

# 1           **Connecting Program Assessment of Teaching Talent** 2           **Training in Higher Education to the Reality of the** 3           **Schools: A Case Study of a Teacher Preparation Program**

4  
5           *Amid the broader accountability movement in education, the evaluation of teacher*  
6           *preparation programs (TPPs) must be systematically aligned with the performance*  
7           *assessments of their graduates by school leaders in authentic educational contexts.*  
8           *Existing research indicates that linking TPP evaluations to graduate effectiveness*  
9           *ratings represents a promising mechanism for enhancing program accountability*  
10           *and continuous improvement. Concurrently, scholarship underscores the necessity*  
11           *of embedding equity as an explicit objective within TPP evaluation frameworks.*  
12           *This study investigates how a TPP at a U.S. university conceptualized and*  
13           *operationalized a program-level assessment system that integrates employer*  
14           *evaluations of graduates within culturally and linguistically diverse school*  
15           *environments. Data sources include internally developed assessments and*  
16           *comprehensive annual reports from the state's Department of Education.*  
17           *Comparative analyses were conducted to identify programmatic strengths and*  
18           *growth areas in preparing teachers for culturally and linguistically diverse*  
19           *classrooms. Findings demonstrate the feasibility of connecting TPP evaluation*  
20           *systems to employer feedback and highlight the imperative of incorporating equity*  
21           *considerations into assessment design, implementation, and interpretation to*  
22           *inform programmatic refinements. Although situated in the U.S. context, the study*  
23           *addresses a global challenge: the increasing diversity in primary and secondary*  
24           *classrooms. Consequently, the implications extend beyond national boundaries,*  
25           *offering insights for international teacher education reform.*

26  
27           **Keywords:** *Employment Context, Higher Education, Program Assessment,*  
28           *Teacher Education, Teaching Talent Training*

## 29 30 31           **Introduction**

32  
33           Existing research studies regarding the evaluation of teacher preparation  
34           programs (TPPs), sometimes more broadly referred to as educator preparation  
35           programs (EPPs), have given attention to “identifying the strengths and weaknesses  
36           of evaluation metrics based on rigorous criteria for accuracy and utility,” thus alerting  
37           researchers to argue for equity to be established as an explicit goal and a desired  
38           outcome of teacher preparation evaluation (Cochran-Smith & Reagan, 2021, p. 2). In  
39           the United States, standards-based policies have become firmly embedded across all  
40           aspects of the “Learning to Teach Professional Sequence” over the past two decades,  
41           with significant implications for how teachers are prepared and how TPP graduates  
42           are evaluated (Bales, 2015). Concurrently, the broader accountability movement in  
43           education emphasizes holding teachers, schools, and districts responsible for  
44           performance outcomes (Hood et al., 2022). As a result, meaningful evaluation of TPPs  
45           requires explicit connections between program assessments and how school leaders  
46           evaluate graduates once they enter the workforce. For program improvement to be  
47           relevant and actionable, TPP assessments must be aligned with candidates’

1 demonstrated performance in school settings. As the cultural and linguistic diversity  
2 of students from kindergarten through to twelfth grade (K-12) learners increases,  
3 equity is a crucial evaluation factor for future teachers in a TPP and for the impact  
4 they have on their K-12 learners.

5 The call to explicitly center equity in TPP evaluation has intensified in recent  
6 years. The National Academy of Education’s 2024–2025 consensus report (National  
7 Academy of Education, 2024) on evaluating and improving TPPs emphasizes equity-  
8 centered program improvement, careful selection of measures connected to teaching  
9 quality, and the importance of contextualizing outcomes, including the demographics  
10 of students served by graduates. This report synthesizes evidence and offers logic  
11 models to guide evaluation that balances accountability with improvement, and it  
12 reinforces a central theme of this study: the need to consider the school employment  
13 context and the culturally and linguistically diverse (CLD) learners in classrooms  
14 when interpreting outcomes for graduates.

15 Existing research emphasizes the need for TPP evaluation systems that foster  
16 candidate growth, employ balanced evidence portfolios, interpret data contextually to  
17 promote equity, link TPP assessments to employer evaluations of graduates, and  
18 prepare candidates for implementing English learner (EL) policies. However,  
19 empirical studies examining how TPPs achieve these goals remain limited. This study  
20 explores how a teacher preparation program at a U.S. public university conceptualized  
21 and operationalized a program-level assessment system that integrates employer  
22 evaluations of graduates working in CLD school environments. Specific questions  
23 that guided the study are:

- 24
- 25 1. How does a teacher preparation program make connections between its  
26 internal assessment for teacher candidates’ growth and employer evaluations  
27 of its graduates in the school context?
- 28 2. What areas of alignment and tension emerge between teacher candidates’  
29 performance on program assessments and school leaders’ evaluations of  
30 program graduates, particularly with respect to equity and preparation for  
31 culturally and linguistically diverse learners?
- 32

33 This article is organized into six sections. The Introduction situates the study  
34 within current debates on equity, accountability, and teacher preparation program  
35 evaluation. The Literature Review synthesizes research on performance-based  
36 evidence, equity-centered evaluation, and continuous improvement in teacher  
37 preparation, identifying a critical gap in the existing literature. The Methodology  
38 outlines the program context, data sources, and analytic approach of the study. The  
39 Findings present results from program-level assessments and state employer data. The  
40 Discussion interprets these findings through an equity lens and considers implications  
41 for program improvement and evaluation practice. The Conclusion summarizes key  
42 contributions, notes limitations, and highlights directions for future research.

43  
44  
45

## 1 Literature Review

2  
3 Although EPP is a broader umbrella term under which TPPs typically fall, the  
4 terms are often used interchangeably in the literature; therefore, this paper uses EPP  
5 and TPP interchangeably to refer to teacher preparation programs. Within this body  
6 of literature, recent scholarship converges on the importance of equity-centered  
7 program improvement, culturally responsive and sustaining pedagogy, and balanced  
8 assessment systems for TPPs/EPPs (Bardelli et al., 2023; National Academy of  
9 Education, 2024, 2025; Olson & Toch, 2024; REL Pacific/IES, 2025). Building on  
10 these converging perspectives, this synthesis situates the present case study within  
11 evolving national and international discourse by integrating research on performance-  
12 based evidence and its limitations, equity and standardized measures, continuous  
13 improvement frameworks, and evaluation approaches tailored to CLD learners.  
14 Together, these strands inform both program-level redesign and policy-aligned  
15 accountability practices, with implications for EPPs serving K–12 multilingual  
16 learners.

### 17 *Performance Evidence and Its Limits*

18  
19  
20 Goldhaber et al. (2013) conducted one of the earliest statewide efforts to connect  
21 teacher training institutions with student achievement. Through an analysis of  
22 multiple teacher training programs, they found that “the best assessments of teachers  
23 are those based on actual classroom performance rather than pre- or in-service  
24 credentials” (Goldhaber et al., 2013, p. 42). This finding catalyzed a decade of  
25 attention to performance assessments for TPPs, with studies exploring the validity,  
26 reliability, and practical utility of such measures for program evaluation and  
27 accountability (Barron, 2015; Bastian et al., 2016, 2018, 2023; Knight et al., 2014).  
28 Nevertheless, for performance evidence to inform program improvement, the  
29 evidence must be systematically connected to candidate assessments within TPPs,  
30 creating coherence between what programs teach, how candidates demonstrate  
31 proficiency, and how graduates are evaluated in schools.

32 Although U.S. researchers began examining relationships between TPPs and  
33 graduate outcomes in the 1980s as policy interest grew in program quality, notable  
34 research gaps exist. For instance, Bastian et al. (2018) note that “little is known about  
35 the relationships between TPPs and the evaluation ratings of program graduates” (p.  
36 429) by school leaders. This gap underscores a persistent need to link program features,  
37 such as clinical placements, supervision quality, feedback systems, and coursework  
38 integration, to how graduates perform and grow in real classroom settings.

