International Education and Culture Studies
ISSN 2766-6778
Vol. 6, No. 4, 2026
Published by STSL Press

Original Paper

Using Candidate and Employer Feedback for Continuous Improvement in a University-Based Teacher Preparation Program: A Descriptive Program Evaluation

John Peter Petrone, Michael F. Shaughnessy, Kari Lemke

Eastern New Mexico University

Abstract

Teacher preparation programs need evidence that is useful for accountability and continuous improvement, yet no single measure can adequately represent program quality. This descriptive program evaluation examined stakeholder feedback from a university-based teacher preparation program in the rural U.S. Southwest. Teacher candidates responded to a Qualtrics survey across three administrations conducted between fall 2023 and spring 2025, and 25 employer respondents representing 17 schools completed the employer portion of the Qualtrics instrument. Both groups responded using seven-point Likert-type response options. Employer ratings across 14 professional competencies ranged from 6.17 to 6.83, and 90% of responses to the overall satisfaction item indicated satisfaction with program graduates. Candidate ratings were also favorable: clinical-experience and support ratings ranged from 5.60 to 6.70, instructional-readiness ratings ranged from 6.27 to 6.87, and faculty-support ratings ranged from 6.13 to 6.37. The highest candidate rating concerned lesson planning, whereas comparatively lower ratings concerned seminars, cohort meetings, and communication with families. Employer ratings identified reflective practice and acceptance of constructive criticism as relative growth areas. Because the evaluation relied on self-report data, summary-level descriptive statistics, and a single institutional context, the findings should be used to inform local improvement rather than to support causal claims. The study demonstrates how candidate and employer feedback can be used together to identify both program strengths and actionable priorities.

Keywords: teacher preparation, program evaluation, teacher candidates, employer feedback, continuous improvement

Using Candidate and Employer Feedback for Continuous Improvement in a University-Based Teacher Preparation Program

Evaluating a teacher preparation program is both necessary and methodologically difficult. Programs are expected to demonstrate that candidates acquire the knowledge, skills, and professional dispositions needed for classroom practice, while also using evidence to improve coursework, clinical experiences, and candidate support. These purposes overlap but are not identical. Evidence adequate for regulatory reporting may not explain why a particular program feature is working, and evidence useful to faculty may not support comparisons across institutions. A responsible evaluation therefore begins by identifying its purpose, selecting evidence suited to that purpose, and limiting claims to what the evidence can support.

The National Academy of Education (2024) identifies three broad purposes for teacher preparation evaluation: program improvement, accountability, and consumer information. Its consensus report further emphasizes that program quality is multidimensional, encompassing coherence and alignment, curriculum, instructional methods, clinical experiences, candidate recruitment and support, and faculty capacity. Similarly, the Council for the Accreditation of Educator Preparation (CAEP, 2022) calls for quality-assurance systems built on multiple measures and sustained stakeholder involvement. These frameworks position evaluation as a recurring cycle of evidence gathering, interpretation, action, and reassessment rather than as a single compliance event.

The use of multiple measures is especially important because common indicators capture different aspects of quality and carry different limitations. Worrell et al. (2014) described surveys, standardized observations, and measures of student learning as potentially useful sources of evidence while stressing validity and appropriate interpretation. Cochran-Smith et al. (2015) likewise characterized teacher preparation research as a broad and uneven field in which conclusions depend heavily on the questions asked, the outcomes selected, and the contexts examined. In a peer-reviewed synthesis of 166 impact evaluations, Mancenido (2024) found that relatively few studies adequately addressed internal validity, external validity, and measurement. No isolated survey result, observation score, or outcome measure can therefore function as a comprehensive judgment of program effectiveness.

Stakeholder Surveys as Improvement Evidence

Candidate and employer surveys remain useful because they provide perspectives that administrative records cannot. Candidates can report on program coherence, clinical-placement quality, faculty access, and their perceived readiness to carry out specific professional tasks. Employers can assess how graduates perform in school settings and whether preparation appears relevant to assigned responsibilities. CAEP's current standards explicitly identify employer and completer satisfaction as evidence of program impact and call for both internal and external stakeholders to participate in continuous improvement (CAEP, 2022).

Survey evidence must nevertheless be interpreted carefully. Candidate perceptions of preparation quality have shown modest associations with later measures of teacher effectiveness and retention, but those relationships are not uniform and can be affected by selection, school placement, timing, and differential response rates (Bastian et al., 2021). Institution-developed candidate surveys can support program reflection when items are aligned to program outcomes and when reliability and validity are examined (Bhatnagar et al., 2014). In the absence of such psychometric evidence, item ratings are best treated as stakeholder perceptions rather than direct measures of teaching competence. This distinction is particularly important when respondents provide uniformly high ratings, because ceiling effects may conceal meaningful variation.

