

# Performance Assessment of Aspiring School Leaders Grounded in an Epistemology of Practice: A Case Study

This manuscript has been peer-reviewed, accepted, and endorsed by the International Council of Professors of Educational Leadership (ICPEL) as a significant contribution to the scholarship and practice of school administration and K-12 education.



**Jessica E. Charles**  
*Bank Street College*

**Rebecca Cheung**  
*University of California, Berkeley*

**Kristin Rosekrans**  
*University of California, Berkeley*

*There is increasing interest in the field of leadership preparation about the opportunities that robust performance assessments may provide to capture and evaluate the complexity of school administrators' work. Heretofore, the conversation about administrator performance assessment in leadership preparation has mainly centered on the development and impact of large statewide assessments that grow out of a Cartesian epistemology of individual knowledge possession, in which individuals must demonstrate mastery of a set of static knowledge and skills. We analyzed the characteristics of a performance assessment system that deliberately accounts for the organizational complexity of practice and knowledge generation in its design. Candidates are assessed by faculty and coaches on state-wide and program standards, but instead of producing evidence of their practice as individuals, they are assessed within simulated practice-based scenarios that require them to both draw on their extant individual and collective knowledge and build and act on new knowledge as they move through the simulation. Our analysis enables us to dimensionalize issues related to state mandated performance assessments and their implementation by preparation programs.*

There is increasing interest in the field of leadership preparation about the opportunities that robust performance assessments may provide to capture and evaluate the complexity of school administrators' work. For example, major efforts have been invested into the development and adoption of tools designed to measure the effectiveness of practicing school leaders such as the Vanderbilt Assessment for Leadership in Education (Val-Ed) and Comprehensive Assessment of Leadership for Learning (CALL). Both the Val-Ed and the CALL are quantitative tools that collect anonymous, evidence-based feedback from multiple stakeholders. The Val-Ed allows results to be organized according to a leadership framework and/or evaluation standards while the CALL focuses on providing actionable, formative feedback including customized strategies and action plans for school improvement (CALL, 2018; IOEducation, 2018). These and other valid and reliable assessments have been integrated into the fabric of principal evaluation throughout the nation.

Related to this practice, several states, such as California, Connecticut, Florida, and Massachusetts, have adopted or begun the process of adopting performance assessments, often called administrator performance assessments or APA, to measure the competence of aspiring school leaders in preparation programs. This trend follows a long-standing practice in teacher education, particularly the relatively rapid adoption of the edTPA by multiple states to assess the readiness of novice teachers to enter the classroom, performance assessments are being used to provide accountability for teacher licensure (Au, 2013; Sato, 2014; Price, 2016). Through these new performance assessments, states seek to assess leadership candidates' preparedness for domains such as vision for student achievement, instructional leadership, observation and mentoring of teachers, and engaging parents and other stakeholders. Understandably, these policy decisions have resulted in deep and significant psychometric and assessment design investments related to how performance assessments can be designed for accountability purposes in a valid and reliable manner (Cizek & Bunch, 2007; Meherens, 1992; Messick, 1995).

The leadership preparation field has long utilized assessments that help them determine how well their candidates are able to engage in leadership behaviors that will lead to successful outcomes for their schools. And, performance assessments have been used within programs to measure the capacity of leadership candidates for many decades (Wendel & Uerling, 1989; Wendel & Sybouts, 1988). Generally speaking, a performance assessment can assess the performance of any skill or area of knowledge across a range of less authentic to very authentic contexts (Palm, 2008; Haertel, 1999). For example, a culinary student might be asked to prepare an egg souffle in a test kitchen where he or she will be scored on the quality of the souffle he produces. Such a performance assessment would help a scorer know whether or not the student had acquired the skills to effectively prepare souffles. At the other end of the spectrum, a student could be assessed on his ability to work with an entire kitchen staff to prepare and serve an egg souffle for multiple customers at a busy restaurant. Under those conditions, the student's ability to prepare the souffle under the unpredictable and complex circumstances of a restaurant kitchen would also be assessed. Similarly, performance assessments of educational leaders can range from assessing important, but discrete, tasks, such as creating a meeting agenda, to assessing how well a principal leads a meeting amidst systemic pressures, in spite of organizational constraints, and within a network of human relationships. In this way, authenticity in a performance assessment is not dependent on whether or not it happens in a school setting, but, rather, on the degree to which it surfaces complex organizational conditions under which leadership tasks must be performed.

Professional and vocational preparation programs regularly use formative and summative performance assessments to determine how well their candidates perform discrete and integrative tasks under varying levels of uncertainty. For example, a medical professional's ability to perform

a physical exam can be assessed using a simulator or visual inspection (Johnson, 2007, Rose, 1999), architects are assessed as they learn to account for the slope of the land on which they are building (Schön, 1987), and nurses are assessed on their developing sense of salience, or ability to pay attention to the important aspects of a patient's care (Benner et al., 2012). School leadership candidates, like these professionals, can be assessed in the field by their field supervisors or leadership coaches as they are becoming leaders through observations and conferences, and are also assessed within their course and program structures through various methods such as traditional papers and projects. In addition, it is common for candidates to make presentations, a form of performance assessment. Thus, what is new about statewide performance assessments is not that leadership candidates are being assessed or that performance assessments are being used to conduct the assessment. Instead, it is the external nature of the assessment, which calls for students to provide evidence of their practice to an external and blind scorer who is not familiar with the candidate's school context or their program that is new, as well as the high stakes use of the assessment in the licensure process.

### **Purpose and Context**

The central question of this paper centers on the design of state-mandated administrator performance assessments (APA) for licensure. Specifically, what are the assumptions and orientations that affect performance assessment design for aspiring leaders? And, how do those assumptions and orientations manifest in the development and goals? We will focus on the California administrator performance assessment (CalAPA), the first statewide standardized measure of readiness for aspiring administrators in California that will be fully implemented in 2019. The recent decision in California to introduce an APA was strongly influenced by the rapid adoption of the edTPA and its proponents. Building from the theory that a performance assessment of teachers could ensure a baseline of quality in the workforce, policymakers advocated to extend this type of assessment to burgeoning school leaders (Fensterwald, 2012). In fact, the California Teacher Credential Commission, the agency in the executive branch of the California state government serving as the official accrediting body charged with overseeing all of the licensing and credentialing of professional educators in the state, specifically stated that one of the intents of its new California administrator performance assessment (CalAPA), is “to ensure a minimum threshold of leader readiness rather than to define exemplary practice” (CTC minutes, 2015).