39 Recent literature introduces two critical nuances. First, rather than focusing only  
40 on initial effectiveness, analyses of graduates’ growth trajectories reveal meaningful  
41 differences across TPPs (Bardelli et al., 2023; National Academy of Education, 2024).  
42 Features that predict starting levels of effectiveness are not identical to those  
43 predicting early-career growth (e.g., clinical placement characteristics versus ongoing  
44 support and mentoring). This distinction is especially salient for equity because  
45 contexts serving CLD learners often require graduates to demonstrate more advanced  
46 competencies in culturally responsive practice, linguistically informed instruction,

1 and literacy assessment for multilingual learners. However, the field experience  
2 contexts offered to teacher candidates may not consistently demand or cultivate the  
3 same level of CLD instructional competencies. Second, there is national consensus  
4 that emphasizes building evaluation systems that support continuous improvement  
5 while avoiding punitive interpretations that may inadvertently disincentivize  
6 placements in high-need schools serving multilingual learners (National Academy of  
7 Education, 2024, 2025).

### 8 9 *Equity and Standardized Measures*

10  
11 A complementary line of inquiry examines the role of standardized measures,  
12 including value-added models for teacher evaluation, in program accountability and  
13 improvement. Studies suggest that linking TPP evaluations to graduates' evaluation  
14 ratings can be promising, provided systems address measurement validity and  
15 contextual factors (Bastian et al., 2018; Ronfeldt & Campbell, 2016). Bastian et al.  
16 caution that if TPP evaluation systems do not incorporate the school context,  
17 programs whose graduates disproportionately work in high-need schools may be rated  
18 lower, unless the systems explicitly account for the employment context and student  
19 population needs.

20 The broader debate on standardized testing and value-added models remains  
21 active. A systematic review by Amrein-Beardsley et al. (2023) urges caution in the  
22 use of value-added models and advocates for multi-measure interpretations due to  
23 validity concerns, emphasizing the risk of interpretive pitfalls when measures of  
24 student learning are tied closely to standardized tests, particularly for teachers of  
25 multilingual learners. At the same time, Koedel (2025) argues that value-added  
26 models can be useful when applied carefully and combined with observational and  
27 other evidence, situating value-added models as one piece of a balanced evidence  
28 portfolio rather than a singular metric. Reflecting this balance, Olson and Toch (2024)  
29 recommend decoupling high-stakes accountability from instructional assessment to  
30 better serve equity goals. Collectively, this body of research supports the use of  
31 multimodal, performance-based evidence, with employer metrics interpreted  
32 contextually rather than in isolation. The policy takeaway is not to discard growth  
33 measures, but to embed them within a balanced, culturally responsive assessment  
34 system and adjust interpretations for school context, the stance adopted in this study.

### 35 36 *Evaluation of EPPs for Continuous Improvement*

37  
38 The National Academy of Education's consensus report (2024) and subsequent  
39 recommendations (2025) provide actionable frameworks that balance accountability  
40 with program improvement. They call for evidence coherence, meaning alignment  
41 across candidate assessments, graduate performance data, and program goals, coupled  
42 with equity considerations and contextual interpretation of outcomes. Central to these  
43 frameworks is the use of data systems to support continuous improvement cycles,  
44 emphasizing teacher learning over time rather than focusing exclusively on initial  
45 performance snapshots.

1 Similarly, research on teacher effectiveness growth reinforces the importance of  
2 program features that scaffold ongoing development. Bardelli et al. (2023) document  
3 significant early-career growth differences across TPPs and show that program  
4 features correlate differently with initial effectiveness versus growth trajectories. For  
5 programs aiming to strengthen CLD competencies, their findings imply prioritizing  
6 clinical features that support practice growth (e.g., placement quality, structured  
7 feedback systems, mentor capacity, and coherent coursework-practicum integration).

#### 8 9 *EPP Evaluation in CLD Contexts*

10  
11 Levy-Feldman (2025) argued that redesigning assessment practices to better  
12 support learners from diverse sociocultural backgrounds can strengthen teaching and  
13 learning while advancing equity. Within teacher preparation, where culturally  
14 responsive and linguistically sensitive practices (Berg & Huang, 2015) have gained  
15 prominence, research indicates that teacher candidates' learning experiences  
16 frequently emphasize shifts in attitudes, beliefs, and professional dispositions rather  
17 than technical competence alone (Miller et al., 2023). This shift reflects broader  
18 values-based reconceptualization of teacher quality. As Curtis (2015) contends,  
19 competence encompasses not only instructional knowledge and behavior but also  
20 character, ethical commitments, and sensitivity to cultural, social, and political  
21 contexts (p. 314).

22 Policy briefs and professional guidance further reinforce culturally responsive  
23 and sustaining pedagogy as a foundational orientation for evaluation and  
24 improvement in culturally and linguistically diverse (CLD) contexts. State and  
25 national resources (California Department of Education, 2025; Keith et al., 2023; REL  
26 Pacific/IES, 2025) offer actionable strategies to honor and sustain students' languages  
27 and cultural practices, operationalizing frameworks advanced by Ladson-Billings  
28 (2014) and Paris and Alim (2014) in today's classrooms. These resources foreground  
29 evaluation practices that enable students to demonstrate learning through culturally  
30 meaningful and contextually valid forms. Such orientations are consistent with  
31 research linking teacher ethics and multicultural competence to definitions of quality  
32 teaching, suggesting that effectiveness in multi-ethnic and multiracial contexts should  
33 be central to how teacher competence is evaluated rather than treated as a peripheral  
34 skill (Boon & Lewthwaite, 2016).

35 Illustrative programmatic cases further demonstrate how these principles can be  
36 operationalized within teacher preparation. Project ELEECT (Donley et al., 2023), for  
37 example, positions teacher research as a lever for candidate and novice teacher  
38 learning about culturally and linguistically responsive practice. Related assignments,  
39 such as Impact on Learning Studies, provide structured opportunities for candidates  
40 to analyze instructional impact with multilingual learners by examining feedback  
41 cycles, formative assessment practices, and differentiated scaffolds. Such designs are  
42 particularly instructive for evaluation systems, as they generate situated evidence of  
43 candidate learning in CLD contexts where feedback and formative assessment have  
44 historically posed challenges.

45 Taken together, the evolving literature on TPP/EPP evaluation and improvement  
46 increasingly converges on a balanced, equity-centered approach that values

1 performance evidence, growth over time, and contextual interpretation. For programs  
2 serving CLD learners, evaluation systems must honor culturally responsive pedagogy  
3 while leveraging multimodal evidence to inform continuous improvement. By  
4 integrating clinical quality, feedback-rich learning opportunities, policy-aligned  
5 assessment literacy, and context-aware analytics, EPPs can advance equitable  
6 outcomes for teacher candidates and the multilingual learners they serve (Bardelli et  
7 al., 2023; National Academy of Education, 2024, 2025; Olson & Toch, 2024; REL  
8 Pacific/IES, 2025). Despite robust conceptual and policy guidance, there remains a  
9 critical need for empirical research documenting how educator preparation programs  
10 operationalize equity-centered evaluation principles in assessment practice,  
11 particularly in CLD contexts.

12  
13

## 14 **Methodology**

15

### 16 *Context of the Study*

17

18 The current study examines how equity can and must be incorporated into a  
19 standards-driven and data-informed TPP assessment system, particularly within  
20 culturally and linguistically diverse school settings. This focus emerged from recent  
21 efforts by the largest TPP in a western U.S. state to engage in data informed program  
22 improvement, which revealed several equity related concerns for its evaluation  
23 system. To distinguish this university teacher preparation program from TPPs more  
24 broadly, it is referred to here as UTPP.