Clinical experiences also merit focused evaluation. Research syntheses link stronger preparation and retention outcomes to coherent clinical experiences, effective cooperating teachers, productive school environments, and high-quality coaching (Ronfeldt, 2021). Program-level feedback about placement quality, supervision, seminars, and school collaboration can therefore identify where clinical structures support candidates and where additional attention is warranted. Darling-Hammond et al. (2010) demonstrated the value of combining candidate and graduate perceptions with other forms of program evidence during a sustained redesign process. Their work illustrates that survey results are most useful when they are connected to a theory of program improvement and examined alongside additional outcomes.

Program Context and Purpose

The evaluated program is housed at Eastern New Mexico University, a regional public university serving a largely rural area of the U.S. Southwest. The teacher preparation program includes pathways in early childhood, elementary, and secondary education. Candidate progression includes an admissions review, a minimum grade point average requirement, an interview, methods coursework, preparation related to culture and learner differences, supervised clinical practice, and a culminating student-teaching experience. These transition points are intended to assess readiness while providing structured support as candidates move from coursework into professional practice.

The present study reports an initial analysis of stakeholder data collected through a Qualtrics survey instrument. Its purpose is not to establish causal program effects or to rank the program against other institutions. Rather, the study asks whether candidate and employer feedback yields a coherent description of perceived strengths and actionable priorities for continuous improvement. Three research questions guided the evaluation:

1. How do employers perceive program graduates' professional competencies and overall preparation?

2. How do teacher candidates perceive the quality of their clinical experience, their readiness for professional practice, and the support they receive from program faculty?

3. What program strengths and improvement priorities emerge when candidate and employer feedback are considered together?

Method

Design

This study used a descriptive program-evaluation design. The analysis combined two stakeholder perspectives collected for institutional improvement: feedback from teacher candidates engaged in student teaching and feedback from school-based employers. Mancenido (2024) distinguishes impact evaluations, which are designed to estimate the causal effects of particular preparation practices, from process-oriented evaluations that examine implementation and stakeholder experience. The present study belongs to the latter category. It uses local evidence to identify strengths, examine program delivery, and establish priorities for subsequent inquiry. Because the data are observational and self-reported, the analysis is descriptive rather than causal.

Participants and Data Sources

Candidate data were collected across three survey administrations between fall 2023 and spring 2025. The response distribution reported for the professional-learning item indicated 30 candidate respondents. The respondents were teacher candidates completing student teaching and were asked about placement and program supports, perceived readiness for teaching responsibilities, preparation to work with learners from varied backgrounds, collaboration, commitment to continued learning, and faculty support.

The employer dataset included 25 respondents representing 17 schools. Sixty-eight percent reported that their school had hired graduates of the university as teachers or counselors. The employer portion of the survey asked respondents to rate graduates across 14 professional competencies and to report overall satisfaction. The available summary did not include respondent roles, school-level demographic characteristics, years of experience with program graduates, or a response-rate denominator.

Instrument and Data Collection

Respondents completed the program-evaluation survey through Qualtrics. Stakeholder-specific items used a seven-point Likert-type scale: 1 (strongly disagree), 2 (disagree), 3 (somewhat disagree), 4 (neutral), 5 (somewhat agree), 6 (agree), and 7 (strongly agree). Candidate items addressed clinical placement and support, instructional readiness, work with varied learners, professional collaboration, continued learning, and faculty support. Employer items addressed professional conduct, communication, collaboration, instructional preparation, assessment, classroom management, reflection, and responsiveness to feedback. The survey also included an overall employer satisfaction item.

The instrument was developed for program evaluation and administered electronically. The archival summary did not report pilot testing, evidence of content validity beyond alignment with program concerns, internal-consistency estimates, or factor structure. Accordingly, the present analysis reports individual items and conceptually related domains but does not treat those domains as validated scales.

Analysis

The evaluation used descriptive statistics. Item means were calculated for the seven-point ratings, and percentages were reported for overall employer satisfaction and the candidate continued-learning item. Results were organized by stakeholder and by content domain. To identify practical priorities, the analysis considered both absolute ratings and relative position within each set of uniformly positive scores. No inferential tests were conducted because the evaluation was not designed for between-group comparison, and the available summary did not include standard deviations or candidate-level records. Mean values originally reported to three decimal places were rounded to two decimals for consistent presentation.

Results

Employer Perceptions of Graduate Preparation

Employer ratings were strongly favorable across all 14 competencies, with means ranging from 6.17 to 6.83 on the seven-point scale (Table 1). The highest-rated area was evaluation of student progress (M = 6.83), followed by professional ethics and expectations (M = 6.78), adaptability (M = 6.74), and communication skills (M = 6.72). Employers also rated graduates highly in rapport with students and colleagues, respect for the school community, collaboration, instructional preparedness, classroom management, use of assessment data, and professional responsibility.