The CalAPA is structured around tasks situated in three leadership cycles that are completed at three different periods during a candidate's preliminary credential program. Each task focuses on the roles and responsibilities of today's education leaders, using an investigate, plan, act, and reflect leadership sequence. Completion of each task requires that candidates either be in a school site-placement or have access to a school site where they can complete the work necessary for the CalAPA. The assessment comprises the following three leadership cycles focused on school site level work:

- Cycle 1: Planning School Improvement — Conducting data-based investigations, and planning and facilitating collaborative data inquiries that support equity and school improvement.
- Cycle 2: Facilitating Professional Learning — Facilitating collaborative learning among a small team of teachers to improve student learning.
- Cycle 3: Supporting Teacher Growth — Coaching an individual teacher to improve teaching and learning.

Emphasis on multiple modalities for evidence across these three leadership cycles allows candidates to submit evidence in various formats: annotated video, written plans for implementing academic priorities, observation notes and feedback on teaching practice, and narrative responses and reflections about practice. Elements requiring video must be directed, specific, and annotated (Kearney et al, 2018; CALAPA, 2018).

First, we will describe the impact of different epistemologies on performance assessment design. Then, we will describe the nature of leadership enacted and the related implications for assessment design. Finally, we provide performance assessment examples from UC Berkeley's Principal Leadership Institute that illustrate how assessments can be grounded within an epistemology of practice and discuss implications for preparation programs in the context of mandated state assessments.

### **Epistemologies of Organizational Learning**

Scott Cook and John Brown's theory of organizational learning (1999) distinguishes between an epistemology of possession and an epistemology of practice. Specifically, Cook and Brown argue that organizations, and individuals within organizations, learn as a system. They see the Cartesian perspective as limited by its individualistic approach to understanding knowledge creation, and argue that to fully investigate how individuals and organizations learn, one must account for individual knowledge *and* group knowledge, as well as explicit *and* tacit forms of knowledge. According to Cook and Brown, there are four types of knowledge: explicit individual knowledge, explicit group knowledge, tacit individual knowledge and tacit group knowledge. Individual knowledge is what one person personally possesses, while group knowledge is what people know together. Moreover, explicit knowledge is what we know that can be named, while tacit knowledge is what we know that is not easy to communicate to others but is vital to the enactment of complex practice.

They further argue that while none of these types of knowledge can be transformed into the other, they do work in what they label a "generative dance" to produce new knowledge. They call that "knowing." In the moment, individuals, who are part of larger organizations, draw upon what they know individually and collectively to respond to problems of practice. School leaders, for example, likely draw from explicit knowledge of theoretical perspectives, school data and state standards, while at the same time drawing on their tacit knowledge of how to navigate collegial relationships, or how to enact authority given their gender or racial positioning within a particular school context. The way moment-to-moment decisions are enacted are then a result of what the leader knows about, what he or she knows how to do, and what that leader knows about how to lead within his or her school at a particular time and place. In assessing readiness for school leadership, then, it may be more important to assess new leaders' capacity for "knowing," than to find out what they "know."

Cartesian perspectives dominate many assessment designs. Examples include multiple choice exams, short answer responses, and the individualized nature of the assessments themselves. Like most traditional tests, statewide administrator performance assessments such as the CalAPA grow out of a Cartesian epistemology of individual knowledge possession, in which individuals must demonstrate mastery of static knowledge and sets of skills. While the CalAPA requires candidates to engage with their colleagues at a school site, the submission items are artifacts submitted after the fact. Meaning, while they ask candidates to capture their practice in

organizational settings, the content of the assessment tasks rely entirely on the individual being assessed to select a video clip or clips, provide analysis and present that to the scorers. Scorers then rate individuals for both their performance and their individual ability to respond to the tasks within the assessment. Within the current design, the CalAPA largely replicates the traditional multiple choice and constructed response exams by substituting video evidence for information that previously would have been collected in writing. In essence, many of the limitations that traditional paper-and-pencil tests have posed for assessing leadership behaviors are replicated albeit with expanded menus of artifacts. To summarize, although performance assessments such as the CalAPA place a clear value on practice, they are generally built from an epistemology of possession, partly because they rely on materials, videos and artifacts filtered by the candidate, which he or she curates to meet the given standards. For example, there is strong potential for the selection of non-representative video clips, inadequate explanation or consideration of contextual factors, and the inability for scorers to see the practice in the video clips they are presented as part of a larger system of practice, embedded in specific organizations and communities (Haertel, 1999). This approach privileges individual knowledge and explicit knowledge, such as written reflections on practice, rather than the enactment itself.

Performance assessments designed to capture discrete individual knowledge and practice do not sufficiently take into account the complexity of leader practice. Because school leaders are embedded within multiple organizational layers, including the district, the school and various other professional groups, assessing their development as individual leaders should account for how they build and use knowledge in interaction with those organizational layers.

What do performance assessments grounded in an epistemology of *practice* look like? How do they benefit candidates and programs differently? We provide an example in use at the University of California, Berkeley. Developed by the Principal Leadership institute, their Assessment Center model deliberately accounts for the organizational complexity of practice and knowledge generation in its design. [Important Note: the PLI has been refining its performance assessment practices over two decades. Through this time, they have continued to use the name Assessment Center. We ask the reader to suspend assumptions about the term that may be related to earlier iterations of performance assessment in the field.] In this system, the assessment process that leaders-in-training experience looks quite different from the newly developed Cal APA. Candidates are assessed by faculty and field supervisors (called coaches) on state-wide and program standards, but instead of producing evidence of their practice as individuals, they are assessed within simulated practice-based scenarios that require them to both draw on their extant individual and collective knowledge and build and act on new knowledge as they move through the simulation. Assessment Center is a case worthy of analysis because it illuminates how an assessment constructed from an epistemology of practice, rather than an epistemology of possession, can work in the service of candidate and program learning, as well as for the development of the larger field of school leader preparation. Specifically, Assessment Center reflects an epistemology of practice for three reasons: 1) its focus on “approximations to practice” simulations (Grossman et al., 2009), which require candidates to engage in enactment of leadership, drawing on both tacit and explicit knowledge, 2) the emphasis that it places on practicing distributed leadership (Spillane, 2012), in which group knowledge, not solely individual knowledge, is assessed (this reflects a recognition of professional knowledge as embedded in the organizational relationships of the school and educational context); and 3) the orientation to ongoing program and professional learning that the assessment embodies.