25 The university that houses the UTPP has the longest history of teacher preparation  
26 in the state, originating as the state's normal school more than 130 years ago. Like all  
27 TPPs in the U.S., however it faces increasing scrutiny from the public and state  
28 regulators. State policy requires all licensure areas to meet the Teacher Quality  
29 Standards (TQS) and the English Language Learner Educator Preparation Standards,  
30 in addition to discipline-specific standards. The state's K–12 system has also adopted  
31 a teacher evaluation rubric aligned with the TQS. As a result, policymakers, the  
32 public, and TPPs alike closely monitor how well programs such as the UTPP prepare  
33 future teachers to meet these requirements and to be viewed as competent by K–12  
34 schools.

35 Building on these accountability expectations, recent updates to the state's  
36 Educator Preparation Program (EPP) Report dashboards have expanded the annual  
37 report to include interactive tools linking new teachers' employment, performance,  
38 and retention to their preparation programs. These enhancements shape how the  
39 UTPP and other programs are reviewed and signal the state's growing emphasis on  
40 connecting teacher preparation to workforce readiness outcomes.

41 In addition to these state EPP dashboard updates, several state policy changes  
42 have further-clarified proficiency thresholds and English Language Development  
43 requirements. These policies reinforce the need for TPPs to prepare candidates for  
44 English Learner policy implementation. They also influence candidate assessment  
45 design, particularly in aligning dispositions and practices with the English Language  
46 Learner Educator Preparation Standards and affect how graduates' measure of student

1 learning data are interpreted in multilingual, high-need districts. Collectively, these  
2 shifts provide broader expectations for programs, promote more consistent  
3 benchmarking, and offer an opportunity to interpret performance evidence within the  
4 context of candidates' future teaching environments.

5 To promote transparency in program improvement efforts and meet the state's  
6 teacher preparation program reauthorization requirements, the UTPP engaged in  
7 program assessment revision that explicitly tied assessment practices to the following:  
8 (1) the state teacher licensure standards including English Language Learner Educator  
9 Preparation Standards, (2) how school leaders evaluate first-year graduates from the  
10 UTPP and (3) value-added assessment implementation to demonstrate teacher  
11 candidates' growth during the program. The annual data provided by the state about  
12 graduates' employment context and their teaching performance rated by school  
13 leaders provided an opportunity for an examination of equity issues in the evaluation  
14 of the UTPP.

15 The UTPP examined in this study has multiple licensure areas, all of which rely  
16 on similar service courses and clinical experiences to address the TQS and the English  
17 Language Learner Educator Preparation Standards. Consistency and coherence in  
18 program assessment was challenging for years, and the tools used in the past were not  
19 aligned with the state's licensure standards. In academic year 2022–23, led by a  
20 faculty team of ten members representing 12 different licensure areas, five existing  
21 assessment tools were revised or redesigned to explicitly align with the state  
22 standards: professionalism disposition assessment, Lesson Planning, Lesson  
23 Observation, Teacher Work Sample, and Program Completer Survey. These  
24 performance-based assessments are designed to be implemented with teacher  
25 candidates in the field during clinical experiences. The Lesson Plan and Lesson  
26 Observation rubrics were structured according to the rubric recommended by the state  
27 to be used by K-12 school leaders evaluating their teachers. All newly revised or  
28 developed assessment tools were implemented across all 12 licensure areas in  
29 academic year (AY) 2023–24. Beginning fall 2023, the two-year effort provided  
30 extensive data leading to an opportunity to examine equity issues in the assessment  
31 system used to evaluate candidates in the UTPP and in the context of how the UTPP  
32 graduates are evaluated by the K-12 school leaders.

### 33 34 *Data Sources and Scope*

35  
36 Data collected and analyzed for this case study refer to multiple aspects of the  
37 UTPP assessment and evaluation and include a focus on aspects related to equity  
38 issues. Specifically, data will be examined in the context of candidates' ability to work  
39 with CLD learners.

40 This study draws on two data sources. One source comes from the assessments  
41 developed and implemented by the UTPP. Data collection and aggregation was  
42 gathered via Student Learning and Licensure (SLL), an electronic data management  
43 system specifically designed for TPPs with the capability to tie specific performance  
44 indicators to required standards when aggregating data. During AY 2023–24 (i.e., fall  
45 and spring semesters) completed assessments totaled: Program Completer Survey (n  
46 = 612), Teacher Work Sample (n = 143), Lesson Observation (n = 1460), and Lesson

1 Plan (n = 436). SLL generated ratings of teacher candidates' performance based on  
2 performance indicators as well as required TQS and the English Language Learner  
3 Educator Preparation Standards, grouping the results according to assessors for  
4 comparison purposes for each semester. The Program Completer Survey is completed  
5 once by teacher candidates, field mentors and university supervisors. The Teacher  
6 Work Sample is evaluated once by university supervisors. The Lesson Observation  
7 and Lesson Plan rubrics are used multiple times in the program by mentor teachers  
8 and university supervisors. Using the same assessment tools at different stages enables  
9 viewing the impact of the program on teacher candidates' growth.

10 A second source of data comes from the State EPP Report Dashboard generated  
11 by the state's Department of Education, which provides comprehensive annual EPP  
12 data regarding all EPPs in the state. The State EPP Report includes multiple aspects  
13 of each EPP, such as enrollment and completion rate, licensure test passing rate, new  
14 teacher employment rate, employment context, new teacher performance and  
15 retention, etc. The report also includes the employment context of graduates and the  
16 evaluation rating of new teachers by the school leaders linking new teacher  
17 performance to the specific EPP from which they graduated. Though a two-year lag  
18 exists in the publishing of the EPP report, the report provides insight into graduates'  
19 teaching performance as viewed by employers. Because employers follow the same  
20 state licensure standards, an opportunity exists to compare the UTPP assessment of  
21 its candidates and the K-12 schools' assessment of the UTPP graduates. The state EPP  
22 data is tied to the Teacher Quality Standards (TQS) but not to the English Language  
23 Learner Educator Preparation Standards. Importantly, due to a 2023 cybersecurity  
24 incident, more recent traditional program data are delayed, so statewide totals for  
25 2021–22 and 2022–23 reflect alternative EPP data only; highlights from 2020–21  
26 remain the most recent available statewide summaries. This context is considered  
27 when comparing PCS and EPP outcomes.

### 28 29 *Analytic Approach*

30  
31 While data were collected and aggregated, the faculty team simultaneously  
32 engaged in data analysis, which provided an opportunity for action research to seek  
33 understanding of how programs prepared teacher candidates to address the state  
34 standards and how the assessment system aligned with school leaders' evaluation of  
35 program graduates. Required UTPP assessments were completed and gathered via  
36 SLL on a semester basis. SLL organized data via assessment tools, assessors, and  
37 programs. Assessments were aggregated according to the performance indicators as  
38 established in the assessment tools and the required standards already matched to the  
39 performance indicators in SLL. Then, SLL was used to generate reports by  
40 performance indicators or required standards to show the percentage of the teacher  
41 candidates meeting or exceeding expectations at different stages of the program.  
42 Program growth areas were identified based on the performance indicators or required  
43 standards that consistently showed lower percentage of teacher candidates meeting or  
44 exceeding expectations across semesters or assessment tools.

45 To interpret findings through an equity lens, we triangulated three data sources.  
46 First, we examined program-level performance by Teacher Quality Standards (TQS)

1 and the English Language Learner Educator Preparation Standard elements across the  
 2 Lesson Plan, Lesson Observation, Teacher Work Sample, and Program Completer  
 3 Survey. These data were disaggregated by assessor (i.e., teacher candidate, mentor  
 4 teacher, and university supervisor) and by program stage (i.e., practicum and student  
 5 teaching). Second, we analyzed employer evaluation ratings by TQS and Measures  
 6 of Student Learning from the State EPP Report. Third, we considered employment  
 7 context indicators, including multilingual learner densities. Analytic memos were  
 8 then drafted to identify areas of alignment or tension between candidate performance  
 9 growth and employer ratings.

## 12 Findings

14 This section presents a summary of the data analysis focusing on the strengths  
 15 and growth opportunities evidenced in the following data: Lesson Plan, Lesson  
 16 Observation, Program Completer Survey as well as a comparison with data from the  
 17 State EPP Report, and a consideration of the contexts of teaching.