Reflective practice (M = 6.22) and acceptance of constructive criticism (M = 6.17) were the two lowest-rated items. These values still fall above the scale point labeled agree, so they should not be characterized as deficiencies. Rather, their relative position within an otherwise compressed distribution identifies them as areas for closer examination. On the overall satisfaction item, 90% of responses indicated satisfaction with program graduates and 10% indicated dissatisfaction. The source summary attributed the negative responses to an individual case and did not identify a recurring concern across schools; however, the available data do not support dismissing those responses without follow-up.

Table 1. Employer Ratings of Graduates' Professional Competencies

Competency	***M***
Evaluation of student progress	6.83
Professional ethics and expectations	6.78
Adaptability	6.74
Communication skills	6.72
Rapport with students	6.65
Rapport with colleagues	6.63
Respect for the school community	6.60
Collaboration with colleagues	6.58
Instructional preparedness	6.55
Classroom management	6.52
Use of assessment data	6.50
Professional responsibility	6.48
Reflective practice	6.22
Acceptance of constructive criticism	6.17

Note. Items were rated on a seven-point scale from 1 (strongly disagree) to 7 (strongly agree). Employer respondents represented 17 schools (N = 25).

Candidate Perceptions of Clinical Experience and Support

Candidate ratings of the student-teaching experience ranged from 5.60 to 6.70 (Table 2). Quality of placement received the highest rating (M = 6.70), followed by university supervision and placement diversity (both M = 6.20). The application and placement process received a mean of 6.07. The initial

seminar (*M* = 5.83), cohort meetings (*M* = 5.67), and Jump Start sessions (*M* = 5.60) received the lowest ratings in this group. Although these means remain above somewhat agree, they show more room for improvement than the placement and supervision ratings.

Table 2. Candidate Ratings of the Student-Teaching Experience

Program element	***M***
Quality of placement	6.70
Quality of university supervision	6.20
Diversity of placement	6.20
Application and placement process	6.07
Quality of initial seminar	5.83
Quality of cohort meetings	5.67
Quality of Jump Start sessions	5.60

Note. Items were rated on a seven-point scale from 1 (strongly disagree) to 7 (strongly agree).

Candidate Perceptions of Readiness

Candidates reported high perceived readiness across instructional, learner-diversity, and collaboration items (Table 3). Within instructional readiness, creating a lesson plan received the highest rating (*M* = 6.87), followed by using engaging teaching methods (*M* = 6.72), effectively teaching in the content field (*M* = 6.67), and fostering classroom collaboration (*M* = 6.67). Theories of human development received the lowest instructional-readiness rating (*M* = 6.27), though it remained above agree.

Ratings related to varied learners ranged from 6.53 to 6.73. Assessment of student learning received the highest rating (*M* = 6.73), whereas communication with families received the lowest rating (*M* = 6.53). Candidates also reported high confidence in working with school colleagues (*M* = 6.77) and administrators (*M* = 6.57); working with parents was lower by comparison (*M* = 6.30). The repeated relative position of family- and parent-communication items across domains makes family engagement a more compelling improvement priority than either item would suggest in isolation.

Table 3. Candidate Ratings of Perceived Readiness for Professional Practice

Domain	**Item**	***M***
Instruction	Create a lesson plan	6.87
Instruction	Use engaging teaching methods	6.72
Instruction	Effectively teach in my content field	6.67
Instruction	Foster classroom collaboration	6.67
Instruction	Develop curricula in my content field	6.57
Instruction	Manage my classroom	6.40
Instruction	Use theories of human development	6.27
Varied learners	Assess student learning	6.73

Varied learners	Work with varied academic backgrounds	6.70
Varied learners	Address varied learning preferences	6.70
Varied learners	Enhance student learning	6.70
Varied learners	Improve classroom management	6.63
Varied learners	Work with varied ethnic backgrounds	6.57
Varied learners	Communicate with families	6.53
Collaboration	Work with school colleagues	6.77
Collaboration	Work with school administrators	6.57
Collaboration	Work with parents	6.30

Note. Items were rated on a seven-point scale from 1 (strongly disagree) to 7 (strongly agree). The label varied learners is used here to summarize items referring to different populations and backgrounds.

Professional Learning and Faculty Support

Candidates expressed a strong commitment to continued professional learning. On the future learning and development item, 83.33% selected strongly agree, 13.33% selected agree, and 3.33% selected somewhat agree; no respondent selected a neutral or negative option. Perceptions of faculty support were also favorable. Candidates rated faculty involvement in progress toward program completion at 6.37, accessibility at 6.33, and efforts to know candidates personally at 6.13 (Table 4).