## University of California, Berkeley Principal Leadership Institute Assessment Center

Founded in 1999, UC Berkeley’s Principal Leadership Institute (PLI) has three areas of work: preparation, induction, and leadership outreach. All programs are designed based on the principles of equity and social justice and focus on improving education for the most vulnerable and historically underserved public school students. In 19 cohorts, Berkeley PLI has prepared over 600 educational leaders who are 50% students of color, 95% working in public education, and 88% working in the Bay Area. The preparation program is a rigorous 15-month MA program for working teacher leaders who are interested in pursuing formal leadership as a school administrator that includes the preliminary licensure requirements.

One of the hallmark practices of the program are day long performance assessment events, known as PLI Assessment Center. Unlike many performance assessments or earlier models of assessment centers, the current PLI Assessment Center does not rely on artifacts of practice, but, rather, creates opportunities for candidates to simulate deliberate aspects of practice to demonstrate individual and group knowledge. Specifically, Assessment Center consists of two major performance events, during which candidates participate in simulated scenarios - that are embedded in an overarching case of a fictional school - that approximate the real work of school leaders. The first Assessment Center occurs at the halfway point of the 15-month program and requires candidates to work individually and in teams on scenarios related to instructional leadership and interpreting data for the purpose of school improvement. The second Assessment Center occurs at the three-quarter point of the program and centers on a mock expulsion hearing as well as analyzing school wide strengths and needs from the perspective of a new principal, in which candidates must demonstrate multiple competencies related to legal and policy content as well as systemic analysis. Both events also require them to showcase individual and group-related skills and knowledge aligned to the coursework they have completed up to that point in the program.

Table 1  
*PLI Assessment Center Map*

<b>PLI Assessment Center</b>			
<b>Semester</b>	<b>Leadership Competencies</b>	<b>Activities</b>	<b>Assessors</b>
Fall	<ul style="list-style-type: none"> <li>● Instructional Leadership</li> <li>● Supervision and Evaluation of Teaching</li> <li>● Interpreting Data for School Improvement</li> </ul>	<ul style="list-style-type: none"> <li>● Individual and group case analysis</li> <li>● Group analysis of instructional coaching</li> <li>● Group presentation of a plan of action</li> <li>● Individual analysis of teaching</li> <li>● Post observation teacher conference simulation</li> </ul>	<ul style="list-style-type: none"> <li>● Instructors</li> <li>● Field Supervisor/Coaches</li> <li>● Peer Observation</li> <li>● Selected Outside Guests</li> </ul>
Spring	<ul style="list-style-type: none"> <li>● Educational Law related to</li> </ul>	<ul style="list-style-type: none"> <li>● Group presentation for mock expulsion hearing</li> </ul>	<ul style="list-style-type: none"> <li>● Instructors</li> <li>● Field</li> </ul>

	expulsion hearing <ul style="list-style-type: none"> <li>● Interpreting Data for School Improvement</li> <li>● Organizational/Systemic Analysis</li> <li>● School improvement planning</li> </ul>	<ul style="list-style-type: none"> <li>● Individual and group case analysis</li> <li>● Integration of analysis into specific leadership strategies (i.e. Professional development plan development or interview protocol)</li> <li>● Group presentation</li> </ul>	Supervisor/Coaches <ul style="list-style-type: none"> <li>● Peer Observation</li> <li>● Selected Outside Guests</li> </ul>
--	---	--	--

We conducted a year-long descriptive study of the PLI Assessment Center system that analyzed the stated purposes, the design of the Assessment Center model, and the experiences of students, faculty and staff during Assessment Center activities. We highlight three findings that demonstrate the affordances of a performance assessment based in an epistemology of practice. First, Assessment Center creates opportunities for candidates to demonstrate tacit knowledge of leadership, which is difficult to surface in traditional written exams and papers, and perhaps, even through written reflection on aspects of one’s own practice. Second, Assessment Center accounts for group knowledge as an essential element of leadership, by creating both group activities and group assessments. Third, Assessment Center creates opportunities for program learning and refinement, because instructors, coaches and the director of the program are closely involved in the creation of the scenarios, and are expected to make adjustments to the individualized education of candidates, as well as the program overall, as a result of participating and collecting data from the assessment.

**Data and Methods**

**Case Selection**

We used an explanatory case study methodology to study an example of an exemplary leadership assessment practice (Yin, 2018; Creswell, 2014). Berkeley PLI’s Assessment Center was selected as a case of authentic administrator performance assessment based in an epistemology of practice because of the deliberate construction of “approximations to practice” which Pamela Grossman and her colleagues defined as “opportunities to engage in practices that are more or less proximal to the practices of a profession,” (2009) as well as the emphasis on group knowledge generation embedded throughout Assessment Center (Cook & Brown, 1999). Specifically, we set out to understand how Assessment Center approached eliciting candidate knowledge for the purpose of assessment. We believed that an assessment concerned with authenticity would be designed with rich opportunities for candidates to display professional “knowing,” which would be visible through the assessment activities themselves, and the interpretations of those activities by the participants, including candidates and assessors. Our goals were to explain how this exemplary program approaches performance assessment, and how that assessment works to both evaluate *and* build candidate, coach and program-level knowledge.

## Research Questions

We asked three research questions:

- 1) How is Assessment Center designed to assess students' individual tacit leadership knowledge?
- 2) How is Assessment Center designed to assess the group knowledge of leadership candidates?
- 3) How does the design of Assessment Center help the program respond to individual and programmatic needs?

## Data collection

Our data consist of interviews before and after each Assessment Center with four PLI candidates/students (n=7) as well as three coaches (n=6)<sup>2</sup> Additionally, we observed and took field notes during each Assessment Center cycle and collected artifacts, including assignments, coach feedback forms, and video clips to contextualize our understanding of the process.

Leadership candidate participants were selected based on a range of factors, including gender, race, experience level and performance in the program, in order to gather a wide variety of perspectives. The coaches we selected as participants had several years of experience with Assessment Center, so their answers would reflect a perspective developed out of deep familiarity with the authentic assessment. Leadership candidates participated in semi-structured interviews to elicit their understanding and experience of Assessment Center. The interviewer asked the following questions, but followed up with probing questions to help her better understand the perspective of the interviewee:

- How did Assessment Center go for you?
- Choose a moment that was meaningful. Tell us about it and explain what you took from it.
- What will you take away from Assessment Center, if anything, as you proceed in your development as a leader?