### 19 *Lesson Planning Results*

21 The Lesson Plan rubric is designed to have multiple elements the candidates  
 22 need to include in their lesson plans to address relevant Teacher Quality Standards  
 23 (TQS). At times, multiple elements address the same TQS, such as TQS 5.10(1) and  
 24 5.01(2). For example, elements related to 5.10(1) include E1: alignment to academic  
 25 standards, E2: inclusion of lesson objectives, E3: inclusion of formative and  
 26 summative assessments, and E4: alignment with district plans. In this study, the  
 27 Lesson Plan analysis compared the two stages of the program (i.e., practicum and  
 28 student teaching) as shown in Table I.

30 **Table 1.** *Lesson Planning Data Aggregation: Percentage of Candidates Meeting TQS*  
 31 *Expectations*

TQS & Elements	Practicum (%)		Student Teaching (%)	
	F23	S24	F23	S24
5.01(1) E1	100	97	100	100
5.01(1) E2	99	95	100	100
5.01(1) E3	100	95	100	99
5.01(2) E4	100	98	100	100
5.01(2) E5	96	90	<u>90</u>	<u>94</u>
5.01(2) E6	97	94	100	98
5.01(2) E7	96	<u>88</u>	97	95
5.01(3) E8	<u>90</u>	<u>94</u>	100	100
5.02(1) E9	<u>94</u>	<u>89</u>	97	96
5.02(2) E10	<u>91</u>	<u>93</u>	96	97
5.02(3) E11	<u>90</u>	<u>88</u>	100	95
5.03(1) E12	96	96	97	98
5.03(2) E13	96	<u>93</u>	97	97
5.03(3) E14	97	99	100	100

<u>5.03(4) E15</u>	<u>91</u>	<u>92</u>	<u>93</u>	98
5.03(5) E16	96	<u>94</u>	97	100
5.03(6) E17	99	96	100	100
mean	96	<u>94</u>	98	98

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

In Table 1, elements associated with each TQS for which fewer than 95% of candidates met expectations are italicized and underlined to identify growth areas. Overall findings show a clear progression in ratings from the practicum to student teaching, the final stage of the program. During the practicum stage, fewer than 95% of teacher candidates met expectations across several TQS. By the student teaching stage, however, more than 95% of teacher candidates consistently met expectations in nearly all areas, with two notable exceptions.

One area of continued concern is TQS 5.01(2), which focuses on developing and implementing lessons that connect to multiple content areas or disciplines while emphasizing literacy and mathematics. Specifically, teacher candidates demonstrate needed growth in TQS 5.01(2), as indicated by Element 5 (i.e., lesson plan connect to a variety of content areas/disciplines and emphasize mathematics and literacy), which requires making content-specific academic language accessible to students with varying levels of English proficiency. The second area in which candidates continue to struggle at the student teaching stage is TQS 5.03(4), Element 15, which is to establish and communicate high expectations and use processes to develop critical thinking and problem-solving skills. Both areas are identified as growth areas and are aligned with multiple English Language Learner Educator Preparation Standards, including Standard 5.10 (language acquisition), Standard 5.11 (literacy development), and Standard 5.12 (instructional strategies for culturally and linguistically diverse learners). Additional details on the TQS and the English Language Learner Educator Preparation Standards are provided in the Appendix.

### *Lesson Observation Results*

The Lesson Observation rubric also includes multiple elements that candidates must demonstrate to address relevant Teacher Quality Standards (TQS). Similar to the Lesson Plan analysis, candidates' performance on the Lesson Observation was compared across two program stages, practicum and student teaching, as shown in Table 2.

1 **Table 2.** *Lesson Observation Data Aggregation: Percentage of Candidates Meeting*  
 2 *TQS Expectations*

TQS	Practicum				Student Teaching			
	Mentor Teacher (%)		University Supervisor (%)		Mentor Teacher (%)		University Supervisor (%)	
	F23	S24	F23	S24	F23	S24	F23	S24
5.01(1)	100	96	95	95	98	97	99	97
<u>5.01(2)</u>	<u>92</u>	<u>87</u>	<u>92</u>	<u>89</u>	95	<u>93</u>	97	<u>93</u>
5.01(3)	96	<u>88</u>	95	<u>94</u>	95	<u>94</u>	98	96
<u>5.02(1)</u>	<u>91</u>	<u>82</u>	<u>91</u>	<u>87</u>	<u>93</u>	<u>92</u>	95	<u>94</u>
<u>5.02(2)</u>	96	96	<u>93</u>	<u>93</u>	98	<u>93</u>	98	<u>94</u>
5.02(3)	<u>93</u>	<u>93</u>	96	<u>93</u>	97	<u>88</u>	96	96
5.03(1)	<u>91</u>	<u>89</u>	97	<u>89</u>	97	99	98	95
<u>5.03(2)</u>	<u>91</u>	<u>83</u>	<u>87</u>	<u>84</u>	95	<u>92</u>	96	<u>93</u>
5.03(3)	95	<u>92</u>	99	98	99	98	99	99
<u>5.03(4)</u>	<u>92</u>	<u>82</u>	<u>86</u>	<u>87</u>	<u>92</u>	<u>87</u>	97	<u>94</u>
<u>5.03(5)</u>	<u>93</u>	<u>83</u>	<u>85</u>	<u>90</u>	95	<u>87</u>	95	<u>90</u>
5.03(6)	<u>93</u>	<u>89</u>	<u>93</u>	<u>93</u>	95	<u>93</u>	98	95
5.04(1)	99	96	99	99	99	98	99	100
mean	<u>94</u>	<u>89</u>	<u>93</u>	<u>92</u>	96	<u>93</u>	97	95

3  
 4 For the Lesson Observation assessment, TQS for which fewer than 95% of  
 5 candidates met expectations are italicized and underlined to identify growth areas. As  
 6 shown in Table 2, at the practicum stage, a greater number of TQS falls below the  
 7 95% benchmark, whereas at the student teaching stage, most TQS are met by more  
 8 than 95% of teacher candidates. Based on candidates' demonstrated growth from  
 9 practicum to student teaching, six areas that continue to show two or more ratings  
 10 below 95% during student teaching are identified as growth areas. All six areas are  
 11 related to working with culturally and linguistically diverse learners: content-area  
 12 literacy (TQS 5.01(2)); fostering a predictable learning environment (TQS 5.02(1));  
 13 addressing multiple aspects of diversity (TQS 5.02(2)); using formal and informal  
 14 methods to assess student learning (TQS 5.03(2)); developing critical-thinking and  
 15 problem-solving skills (TQS 5.03(4)); and providing opportunities for students to  
 16 work in teams (TQS 5.03(5)). These growth areas are aligned with the English  
 17 Language Learner Educator Preparation Standards related to second language  
 18 acquisition, literacy for content learning, and research-based instructional and  
 19 assessment strategies. Increased and more intentional use of CLD strategies may  
 20 further strengthen candidates' performance in lesson delivery.

21  
 22 *Program Completer Survey vs. State EPP Report*

23  
 24 To continue the data-driven reflection and inform strategic program revision  
 25 aimed at better preparing teacher candidates to meet school needs, we used the State  
 26 EPP Report to compare the evaluation of teacher candidates at the end of their

1 programs to the school evaluation of the UTPP graduates. Specifically, the UTPP’s  
 2 Program Completer Survey data related to the TQS, including assessments completed  
 3 by mentor teachers, university supervisors, and teacher candidates, were compared to  
 4 the available EPP Report. Table 3 shows the comparison based on the four TQS. Note  
 5 that the State EPP Report only has the data from 2019 to 2021.

