Nebo School District, *ACT Preparation Training (teachers)*	2016-2017
Canyons School District, *District ELA Curriculum Alignment Committee*	2014
Nebo School District, *Student Educational Travel*	2010-2013

PROFESSIONAL ORGANIZATION MEMBERSHIP

Association for Supervision and Curriculum Development (ASCD)
American Association of University Professors (AAUP)
National Council of Teachers of English (NCTE)
International Literacy Association (ILA)