Coaches were asked to ground their answers in their work with specific candidates, in order to elicit the most specific information possible,. The interviewer followed up with appropriate probing questions as they responded to the following prompts:

- Please think about one coachee in particular and what experiences and observations from the Assessment Center, if any, you will use in your coaching with that student.
- Since Assessment Center provides a different environment from the one in which you usually observe your coachee, how, if at all, does your participation in Assessment Center inform your understanding of your candidate's leadership development?

## Data Analysis

We analyzed our data in four phases, which enabled us to attend to emerging themes related to our theoretical frame. In our first stage, we organized our interview transcripts, field notes and documents into three categories: evidence of assessment of tacit knowledge, explicit knowledge, and group knowledge. We used our video footage to help contextualize our other data, and as a

---

<sup>2</sup>Coaches are experienced educators who, as part of PLI, are assigned four to five PLI candidates/students for the duration of the program to guide them in applying theory to practice in their work sites and contribute to assessing their progress along with the instructors.



reference point to clarify questions that arose during our analysis. In our second phase, we coded for evidence of the stated purposes of Assessment Center, its design, and the experiences of the candidates and coaches (who served as scorers). In our third phase, we analyzed our interview transcripts for the meaning participants made of the process as they experienced it. We triangulated our data across multiple participants (candidates, coaches and program staff) and across methods (interviews and document analysis) to ensure accuracy of our results (Patton, 1999; Yin, 2018). In a final stage, we integrated our analysis of these categories to present a holistic picture of Assessment Center in response to our research questions.

## Findings

Our analysis enables us to dimensionalize the aspects of Assessment Center that demonstrate its strength in assessing individual candidates in authentic scenarios, and its focus on assessing tacit, as well as emerging group knowledge as it unfolds in these scenarios. We were also able to see how Assessment Center contributed to ongoing program development and individualized feedback and support for leadership candidates.

### Eliciting Tacit Knowledge

It is clear that Assessment Center requires candidates to put into practice explicit and tacit knowledge to grapple with the leadership scenarios with which they are presented. Candidates are often required to use explicit knowledge that they have gained during courses by citing texts and data that they have encountered. They also draw on school law and appropriate procedures and protocols for interacting with colleagues and students during Assessment Center. How they use these pieces of explicit knowledge, however, requires them to draw on tacit knowledge for enactment of leadership in the moment.

Cook and Brown's conception of tacit knowledge is helpful here. They describe it as knowledge that is gained through the generative dance of knowing, but which the individual retains in order to enact it again. They give an example of the knowledge needed to ride a bicycle to illustrate their point. When a person learns to ride a bicycle, they argue, they have explicit knowledge of how a bicycle works. However, it is not until they actually get on and learn to ride that a tacit understanding of how their own body feels and works while riding is developed. While a bicycle rider is only "knowing" how to ride a bicycle in the moment of riding, a tacit knowledge of how to enact bicycle riding is retained by the rider for use at a later time.

For leadership candidates, it is hoped that tacit knowledge of leadership is gained through course assignments that require approximations to practice, fieldwork experiences, and elsewhere in the program. These experiences are designed to cultivate tacit knowledge in the candidate, which is then called upon during Assessment Center. We see this through the candidates' reports that the activities feel authentic and require immediate action, thereby necessarily calling upon both explicit and tacit knowledge for leadership enactment. We also see evidence that tacit knowledge is required by the activities in Assessment Center through the coaches' comments about what they are able to learn about their candidates' development, by assessing their enactment of leadership competencies in real time.

**Candidates demonstrate tacit knowledge through realistic leadership experiences.** Leadership candidates remarked on the authenticity and relevance of the Assessment Center experience during all of their post-Assessment Center interviews we conducted with them except

for one. We coded notes for words such as “real” or phrases that otherwise indicated simulation of leadership practice such as “doing something we would do as leaders.” In the interviews that mentioned authenticity and relevance, candidates mentioned this between one and sixteen times during each interview, averaging five mentions per interview. We also coded for places where the candidates judged the experience to be meaningful learning for their leadership development. Candidates described Assessment Center in such terms in nearly every interview, totaling eighteen times in all, averaging two times per interview. Using data triangulation (Patton, 1999; Yin, 2018), we confirmed this finding through coach interviews: a there were a total of sixty mentions of authenticity of the Assessment Center experience across all coach and candidate interviews, and a total of twenty-four descriptions of it as a meaningful learning experience across that interview set.

Moreover, the candidates reported feeling concerned about how well they would perform, indicating that the experience felt consequential to them, despite it being program-embedded and not conducted by a standardized purveyor of professional assessments, such as Pearson or Education Testing Service. Before Fall Assessment Center, most of the candidates we interviewed expressed nervousness about the event, while prior to the Spring Assessment Center, those nerves had primarily been channeled into thorough advanced preparation. More than one candidate discussed having felt quite anxious before Fall Assessment Center, but less nervous and more interested in availing themselves of the learning opportunity during the spring. Mentions about nervousness numbered six across the interviews, while mentions of working to thoroughly prepare numbered ten. One coach also mentioned this phenomenon, stating: “With the initial assessment that we do in the fall, students...go into that one a little more, let’s say, apprehensive. They’re nervous, they go to it with a different mindset...The comment I heard from a lot of them was, we’re ready for this [Spring Assessment Center], we are prepared for this.” On the other hand, some participants also said that they were unable to prepare as much as they would have liked, given their work schedules and job searches. However, we believe this further confirms the finding that those candidates understood the importance of the assessment, despite feeling somewhat underprepared.

Table 2  
*Candidate interview response tabulations*

Interview code (n=9)	Total number of codes	Lowest occurrence within an interview	Highest occurrence within one interview	Average across all interviews, fall and spring
Real/Authentic (Parent code, no child codes)	49	1	16	5
Meaningful Learning Experience (Parent code, no child codes)	18	0	3	2
Approach to preparation (Parent code)	18	0	3	2
<i>Nervousness in</i>	6	0	2	.7

<i>anticipation of AC (child code Atp)</i>				
<i>Thorough preparation for AC (child code Atp)</i>	10	0	3	1.1
<i>Unable to prepare as much as desired because of other obligations (child code Atp)</i>	2 (spring only)	0	1	.2

The “realness” of the experience for candidates helped us see that Assessment Center requires candidates to call upon and create tacit knowledge as they accessed it during enactment. One student described it in this way:

*...I do really think that in a lot of ways I appreciate Assessment Center because it is authentic, it is an authentic assessment, and it feels real. It feels like you’re doing the work of a site leader, you’re doing the work of an administrator, and it’s not that theoretical piece.*

Another candidate discussed the value of enactment during Assessment Center as a means of eliciting knowledge she may not have otherwise tapped into. Her comment is reflective of many of the interviews with students, coaches and instructors who again and again explained the value of the realism of simulation exercises in which they participated.