6  
 7 **Table 3.** *Comparison between State EPP Report Data and UTPP Program*  
 8 *Completer Survey Data by TQS*

TQS	State TPP (%) 2019-2021	UTPP					
		Mentor Teacher (%)		University Supervisor (%)		Teacher Candidates (%)	
		S24	F23	S24	F23	S24	F23
5.01	93.5	96.8	94.9	99.9	98.9	99.5	98.6
5.02	92.5	97.2	96.3	99.9	98.5	98.4	97.7
5.03	93.3	97.0	92.4	99.5	97.5	98.9	98.4
5.04	92.7	97.4	94.4	99.2	98.4	99.5	97.9
mean	93.0	97.1	94.5	99.6	98.3	99.1	98.2

9  
 10 The comparison identified that all evaluators, meaning school leaders evaluating  
 11 UTPP graduates as first-year teachers and the UTPP stakeholders evaluating the  
 12 UTPP student teachers, agree that more than 92% of the UTPP teacher candidates or  
 13 graduates meet or exceeded TQS expectations. Generally, the evaluation rating by  
 14 university supervisors and teacher candidates was slightly higher than the evaluation  
 15 by school leaders and mentor teachers. University supervisors and teacher candidates  
 16 may have only teacher candidates as reference points, while school leaders and mentor  
 17 teachers have other veteran teachers as reference points. Additionally, because mentor  
 18 teachers see teacher candidates more often than university supervisors, mentor  
 19 teachers’ evaluation may be the result of combining formal and informal observations  
 20 while university supervisors base their evaluation mainly on formal observations  
 21 when the teacher candidates perform better. While teacher candidates and university  
 22 supervisors’ assessments of teacher candidates remain similar from fall 2023 to spring  
 23 2024, mentor teachers’ responses in spring 2024 indicate a greater percentage of  
 24 teacher candidates meeting or exceeding expectations in spring 2024 compared to fall  
 25 2023.

### 26 27 *Employer Ratings and Measure of Student Learning*

28  
 29 One aspect of the State EPP Report was puzzling. The evaluation rating of new  
 30 teachers by the employers is based on the four TQS and a fifth category, Measure of  
 31 Student Learning. Table 4 summarizes the State EPP Report data for the evaluation  
 32 of graduates from the EPP in discussion.  
 33

1 **Table 4. Employer Evaluation of Graduates by the State EPP Annual Data:**  
 2 **Percentage of Teachers Meeting the TQS**

Cohort Year	TQS 5.01 (%)		TQS 5.02 (%)		TQS 5.03 (%)		TQS 5.04 (%)		Measure of Student Learning	
	XX	UTPP	XX	UTPP	XX	UTPP	XX	UTPP	XX	UTPP
20-21	91.3	92.1	91.9	88.9	92.4	90.5	92.7	91.6	90.0	88.8
19-20	94.5	94.8	95.9	96.0	95.2	96.0	93.5	93.7	90.3	89.6
18-19	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
17-18	90	81.3	90.8	84.3	90.0	82.9	90.5	83.4	90.8	87.1
16-17	89.9	85.7	83.3	74	86.2	79.6	79.4	74.5	89.2	82.3

3 Data percentages are rounded to one decimal place. No data were collected for the 2018-2019 cohort  
 4 due to the Covid-19 pandemic. XX=Employer evaluation for all XX graduates, UTPP=Employer  
 5 evaluation for UTPP graduates.

6  
 7 Data analysis revealed several findings. First, ratings by school employers show  
 8 steady and continuous improvement of the UTPP graduates' effectiveness over the  
 9 years regarding all four of the TQS (except for TQS 1 where there was a decrease  
 10 from 2016-17 to 2017-18), thus indicating increased satisfaction by the school leaders  
 11 until 2020-21 when the Covid-19 Pandemic hit. Second, the graduates' effectiveness  
 12 improved from being below the state's average regarding all TQS to being slightly  
 13 above the state average for all TQS until 2020-21. At the same time, in 2019-2020,  
 14 while the graduates were rated at a level above the state average for all TQS, the  
 15 overall rating regarding Measures of Student Learning for the graduates was below  
 16 the state average. Finally, while progress was shown for improvement in Measures of  
 17 Student Learning, the graduates' performance continued to be below the state average.

18 To understand why the graduates rated below the state average for Measures of  
 19 Student Learning, even when they exceeded the state average for all TQS, we  
 20 examined the State EPP Report data from an equity angle. We first reviewed the  
 21 state's definition of Measures of Student Learning as presented in the EPP Report  
 22 dashboard (Colorado Department of Education, 2023–2024). Notably, this definition  
 23 has been revised in recent updates to the dashboard, and the version cited below is no  
 24 longer publicly available:

25  
 26 The change in student achievement against [State] Academic Standards for an individual  
 27 student between two or more points in time, which shall be determined using multiple  
 28 measures, **one of which shall be the results of statewide summative assessments**, and  
 29 which may include other standards-based measures that are rigorous and comparable  
 30 across classrooms of similar content areas and levels. (Emphasis added) (State EPP  
 31 Report Dashboard)

32  
 33 Statewide summative assessments, meaning standardized tests, are mandated as  
 34 a measure of student learning and play a central role in teacher evaluations. Analysis  
 35 of employment context data from the State EPP Report helps explain why UTPP  
 36 graduates met the TQS but scored poorly on the Measure of Student Learning. The  
 37 State EPP Report provides school ratings of new teachers based on TQS but does not  
 38 include information about meeting the English Language Learner Educator

1 Preparation Standards. It also shows that UTPP graduates work in contexts similar to  
2 the statewide average, though they are more likely to be employed in districts with  
3 higher numbers of CLD learners. However, the report does not link employment  
4 context to evaluation results, which risks misinterpretation of teaching effectiveness.  
5 Research consistently shows that standardized tests disadvantage students from racial  
6 minority, language minority, and low-income backgrounds (Bach 2020). CLD K–12  
7 students face challenges on these tests due to factors beyond a teacher’s control  
8 (Amrein & Berliner, 2002; Au, 2016; Bach, 2020). As a result, teachers working with  
9 more CLD students may be just as competent as their peers yet appear less effective  
10 when standardized tests dominate evaluation metrics.

## 11 12 13 **Discussion**

14  
15 Standardized tests have long posed challenges for students from racial minority,  
16 language minority, and low-income backgrounds. Consequently, when statewide  
17 assessments are embedded as required components of Measures of Student Learning,  
18 teachers working in CLD classrooms may appear less effective than they are in  
19 practice. This concern extends beyond UTPP and has prompted recent analyses and  
20 policy reports (Colorado Department of Education, 2024–2025; Olson & Toch, 2024)  
21 to call for statewide systems that better support equitable instruction and to reconsider  
22 the high-stakes role of summative assessments.

23 Within this broader context, UTPP’s assessment data reveal specific areas where  
24 teacher candidates need stronger preparation to serve CLD learners. The patterns  
25 observed parallel national trends documented in recent reviews (Donley et al., 2023;  
26 Sandoval et al, 2023) and existing research in the field of teacher education (Boon &  
27 Lewthwaite, 2016). Across assessment tools, growth areas consistently relate to  
28 candidates’ readiness for linguistically responsive teaching. Lesson Planning results  
29 indicate ongoing difficulty in integrating literacy and mathematics across content  
30 areas, an especially complex task for novice teachers working in linguistically diverse  
31 classrooms. These findings align with research emphasizing the need for structured  
32 support that helps candidates anticipate academic language demands and incorporate  
33 students’ cultural and linguistic assets into disciplinary learning (National Academy  
34 of Education, 2025; REL Pacific/IES, 2025).

35 Lesson Observation data further highlight the need to strengthen candidates’  
36 formative assessment practices and their ability to facilitate collaborative, cognitively  
37 rigorous learning. These competencies are central to culturally sustaining pedagogy,  
38 which affirms students’ linguistic and cultural identities while promoting deep  
39 engagement. Recent state and national briefs recommend strategies such as balanced  
40 assessment systems and intentional structures for student collaboration, approaches  
41 that can be integrated into clinical experiences through targeted coaching (California  
42 Department of Education, 2025; REL Pacific/IES, 2025).