*I just can’t say enough about how much our work as leaders in education rely on our ability to take information and quickly do something with it, and to present things in a way that makes people feel calm and empowered at the same time, and we have lots of different types of people and expectations and responsibilities, and you can’t get that from taking a test. You just, you can’t. You can’t just be given something and write down what I would say or whatever, because you’re always going to sound better on paper than you are when you’re having to talk to someone out loud and go through and respond to somebody and be quick on your feet. So I think it’s incredibly powerful to do the assessments this way and to give us real experiences that we can take with us into leadership. You couldn’t do that any other way. So that would be an add-on to me, just to take that away.*

Here the student points out the value she sees in the simulated experiences of Assessment Center. Her comment that “having to talk to someone” and “be quick on your feet” is preferable to being asked to “write down what I would say,” shows that candidates are required by Assessment Center to demonstrate their knowledge through their behavior, not just their written reflections, in real time. This student, like many of the others we interviewed, saw this as both an opportunity to demonstrate knowledge in a challenging performance environment, as well as a learning experience that helped her become a better leader. In other words, candidates saw Assessment Center as a moment of “knowing,” in which they called upon tacit and explicit forms of knowledge and created new knowledge through the “generative dance” in which Assessment Center required them to engage.

***Coaches see aspects of practice that were invisible before.*** Another theme that emerged from our interviews with coaches about the Assessment Center experience was an identification of

simulations as an opportunity to witness tacit knowledge in action, or identify tacit knowledge that was lacking in candidates. The tables below illustrate that while coaches highlighted different aspects of the Assessment Center experience in their interviews, they all emphasized the unique opportunity Assessment Center provided to witness candidates whom they were coaching (their “coachees”) perform aspects of a school leader’s role in a purposeful, but realistic context. This context enabled them to learn about their coachees, and attend to their leadership development through coaching. They stated in multiple ways that Assessment Center helped their students surface knowledge that they were unable to access through courses or even site visits, and that the constructed scenarios highlighted both strengths and weaknesses in candidates’ leadership skills that otherwise would have remained hidden from view. Interestingly, they often mentioned the importance of both the contingent and collaborative nature of the Assessment Center process, noting what they were able to learn about coachees as they related to their peers throughout performative group activities, such as the mock expulsion hearing and the case study discussion.

Table 3  
*Pre-Spring Assessment Center Coaching Interview Response Tabulations*

Pre-Spring AC Coach Interview coding (n=3)	Total number of codes	Lowest occurrence within an interview	Highest occurrence within one interview	Average across all interviews, fall and spring
Anticipated learning something specific about candidates during AC	8	1	5	2.6
Anticipated being able to “see” something new about a candidate during AC	5	2	3	1.7
Anticipated AC being an authentic learning event	8	1	5	2.6
Anticipated AC being an opportunity for candidates to learn important leadership skills	5	1	3	1.7
Anticipated AC being an opportunity for candidates to work collaboratively with others	7	0	5	2.3

Anticipated AC being an opportunity for candidates to contribute individual strengths to the whole group	2	0	2	.7
Anticipated AC being an opportunity to build candidate confidence	4	0	4	1.3
Anticipated using AC as a means of reflecting on coaching and to help the candidate reflect	4	0	4	1.3

Table 4  
*Post-Spring Assessment Center Coaching Interview Response Tabulations*

Post-Spring AC Coach Interview coding (n=3)	Total number of codes	Lowest occurrence within an interview	Highest occurrence within one interview	Average across all interviews, fall and spring
Mentioned learning something specific about candidates during AC	11	2	5	3.7
Mentioned being able to “see” something new about a candidate during AC	21	5	8	7
Mentioned AC being an authentic learning event	3	0	3	1
Mentioned AC being an opportunity for candidates to learn important leadership skills	1	0	1	.3
Mentioned AC being an opportunity for candidates to work collaboratively with	6	0	4	2

others				
Mentioned AC being an opportunity for candidates to contribute individual strengths to the whole group	2	0	2	.7
Mentioned AC being an opportunity to build candidate confidence	10	2	5	3.3
Described using AC to inform developmental coaching approach	12	3	5	4

In pre-assessment interviews, coaches described wanting to know more about their coachees through the process of Assessment Center, and in post-Assessment Center interviews, they described new insights they had gleaned about their leadership development. One coach described Assessment Center as a “different venue” in which you see candidates in a “different light.” This same coach had worried about a candidate’s ability to keep up with the coursework in the program saw her enact leadership knowledge during Assessment Center that hadn’t been visible to her before. After observing her performance during the mock expulsion hearing, she remarked:

*I was particularly struck by a... coachee, who is struggling in terms of keeping up with PLL, for a lot of reasons. A lot of extremely valid reasons... But she was just sort of at the top of her game, and she did the closing statement in the expulsion hearing, and she was terrific. She also took over facilitating her group when they were working on this case study. And so it's very re-affirming to see what incredible talent she has.*

An example from another coach pointed to Assessment Center’s power to assess tacit knowledge for leadership. The candidate was not struggling with coursework, but, rather, excelled in the traditional academic sense. Spring Assessment Center provided this coach with an opportunity to see this candidate’s leadership knowledge in action, rather than to rely only on her written expression of knowledge.

*It's interesting that watching her in small groups and what have you in my class, she contributed but she wasn't very outspoken when it came time to, let's share out. She didn't do a lot of that. And I was really impressed... She's a good student, don't get me wrong. She does really well on her paperwork et cetera, but watching her in her element, because she is the lead PD, and the way she handled it, she was confident, there was some humor there, she did an outstanding job. She's another one that stood out for me in that sense, because I was really impressed with the way she came across.*

However, codes were remarkably less frequent, perhaps because interviews focused more on the coaches’ experience of the event and how they used it for their own practice, for demonstrating

candidates' individual strengths in a group context (.7) and as a learning event for candidates (.3). If isolating particular leadership skills is a sole purpose of performance assessment, it will be important to tease out the elements of events such as those that facilitate program, coach and candidate "knowing," and those which demonstrate "knowledge."

### **Eliciting Group Knowledge**

Another aspect of Cook and Brown's organizational theory of knowledge accounts for the way in which individual and group knowledge work together to inform "knowing" of organizational actors. They argue that knowledge lives within organizations that is larger than individual knowledge that any one person possesses. School leaders do not work in isolation, but, instead, build knowledge for practice with those with whom they work and in the context of the organizational and professional expectations of their role. Assessment Center attends to group knowledge through both the design and the enactment of the activities. Candidates are required to work with others, by design, and are assessed as individuals and as a group. Candidates report growth in their leadership skills and perspectives through these activities.

**Group knowledge as a design element.** Though all activities in Assessment Center are designed to elicit and create group knowledge, group discussions and presentations are perhaps the clearest examples of this. During group discussions and presentations, candidates are expected to build and demonstrate knowledge for leadership as a group. Below is a description of a "Case Study Discussion Protocol." Candidates use this protocol in a group setting to discuss a case study of a leadership dilemma in order to surface the issues and challenges of school leadership in a particular context.

*The purpose of this case discussion is to provide an opportunity for your group to have an initial dialogue about the challenges faced by Ms. Violet and Franklin School. First, you will hear a short report of each group member's initial thoughts about the case as you were instructed identify in the preparation directions. Then, there will be time for open discussion. During this time, we urge you to continue to focus on the underlying issues and leadership challenges.*

*Each individual will have 3 minutes (12 minutes total) to identify the 2-3 most important issues at play in this case, relating them to the course concepts & literature. In order to ensure that each person has the opportunity to share their thoughts, the 3 minute limit on the "whips" will be firmly enforced.*

*The group will have 12 minutes to continue to discuss the case as a group, focusing on the underlying issues and leadership challenges. This open discussion will be left to your group to manage.*

*The instructor/s will have 5 minutes to share feedback and insights into the group's performance*

Figure 1. Assessment Center Case Study prompt

In this activity, candidates have the opportunity to both demonstrate their own knowledge of leadership and the literature they've encountered through coursework, and they are also being

assessed on their interactions with others and their capacity to build knowledge for leadership with others. Unlike some assessments or pedagogical activities that require candidates to discuss their individual knowledge in order to see what they know as individuals, this activity, by design, assesses the group's ability to organize itself for learning, and to build knowledge for leadership as a group.

After an initial share out, the group has several minutes to discuss the case with which they are presented. This is an open discussion, which is "left to your group to manage." Leaving the group to manage itself is not only a way to assess organizational skills, but also to see how would-be leaders position themselves in relation to other adults to create relationships oriented for group learning. After the discussion, the group is then given feedback from observers about how they worked as a group, not as particular individuals who are there to share individual knowledge.

Moreover, throughout student and coach interviews, several participants mentioned the critical role that the program director played not only in designing and requiring such exercises, but in creating the group configurations, as well. Students and coaches understood that the program director often grouped students who needed to work on a particular skill or who needed support or a push from a certain group within the cohort. Across the interviews, two students and two coaches described the program director's deep knowledge of her students and her purposeful approach to designing learning opportunities for each student. One student discussed her understanding that the program director had intentionally matched her with a fellow student whom she found intimidating during fall Assessment Center:

*So...we found out we were going to do...a role-play, so I found out that the person that was going to pretend to be the teacher while I was the administrator and had my planned conversation, the person who I had been set up with to be the disgruntled teacher was actually someone I had admitted to my program director that I am intimidated by, because she's really well-spoken... I admire her, but I definitely feel a little... Yeah. Worried around her that I'm going to mess up or say something... yeah.*

*Our program director, is amazing at making sure we have learning opportunities. She put us together, of course. So she was pretending to be the teacher who was disgruntled. She did a really good job of it. She called me judgmental at one point and a bunch of other things. But I just had to work through it. This kind of thing actually happens as a principal, and it happens in meetings with other teachers and it happens in life, so it was really good to have to remember to stay calm.*

Further, coaches reported that an important aspect of Assessment Center was being able to determine how well candidates were able to collaborate with others in authentic scenarios, which was an aspect of their practice that was difficult to see in the field or during coursework. They mentioned this thirteen times across their interviews.

***Students learn from one another during group activities.*** The group activities also present opportunities for candidates to build knowledge with others in the moment and to demonstrate group and tacit knowledge for leadership during Assessment Center. Students see their own knowledge as situated within a larger body of group knowledge, which is greater than them, but which they can access by working successfully within a group. For example, one student said:

*I think the one task that sort of stood out to me was the, when we had our group conversation surrounding the case study, I think the one thing that stood out, and it was mostly just that*



*we were, everybody sort of came in with their own perspective, and once we were sitting around the table and talking to each other about the case study, it was hard to imagine us having missed anything. Like, everybody brought up something that really meant something to them, and it created a really holistic image of what was going on. I was just impressed once we got rolling, how much people picked up on from the case study, and how many different pieces there were. Yeah. I was very impressed, because there was stuff that I missed, but somebody in the group had picked up on it clearly.*

Here the student describes the value in working on the case study with others because “it was hard to imagine us having missed anything.” He goes on to explain that within what appears to be a fairly well-organized discussion, building group knowledge is greater than the sum of its parts. As with many of the activities in Assessment Center, the process both unearthed candidates’ knowledge and helped them create new knowledge. Group activities such as these placed a value on what could be created from carefully orchestrated sharing, listening and reflecting together, by providing feedback on both content and process.

Coaches saw this, as well. In one interview, a coach thought about the Assessment Center as a place for her coachee to recreate his self-presentation within the context of the group activity. She said,

*I think that the groups are, the group responsibilities are where my coachees will be able to interact and engage with their fellow cohort members. And so in this dynamic, I’m hoping that they will be able to express themselves, articulate their ideas, and of course merge those in the group setting, so that it becomes a holistic presentation. Because I think [to] the path that they’re doing, the expulsion hearing and also the case study presentation, will allow them to present themselves in a way that they are more confident, and then I’ll be able to sense that their contributions are part of the entire group’s presentation. A couple of my, one of my coachees in particular, I know is a little bit shy about maybe asserting himself in a group, so I’m curious as to what his role will be in the group presentation, because there are some roles that are more prominent, others that are tangential, so I’m wondering how he’s going to surface in this group dynamic when they’re combining the two work groups and producing their presentations, where he stands in that setting.*

In her anticipation of the activity, she imagines how her candidate might “merge” his ideas and expressions with others to make a holistic presentation, and wonders how he will “surface” in the group dynamic, which indicates that the Assessment Center is opportunity is not only an opportunity to demonstrate what one knows, but to build what one is learning as one participates in the assessment, which happens within a group setting. Both the authenticity of the scenario and the group dynamics allow this tacit knowledge to build in the moment, and the knowledge that is both created and demonstrated in contingent on those factors.