43 Encouragingly, Program Completer Survey ratings and employer evaluations of  
44 graduates show strong alignment on the Teacher Quality Standards (TQS). This  
45 consistency reflects national guidance to use multiple measures of teaching quality

1 and underscores the importance of examining differences across raters and program  
2 stages as part of continuous improvement.

3 Even so, English Language Learner Educator Preparation Standards 5.10, 5.11,  
4 and 5.12 stand out as persistent areas needing attention. These standards, which focus  
5 on language acquisition, literacy development, and instructional strategies for CLD  
6 learners, mirror longstanding national calls for more explicit preparation in culturally  
7 and linguistically responsive pedagogy. Their recurrence across semesters suggests a  
8 system-wide need for more focused instruction and coached practice.

9 When the UTPP data are considered alongside the State EPP Report, several  
10 implications emerge. While internal assessments identify CLD-related competencies as  
11 areas for growth, state employment data show that UTPP graduates are more likely to  
12 teach in CLD dense schools. This combination underscores the need for an equity  
13 centered approach to program design and evaluation. Programs must intentionally  
14 strengthen preparation for CLD-responsive teaching, and evaluation systems must  
15 contextualize outcomes to avoid misinterpreting teacher effectiveness. Program efforts  
16 such as Project ELEECT (Donley et al., 2023) demonstrate promising approaches by  
17 integrating teacher research assignments that support candidates in intentionally  
18 incorporating academic language and culturally and linguistically responsive pedagogy  
19 into their planning, instruction, assessment, and reflective analysis.

20 Equity concerns also arise in teacher evaluation systems that rely heavily on  
21 growth measures. Researchers caution that when evaluations do not account for  
22 school characteristics, programs whose graduates serve high-need schools may  
23 inadvertently be penalized (Bastian et al., 2018). Systematic reviews likewise raise  
24 concerns about the validity and reliability of value-added measures in diverse contexts  
25 (Amrein Beardsley et al., 2023; Koedel, 2025). Together, these findings reinforce the  
26 need to triangulate teacher growth measures with CLD attuned observation rubrics,  
27 student work artifacts, and community informed indicators.

28 To respond effectively, teacher preparation program design should prioritize  
29 several strategies to support teacher candidates: (1) explicit planning for academic  
30 language demands supported by formative assessment; (2) culturally sustaining  
31 assessment practices, such as performance tasks that draw on students' linguistic  
32 repertoires; (3) clinical coaching protocols that prompt candidates to enact culturally  
33 and linguistically responsive pedagogy in authentic ways; and (4) alignment with  
34 current English Learner policy guidance to prepare graduates for advocacy and  
35 compliance in multilingual contexts.

36 State guidance emphasizes that Measures of Student Learning should include  
37 multiple measures, meaning statewide assessments as well as other rigorous,  
38 comparable forms of evidence, to balance accountability with instructional usefulness.  
39 Yet because UTPP graduates often teach in CLD dense schools where standardized  
40 tests undervalue student learning, Measures of Student Learning results may  
41 underestimate their effectiveness. This tension has fueled statewide and national calls  
42 to reconsider the function of summative tests in teacher evaluation (Colorado  
43 Department of Education, 2024–2025; Olson & Toch, 2024).

44 When employer evaluations show strong performance on the Teacher Quality  
45 Standards but comparatively lower Measures of Student Learning ratings for  
46 graduates in CLD dense districts, stakeholders should interpret results contextually.

1 Decisions related to TPP evaluation should incorporate curriculum embedded  
2 assessments, triangulate employer ratings with observation data and candidate  
3 portfolios (e.g., Program Completer Survey, Lesson Plan, Lesson Observation, and  
4 Teacher Work Sample) and interpret results in light of school context to avoid  
5 discouraging placements in high need schools. As research cautions, failure to account  
6 for employment context can disincentivize placements in high-need settings and  
7 exacerbate inequities (Bardelli, 2023).

8 Recent national analyses further recommend separating accountability functions  
9 from instructional assessment to ensure that tests support equity and learning rather  
10 than impose punitive consequences (Colorado Department of Education, 2024, 2025;  
11 Olson & Toch, 2024;). For UTPP, these shifts underscore the importance of  
12 performance-based tools, contextually relevant assessments, and candidates'  
13 reflective practices that encourage candidates to document their impact in ways that  
14 honor community languages, cultures, and identities.

## 17 **Conclusion**

18  
19 In the U.S., pushed by public concern for quality of teachers and teacher  
20 education, states have increasingly established systems to link teacher performance  
21 scores to the TPP from which teachers graduated (Bastian et al., 2016). The call for  
22 connections between TPP assessment and the evaluation of its graduates by school  
23 employers, as well as a desire to establish equity as a goal and outcome of program  
24 evaluation, inspired the current study. This case study draws on recent consensus  
25 guidance around the use of evidence for TPP improvement and accountability,  
26 contemporary critiques of value-added and standardized assessments, and state-level  
27 policy updates affecting educator preparation and multilingual learners. Using  
28 multiple program-level assessment tools, designed with equity in mind for teacher  
29 candidates working with diverse student populations, and addressing state licensure  
30 standards required of all TPPs, the study shows that it is possible to connect a TPP's  
31 evaluation system to how future employers evaluate its graduates. Additionally,  
32 equity needs to, and can be, intentionally taken into consideration for assessment  
33 design and implementation as well as data analysis to ensure that findings can more  
34 meaningfully impact direct program improvement.

35 There are limitations of the study which include the short data collection period  
36 and the delay in state reporting. Future research should track cohorts longitudinally,  
37 integrate culturally responsive assessment artifacts, and examine inter-rater  
38 calibration across contexts.

39 Although the study is based in the United States, schools and K-12 classroom  
40 teachers in the world are facing ever-increasing diversity. How should a teacher  
41 education program be evaluated considering the context in which their graduates will  
42 be teaching is a question that transcends national boundaries. The findings of this  
43 study shed light on future efforts to further investigate the topic. At the operational  
44 level, TPPs desiring to include equity as a goal or outcome in their program evaluation  
45 system need first to examine their program assessment tools for the inclusion of equity  
46 factors. Even more important is the need to connect the program assessment tools and

1 assessment data analysis to the reality of the working environment and context of the  
2 schools where graduates are hired to teach.

### 5 References

- 7 AACTE (2013). edTPA Support and Assessment Program. <https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/es/edtpa-features-fl.pdf>
- 8 Amrein, A. L., & Berliner, D. C. (2002). An analysis of some unintended and negative  
9 consequences of high-stakes testing. Education Policy Research Unit, Arizona State  
10 University. <http://nepc.colorado.edu/files/EPSSL-0211-125-EPRU.pdf>
- 11 Amrein-Beardsley, A., Ryan Lavery, M., Holloway, J., Pivovarova, M., & Hahs-Vaughn, D.  
12 L. (2023). Evaluating the validity evidence surrounding the use of value-added models to  
13 evaluate teachers: A systematic review. *Education Policy Analysis Archives*, 31, Article  
14 117. <https://doi.org/10.14507/epaa.31.8201>
- 15 Au, W. (2016). Meritocracy 2.0: High-stakes, standardized testing as a racial project of  
16 neoliberal multiculturalism. *Educational Policy*, 30 (1), 39–62.
- 17 Bach, A. J. (2020). High-stakes, standardized testing and emergent bilingual students in Texas.  
18 *Texas Journal of Literacy Education*, 8(1), 18–27.
- 19 Bales, B.L. (2015) Restructuring teacher education in the United States: Finding the tipping  
20 point. *Athens Journal of Education* 2(4): 297- 312. [https://www.athensjournals.gr/edu-  
21 cation/2015-2-4-1-Bales.pdf](https://www.athensjournals.gr/education/2015-2-4-1-Bales.pdf)
- 22 Bardelli, E., Ronfeldt, M., & Papay, J. P. (2023). Teacher preparation programs and graduates’  
23 growth in instructional effectiveness. *American Educational Research Journal*, 60(1),  
24 183–216.
- 25 Barron, L. (2015). Preparing pre-service teachers for performance assessments. *Journal of  
26 Interdisciplinary Studies in Education*, 3(2), 68–75.
- 27 Bastian, K. C., Henry, G. T., Pan, Y., & Lys, D. (2016). Teacher candidate performance  
28 assessments: Local scoring and implications for teacher preparation program  
29 improvement. *Teaching and Teacher Education*, 59, 1–12.
- 30 Bastian, K. C., Lys, D., & Whisenant, W. R. (2023). Does placement predict performance?  
31 Associations between student teaching environments and candidates’ performance  
32 assessment scores. *Journal of Teacher Education*, 74(1), 40–54.
- 33 Bastian, K. C., Patterson, K., & Pan, Y. (2018). Evaluating teacher preparation programs with  
34 teacher evaluation ratings. *Journal of Teacher Education*, 69(5), 429–447.
- 35 Berg, M. & Huang, J. (2015). Improving In-service Teachers' Effectiveness: K-12 Academic  
36 Literacy for the Linguistically Diverse. *Functional Linguistics*, 2 (5). [http://www.fun-  
37 ctionallinguistics.com/content/2/1/5](http://www.functionallinguistics.com/content/2/1/5)
- 38 Boon H.J. & Lewthwaite B.E. (2016) Teacher ethics: The link between quality teaching and  
39 multi-ethnic and multiracial education. *Athens Journal of Education* 3(4): 331 – 344.  
40 <https://www.athensjournals.gr/education/2016-3-4-3-Boon.pdf>
- 41 California Department of Education (2025). *Culturally Sustaining Pedagogy – Professional  
42 Learning* (Practice features that sustain community languages and identities). [[cde.ca.gov](http://www.cde.ca.gov)]  
43 <https://www.cde.ca.gov/ci/pl/culturallysustainingped.asp>
- 44 Cochran-Smith, M., & Reagan, E. M. (2021). Best practices for evaluating teacher preparation  
45 programs. <https://eric.ed.gov/?id=ED615385>
- 46 Colorado Department of Education (CDE) (2023 – 2024). *Educator Preparation Programs  
47 (EPP) Dashboard*. Retrieved January 2024. [https://ed.cde.state.co.us/educatoralent/re-  
48 searchhandimpact/educator-preparation-programs-report/epp-dashboard](https://ed.cde.state.co.us/educatoralent/researchhandimpact/educator-preparation-programs-report/epp-dashboard)
- 49

- 1 Colorado Department of Education (CDE) (2025). *ELPA FAQs (February 2025)* (Funding,  
2 EL definitions, obligations). [[cde.state.co.us](https://cde.state.co.us)]
- 3 Colorado Department of Education (CDE) (2024–2025). *Redesignation Guidance* (ACCESS  
4 proficiency; temporary pause for Alternate ACCESS standard setting). [https://www.cde.state.co.us/cde\\_english/sy24-25redesignationguidance](https://www.cde.state.co.us/cde_english/sy24-25redesignationguidance). [https://ed.cde.state.co.us/clde/eld](https://ed.cde.state.co.us/clde/eldrequirements/redesignation)  
5 requirements/redesignation
- 6  
7 Curtis, E. (2015) The values dimension of quality teachers: Can we prepare pre-service  
8 teachers for this? *Athens Journal of Education* 2(4): 313-330. <https://www.athensjournals.gr/education/2015-2-4-2-Curtis.pdf>
- 9  
10 Donley, K., Wesley-Nero, S., Stephens, C., Ashraf H., Reed D. (2023). *Preparing CLD-*  
11 *responsive teachers through teacher research: Project ELEECT. Journal of Language*  
12 *Teaching*, 3(10), 8–17. <https://jlt.ac/home/article/download/57/64> [https://jlt.ac/home/ar](https://jlt.ac/home/article/view/57)  
13 [ticle/view/57](https://jlt.ac/home/article/view/57)
- 14 Olson, L. & Toch, T. (2024). *None of the Above: A New Vision for State Standardized Testing.*  
15 FutureEd. [[future-ed.org](https://future-ed.org)]
- 16 Goldhaber, D., Liddle, S., & Theobald, R. (2013). The gateway to the profession: Assessing  
17 teacher preparation programs based on student achievement. *Economics of Education*  
18 *Review*, 34, 29–44.
- 19 Hood, S. L., Dilworth, M. E., & Lindsay, C. A. (2022). Landscape of teacher preparation  
20 program evaluation policies and progress. National Academy of Education Committee on  
21 Evaluating and Improving Teacher Preparation Programs. National Academy of  
22 Education. <https://doi.org/10.31094/2021/3/5>
- 23 Keith, H.R., Kinnison, S., Garth-McCullough, R., & Hampton M. (2023). Putting equity into  
24 practice: Culturally Responsive Teaching and Learning. Equity-Minded Digital Learning  
25 Strategy Guide Series. [https://achievethegoal.org/wp-content/uploads/2023/12/Cultu](https://achievethegoal.org/wp-content/uploads/2023/12/Culturally-Responsive-Teaching-and-Learning_Equity-Minded-Digital-Learning-Strategy-Guides_2023.pdf)  
26 [rally-Responsive-Teaching-and-Learning\\_Equity-Minded-Digital-Learning-Strategy-](https://achievethegoal.org/wp-content/uploads/2023/12/Culturally-Responsive-Teaching-and-Learning_Equity-Minded-Digital-Learning-Strategy-Guides_2023.pdf)  
27 [Guides\\_2023.pdf](https://achievethegoal.org/wp-content/uploads/2023/12/Culturally-Responsive-Teaching-and-Learning_Equity-Minded-Digital-Learning-Strategy-Guides_2023.pdf)
- 28 Knight, S. L., et al. (2014). Performance assessment of teaching: Implications for teacher  
29 education. *Journal of Teacher Education*, 65(5), 372–374.
- 30 Koedel, C. (2025). The reliability of growth measures in teacher evaluation. *Journal of*  
31 *Educational Measurement*, 62(1), 45–67.
- 32 Ladson-Billings, G. (2014). Culturally relevant pedagogy 2.0: a.k.a. the Remix. *Harvard*  
33 *Educational Review*, 84(1) 74-84. <https://doi.org/10.17763/haer.84.1.p2rj131485484751>
- 34 Levy-Feldman, I. (2025). The Role of Assessment in Improving Education and Promoting  
35 Educational Equity. *Education Sciences*, 15(2), 224. [https://doi.org/10.3390/educsci15](https://doi.org/10.3390/educsci15020224)  
36 [020224](https://doi.org/10.3390/educsci15020224)
- 37 Miller, R. D., Gerzel-Short, L., Kelly, J. H., Uphold, N., Hsiao, Y., Hovey, K. A., & Wei, Y.  
38 (2024). *Preparing culturally responsive teachers: A systematic review. Multicultural*  
39 *Learning and Teaching*, 19(2), 249–268. [[degruyterbrill.com](https://degruyterbrill.com)]
- 40 National Academy of Education (2024). *Evaluating and Improving Teacher Preparation*  
41 *Programs*. Washington, DC. [https://naeducation.org/wp-content/uploads/2025/01/Evalua](https://naeducation.org/wp-content/uploads/2025/01/Evaluating_and_Improving_Teacher_Preparation_Programs.pdf)  
42 [ting\\_and\\_Improving\\_Teacher\\_Preparation\\_Programs.pdf](https://naeducation.org/wp-content/uploads/2025/01/Evaluating_and_Improving_Teacher_Preparation_Programs.pdf) [https://naeducation.org/wp-con](https://naeducation.org/wp-content/uploads/2025/01/Evaluating_and_Improving_Teacher_Preparation_Programs.pdf)  
43 [tent/uploads/2025/01/Evaluating\\_and\\_Improving\\_Teacher\\_Preparation\\_Programs.pdf](https://naeducation.org/wp-content/uploads/2025/01/Evaluating_and_Improving_Teacher_Preparation_Programs.pdf)
- 44 National Academy of Education (2025). *Evaluating and Improving Teacher Preparation*  
45 *Programs – Executive Summary*. Washington, DC. [https://naeducation.org/wp-content/](https://naeducation.org/wp-content/uploads/2025/02/Evaluating-and-Improving-Teacher-Preparation-Programs-Executive-Summary-FINAL.pdf)  
46 [uploads/2025/02/Evaluating-and-Improving-Teacher-Preparation-Programs-Executive-](https://naeducation.org/wp-content/uploads/2025/02/Evaluating-and-Improving-Teacher-Preparation-Programs-Executive-Summary-FINAL.pdf)  
47 [Summary-FINAL.pdf](https://naeducation.org/wp-content/uploads/2025/02/Evaluating-and-Improving-Teacher-Preparation-Programs-Executive-Summary-FINAL.pdf) [https://naeducation.org/wp-content/uploads/2025/02/Evaluating-](https://naeducation.org/wp-content/uploads/2025/02/Evaluating-and-Improving-Teacher-Preparation-Programs-Executive-Summary-FINAL.pdf)  
48 [and-Improving-Teacher-Preparation-Programs-Executive-Summary-FINAL.pdf](https://naeducation.org/wp-content/uploads/2025/02/Evaluating-and-Improving-Teacher-Preparation-Programs-Executive-Summary-FINAL.pdf)
- 49 Paris, D. & Alim H. S. (2014). What are we seeking to sustain through culturally sustaining  
50 pedagogy? A loving critique forward. *Harvard Educational Review*, 84(1) 85-100.  
51 <https://doi.org/10.17763/haer.84.1.9821873k2ht16m77>