### **Assessment for Organizational Learning**

Another way in which Assessment Center reflects an epistemology of practice is the built-in design for organizational learning. As candidates enact leadership through Assessment Center, knowledge is constructed by the program and its staff alongside the candidates. By interacting with candidates as they respond to the leadership scenarios with which they are presented, coaches and

instructors build knowledge about the candidates, and about their own coaching and teaching. Additionally, the program and Assessment Center, itself, learn from the experience and adapt.

***Assessment Center helps coaches and instructors develop their practice.*** Assessment Center is designed to both assess learning and to simultaneously create opportunities for learning. Coaches and instructors almost unanimously report that Assessment Center helps them learn about their students and their own teaching and coaching. For instance, one instructor noted that Assessment Center helps him prioritize particular aspects of leadership knowledge in his course during an interview after Fall Assessment Center:

*The activities are all collaborative and they require multiple task management and time management... And all of those skills are essential to high-quality educational leadership. And it reinforces my practice in the classroom to be spending time on those things.*

Then, again, this same instructor described a similar sentiment after Spring Assessment Center:

*I find Assessment Center to be incredibly valuable as an assessment tool for me to assess my practice, and again, the course design, and the structures that we use to guide the students in a very short time, in 14 months, from being teachers to being credentialed, authorized, practicing administrators. You know? It's a scary responsibility.*

Another instructor discussed the value in meeting with other coaches and instructors to discuss the candidates' progress during Assessment Center. This meeting is built into the design of Assessment Center, in order for the coaches and instructors to calibrate for the assessment activities themselves, and for them to hone their approach moving forward as individuals and as a program. She said,

*I think I would just underscore the value that I as an instructor gain from the feedback session with the field supervisors that we do during our lunch break. Getting the thematic feedback from the other people who are involved in the process is just, I mean, I've never had that experience as a teacher before, and it's really meaningful for me and my practice.*

***Assessment Center informs program development.*** Assessment Center not only informs the individual practice of coaches and instructors, it also feeds into a cycle of group learning by the entire program. Over time, the program and Assessment Center itself are changed in response to the organizational learning that happens by various constituencies within the Principal Leadership Institute. For example, a few years ago, instructors were disappointed in the number of students who did not use open-ended questions in the simulated post-observation conference. The curriculum was subsequently revised to include more practice and coaches followed up with the individuals directly to ensure more practice in their questioning strategies. The next year, instructors noted improved rates of questioning strategies.

## **Discussion and Implications**

Our findings have important implications for the development of state mandated administrator performance assessments and the programs mandated to implement them. First, this study expands our thinking about the nature of authentic performance tasks and assessment experiences. Using the epistemology of practice frame allows us to see how deliberately designed approximations to practice may have some advantages for assessing candidates' "knowing" over the more widely-used practice of assessing video slices of practice and accompanying candidate reflections. Distinguished

from the individualized nature of typical state performance assessments, Assessment Center affords the program the ability to generate scenarios that require candidates to draw upon individual and group knowledge, and the enactment of practice that is visible to assessors is not mediated through the lens of the candidate who may select the slice of practice to submit. While these approaches both attempt to capture authentic practice, it may be paradoxically true that intentionally designed approximations during which candidates must enact leadership competencies in the presence of coaches and instructors are more suited to revealing a candidate's tacit knowledge for practice than a video of his or her practice in a live setting. Our findings indicate that it would be interesting to compare the dimensions of knowledge for leadership enactment that are visible in a live performance assessment such as Assessment Center and those which are visible through a documented experience upon which a candidate reflects.

Second, the case of Berkeley's Assessment Center raises questions about how current state mandated administrator performance assessments account for group knowledge. Though video clips and descriptions of fieldwork, which are common artifacts required by larger scale assessments currently in use, are reflective of the type of work done with and among other organizational actors, the value that Assessment Center places on both leadership knowledge for working within groups, as well as the knowledge created together by groups, seems difficult to replicate outside a simulated or real-time administrator performance assessment. Because organizational knowledge is key to administrator knowledge and successful leadership, it would be useful to consider the extent to which APA models embrace an epistemology of practice or possession. Given the professional knowledge that is needed for leadership, which draws on both tacit and group knowledge, it may be useful to consider accounting for these in the designs of new APAs.

Third, because Assessment Center is not only a powerful learning tool for candidates, but for their instructors and field supervisor/coaches, as well, it is crucial that the relationship between the administration of an APA and the principal preparation program are closely examined. For example, Assessment Center is embedded into the life cycle of a preparation program, which allows program leadership, instructors and coaches to learn and respond during the program to benefit the learning of candidates. In the CalAPA, for example, each of the three tasks will be scored by separate scorers. In that configuration, the assessors do not have the ability to see growth over time. However, within the Assessment Center model, it is only natural to see the progression of performance over the course of the day. Furthermore, candidates benefit from having assessors who evaluate their performance in Assessment Center and develop their leadership practices during the course of the program. They have more meaningful feedback that is aligned within their program and triangulated to other program assessments. While the CalAPA uses the use of blind external scoring to limit assessor bias, it may also limit the ability of the assessor to give deep, meaningful, and timely feedback.

Fourth, using standardized performance assessments across multiple programs statewide (in California, there are over 60 programs serving extremely different contexts), that is administered by a national testing company, requires the developers to decontextualize and genericize the assessment in ways that can preclude programs, instructors, coaches and students from a more authentic, seamless and inclusive feedback loop. Unlike standardized administrator performance assessments, Assessment Center does not narrow feedback to a numerical score on a specific standard that is provided approximately 6-8 weeks after submission (of course, the submission can be written with a large delay after the actual activities have taken place). In the end, the biggest constraint in creating truly authentic assessments might be the goal of efficiency and attempting to do it "at scale," rather than supporting and building the capacity of individual programs to design and implement

assessments based on an epistemology of practice. As Cohen and Ball elaborate in their paper *Educational Innovation and the Problem with Scale* (2007), “To solve the problem of ‘scaling up’ requires ‘scaling in’- by this we mean developing designs and infrastructure needed to support effective use of an innovation. That, in turn, requires consideration of the problems that have made some sorts of innovation difficult...Scale is relative not just to the universe of possible implementers, but to the scope and depth of what must be done to devise and sustain change.”

Fifth, leadership preparation programs bear the ultimate responsibility to manage and balance the various mandated and non-mandated assessment strategies for their candidates. In the case of Berkeley’s PLI program, they continue their Assessment Center practices alongside the required CalAPA activities. If, in fact, the trend to institute APAs continues and more states use statewide exams to provide minimum competency accountability for the field, what investments do leadership faculty need to make to ensure the inclusion of assessments that more authentically approximate practice in their preparation programs? What are the differences in preparation between those who meet the minimum standard of the APA and those who enroll in programs that engage in more authentic assessment practices?

Finally, unlike other professional fields such as medicine, statewide assessments in education are expensive endeavors for programs and practitioners without the potential of substantial salary increases after licensure. They are costly to aspiring leaders (typically \$350-500/per exam) who are already personally responsible for their licensure expenses and potentially redirect resources from programs given the high stakes nature. In the worst case scenario, external performance assessments raise the stakes, while adding costs and potentially burdening individual school leadership candidates and their preparation programs. How can policy makers and programs work together to ensure that external performance assessments effectively improve the preparation of aspiring leaders, build the capacity of preparation programs, and ensure a stronger leader workforce that all children, especially vulnerable and historically underserved youth, deserve?

## References

- Au, W. (2013). What's a Nice Test Like You Doing in a Place Like This? The edTPA and Corporate Education Reform. *Rethinking Schools*, 27(4), 22-27.
- Benner, Sutphen, Leonard, Day (2010). *Educating Nurses: A Call for Radical Transformation*. San Francisco, CA: Jossey-Bass.
- CALAPA (2018) Retrieved from: [https://www.ctcexams.nesinc.com/TestView.aspx?f=HTML\\_FRAG/CalAPA\\_TestPage.html](https://www.ctcexams.nesinc.com/TestView.aspx?f=HTML_FRAG/CalAPA_TestPage.html).
- CALL (2018) Retrieved from <https://www.leadershipforlearning.org>.
- Cohen, D. K. & Ball, D. L. (2007). Innovation and the problem with scale. In B. Schneider & S. McDonald (Eds.) *Scale up in Education Ideas in Principle* (Volume I) (pp. 19–36). Lanham, MD: Rowman & Littlefield.
- Commission on Teacher Credentialing. (2015, December 3-4). *Commission Meeting Agenda*. Commission on Teacher Credentialing Archives, Sacramento, CA.
- Cook, S. D., & Brown, J. S. (1999). Bridging epistemologies: The generative dance between organizational knowledge and organizational knowing. *Organization science*, 10(4), 381-400.
- Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative sociology*, 13(1), 3-21.
- Creswell, J. W. (2014). *A concise introduction to mixed methods research*. Thousand Oaks, CA: Sage Publications.
- Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage Publications.
- Darling-Hammond, L., LaPointe, M., Meyerson, D., & Orr, M. T. (2007). *Preparing School Leaders for a Changing World: Lessons from Exemplary Leadership Development Programs*. School Leadership Study. Executive Summary. Palo Alto, CA: Stanford Educational Leadership Institute.
- Darling-Hammond, L., & Snyder, J. (2000). Authentic assessment of teaching in context. *Teaching and teacher education*, 16(5-6), 523-545.
- Davis, S., Kearney, K., Sanders, N., Thomas, C., & Leon, R. (2011). *The policies and practices of principal evaluation: A review of the literature*. San Francisco, CA: WestEd.
- Fensterwald, J. (2012, December 10). Credentialing commission imposes tougher test to become school administrator. Retrieved from <https://edsources.org/2012/credentialing-commission-imposes-tougher-test-to-become-school-administrator/23925>
- Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. (2009). Teaching practice: A cross-professional perspective. *Teachers College Record*, 111(9), 2055-2100.
- Haertel, E. H. (1999). Performance assessment and education reform. *Phi Delta Kappan*, 80(9), 662.
- IOEducation (2018). Retrieved from: <https://ioeducation.com/what-you-need/professional-growth/principal-leadership>
- Johnson, E. (2007). Surgical simulators and simulated surgeons: Reconstituting medical practice and practitioners in simulations. *Social Studies of Science*, 37(4), 585-608.
- Kearney, K., Mattson, H., Cheung, R., & Makkonen, R. (2018) *The Initial impact of the Cal APA policy on preparation programs*. San Francisco: WestEd.

- Meherens, W. (1992). Using Performance Assessment for Accountability Purposes. *Educational Measurement: Issues and Practice*, 11(1), 3-9.
- Messick, S. (1995). Standards of Validity and the Validity of Standards in Performance Assessment. *Educational Measurement: Issues and Practice*, 14(4), 5-8.
- Murphy, J. F., Goldring, E. B., Cravens, X. C., Elliott, S. N., & Porter, A. C. (2007). The Vanderbilt assessment of leadership in education: Measuring learning-centered leadership. *Journal of East China Normal University*, 29(1), 1-10.
- Orr, M. T., Pecheone, R. L., Shear, B., Hollingworth, L., & Beaudin, B. (2016). Massachusetts Performance Assessment for Leaders (PAL) technical report: Summary of validity and reliability studies for 2014-15 field trial of PAL. New York: Bank Street College of Education. Retrieved from <http://www.doe.mass.edu/pal/TechnicalReport.docx>.
- Palm, T. (2008). Performance assessment and authentic assessment: A conceptual analysis of the literature. *Practical Assessment, Research & Evaluation*, 13(4), 1-11.
- Patton, M. Q. (1999). Enhancing the quality and credibility of qualitative analysis. *Health Services Research*, 34(5, Pt. 2) 1189.
- Portin, B. S. (2009). Assessing the effectiveness of school leaders: New directions and new processes. Washington, DC: The Wallace.
- Price, T.A. (2016). Teachers teaching under the lens: curriculum surveillance using the edTPA. *Curriculum Studies in Higher Education*. Proceeding book, 9.
- Rose, M. (1999). "Our hands will know": the development of tactile diagnostic skill—teaching, learning, and situated cognition in a physical therapy program. *Anthropology & education quarterly*, 30(2), 133-160.
- Sato, M. (2014). What is the underlying conception of teaching of the edTPA?. *Journal of Teacher Education*, 0022487114542518.
- Schön, D. A. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions*. Jossey-Bass.
- Wendel, F.C. & Sybouts, W. (1988). *Assessment center methods in educational administration: Past, present, and future*. Tempe, AZ: University Council for Education Administration.
- Wendel, F. C., & Uerling, D. F. (1989). Assessment centers — contributing to preparation programs for principals. *NASSP Bulletin*, 73(515), 74-79.
- Yin, R. (2018). *Case study and research and applications: Design and methods*. (6<sup>th</sup> ed.). Thousand Oaks, CA: Sage Publications.