1 REL Pacific/IES (2025). *Culturally Responsive Assessment: Goals, Challenges, and Implications*.  
2 <https://ies.ed.gov/rel-pacific/2025/01/infographic-30> [https://ies.ed.gov/sites/default/files/](https://ies.ed.gov/sites/default/files/migrated/rel/infographics/pdf/REL_PA_Culturally_Responsive_Assessment.pdf)  
3 [migrated/rel/infographics/pdf/REL\\_PA\\_Culturally\\_Responsive\\_Assessment.pdf](https://ies.ed.gov/sites/default/files/migrated/rel/infographics/pdf/REL_PA_Culturally_Responsive_Assessment.pdf)  
4 Ronfeldt, M., & Campbell, S. L. (2016). Evaluating teacher preparation using graduates'  
5 observational ratings. *Educational Evaluation and Policy Analysis*, 38(4), 603–625.  
6 Sandoval, C., Cunningham, J., Jacobs, J., & Warner, S. (2023). *Research and Evidence-Based*  
7 *Best Practices for Preparing Educators for Culturally Responsive Teaching and Leading*.  
8 WestEd. <https://www.isbe.net/Documents/WestEd-Research-Brief.pdf>  
9 [https://www.isbe.](https://www.isbe.net/Documents/WestEd-Research-Brief.pdf)  
10 [net/Documents/WestEd-Research-Brief.pdf](https://www.isbe.net/Documents/WestEd-Research-Brief.pdf)  
11  
12  
13

ONLY FOR REVIEW

1 **Appendix**

2

3 *State Teacher Licensure Standard*

4 *Teacher Quality Standards (TQS)*

5

6 5.01: Teachers demonstrate mastery of and pedagogical expertise in the content they  
7 teach. The elementary teacher is an expert in literacy and mathematics and is  
8 knowledgeable in all other content that he or she teaches (e.g., science, social studies,  
9 the arts, physical education or world languages). The secondary teacher has  
10 knowledge of literacy and mathematics and is an expert in the content area(s) in which  
11 the teacher is endorsed.

12 5.01(1): Teachers provide instruction that is aligned with the Colorado Academic  
13 Standards and their district's organized plan of instruction.

14 5.01(2): Teachers develop and implement lessons that connect to a variety of  
15 content areas/disciplines and emphasize literacy and mathematics

16 5.01(3): Teachers demonstrate knowledge of the content, central concepts, inquiry,  
17 appropriate evidence-based instructional practices and specialized characteristics  
18 of the disciplines being taught.

19 5.02: Teachers establish a safe, inclusive and respectful learning environment for a  
20 diverse population of students.

21 5.02(1): Teachers foster a predictable learning environment characterized by  
22 acceptable student behavior and efficient use of time in which each student has a  
23 positive, nurturing relationship with caring adults and peers.

24 5.02(2): Teachers demonstrate an awareness of, a commitment to, and a respect  
25 for multiple aspects of diversity, while working toward common goals as a  
26 community of learners.

27 5.02(3): Teachers engage students as individuals, including those with diverse  
28 needs and interests, across a range of ability levels by adapting their teaching for  
29 the benefit of all students.

30 5.02(4): Teachers work collaboratively with the families and/or significant adults  
31 for the benefit of students.

32 5.03: Teachers plan and deliver effective instruction and create an environment that  
33 facilitates learning for their students.

34 5.03(1): Teachers demonstrate knowledge about the ways in which learning takes  
35 place, including the levels of intellectual, physical, social, and emotional  
36 development of their students.

37 5.03(2): Teachers use formal and informal methods to assess student learning,  
38 provide feedback, and use results to inform planning and instruction.

39 5.03(3): Teachers integrate and utilize appropriate, available technology to engage  
40 students in authentic learning experiences.

41 5.03(4): Teachers establish and communicate high expectations and use processes  
42 to support the development of critical-thinking and problem-solving skills.

43 5.03(5): Teachers provide students with opportunities to work in teams and  
44 develop leadership.

45 5.03(6): Teachers model and promote effective communication.

1 5.04: Teachers demonstrate professionalism through ethical conduct, reflection, and  
2 leadership.

3 5.04(1): Teachers demonstrate high standards for professional conduct.

4 5.04(2): Teachers link professional growth to their professional goals.

5 5.04(3): Teachers are able to respond to a complex, dynamic environment.

6 5.04(4): Teachers demonstrate leadership in their school, the community, and the  
7 teaching profession.

8

9 *English Language Learner Educator Preparation Standards*

10

11 5.09: Quality Standard I: Educators are knowledgeable about CLD populations.

12 5.09(1): ELEMENT A: Educators are knowledgeable in, understand, and able to  
13 apply the major theories, concepts and research related to culture, diversity and  
14 equity in order to support academic access and opportunity for CLD student  
15 populations.

16 5.09(2): ELEMENT B: Educators are knowledgeable in, understand, and able to  
17 use progress monitoring in conjunction with formative and summative  
18 assessments to support student learning.

19 5.10: Quality Standard II: Educators should be knowledgeable in first and second  
20 language acquisition.

21 5.10(1): ELEMENT A: Educators are able to understand and implement strategies  
22 and select materials to aid in English language and content learning.

23 5.10(2): ELEMENT B: Educators are knowledgeable of, understand, and able to  
24 apply the major theories, concepts and research related to culture, diversity and  
25 equity in order to support academic access and opportunity for CLD student  
26 populations.

27 5.11: Quality Standard III: Educators should understand literacy development for  
28 CLD students.

29 5.11(1): ELEMENT A: Educators are knowledgeable in, understand, and able to  
30 apply the major theories, concepts and research related to literacy development for  
31 CLD students

32 5.11(2): ELEMENT B: Educators understand and implement strategies and select  
33 materials to aid in English language and content learning.

34

35 5.12: Quality Standard IV: Educators are knowledgeable in the teaching strategies,  
36 including methods, materials, and assessment for CLD students.

37 5.12(1): ELEMENT A: Educators are knowledgeable in, understand and able to  
38 use the major theories, concepts and research related to language acquisition and  
39 language development for CLD students.

40 5.12(2): ELEMENT B: Educators are knowledgeable in, understand, and able to  
41 use progress monitoring in conjunction with formative and summative  
42 assessments to support student learning.