Table 4. Candidate Perceptions of Faculty Support and Continued Learning

Area	Item or response	Result
Faculty support	Involved in my progress toward program completion	6.37
Faculty support	Accessible to me	6.33
Faculty support	Took time to get to know me	6.13
Continued learning	Strongly agree	83.33%
Continued learning	Agree	13.33%
Continued learning	Somewhat agree	3.33%

Note. Faculty-support results are means on the seven-point scale. Continued-learning results are percentages for a single item; the distribution corresponds to 30 responses.

Discussion

This evaluation examined what candidate and employer feedback reveals about a university-based teacher preparation program. Across both stakeholder groups, the results were consistently favorable. Employers rated graduates at or above 6.17 on every professional competency, and candidates rated most areas of preparation above 6.00. The two sources also converged around several strengths: instructional planning and assessment, professional conduct, communication, collaboration, adaptability, and support during clinical practice. This convergence is useful because candidates and employers observe preparation from different positions. Candidates describe their experience and perceived readiness, whereas employers observe graduates in school contexts. Agreement across those

perspectives strengthens the case for retaining and further examining the program features associated with those ratings.

At the same time, uniformly high ratings require disciplined interpretation. The results show that stakeholders perceived the program and its graduates positively; they do not demonstrate that every candidate reached the same level of competence or that the program caused later teaching outcomes. Candidate ratings are perceptions recorded during student teaching, and employer ratings may reflect selective hiring, differing levels of contact with graduates, or respondents who are especially connected to the university. High mean scores also suggest possible ceiling effects that may reduce the instrument's ability to distinguish among program elements. For these reasons, the most defensible use of the data is formative: the ratings can guide questions, identify areas for follow-up, and inform improvement decisions.

Implications for Continuous Improvement

Four priorities emerge from the combined evidence. First, the program should examine family engagement and parent communication. Communication with families was the lowest-rated item in the varied-learners group, and working with parents was the lowest-rated collaboration item. Although both means were positive, their recurrence across domains suggests a coherent need. Program faculty could map where family-engagement knowledge and practice appear in coursework and clinical experiences, then add rehearsals, case analyses, or structured opportunities for candidates to communicate with families under supervision.

Second, the program should review the purpose, sequence, and format of the initial seminar, Jump Start sessions, and cohort meetings. These supports received the lowest ratings in the clinical-experience group. Follow-up focus groups or short open-response items could determine whether candidates perceive duplication, scheduling burdens, insufficient relevance, or a weak connection to immediate clinical needs. The goal should not be to remove support simply because ratings are lower, but to ensure that each structure performs a clear and distinct function.

Third, reflective practice and response to constructive criticism deserve more explicit attention. Employers rated both areas positively but lower than other competencies. Because reflection and use of feedback are central to professional growth, the program could examine whether candidates receive repeated opportunities to analyze evidence of student learning, respond to coaching, revise instruction, and document how feedback changed their practice. Employer follow-up should ask for behavioral examples and determine whether the concern is concentrated among novice graduates or linked to particular settings.

Fourth, the survey system itself should be strengthened. CAEP (2022) emphasizes evidence that is relevant, verifiable, representative, cumulative, and actionable. The next administration should preserve item-level response counts, respondent characteristics appropriate to the evaluation purpose, missing-data patterns, and administration-specific results. Program faculty should review the exported Qualtrics instrument for item clarity, double-barreled wording, and alignment with program outcomes. Reliability evidence may be appropriate for any intended multi-item scale, but psychometric analysis should follow a clearly articulated construct rather than be used to justify combining loosely related items. Adding a small number of carefully targeted open-response prompts would also help explain quantitative patterns without making the instrument burdensome.

These actions would move the evaluation from a positive snapshot toward a functioning improvement cycle. For each priority, faculty should specify the change to be made, the evidence expected if the change works, the timeline for reassessment, and the decision rule for retaining or revising the change. Candidate and employer surveys should then be considered alongside clinical-observation data, performance assessments, licensure outcomes, employment and retention data, and evidence of graduates' classroom practice. Such triangulation is consistent with recommendations that teacher preparation evaluation use multiple forms of evidence matched to distinct questions (National Academy of Education, 2024; Worrell et al., 2014).

Limitations and Future Research

Several limitations constrain the conclusions. The evaluation represents one institution and a relatively small employer sample, so the results are context-specific. Response rates were not available, and the employer respondents were drawn from 17 schools but were not described by role, school context, or degree of contact with graduates. The candidate dataset was pooled across three administrations, which may conceal changes over time or differences among program pathways. Demographic and pathway information was not available for disaggregated analysis.

The measures were self-reported, and the available summary did not include standard deviations, item-level response counts for most items, missing-data information, or evidence of instrument validity and reliability. The employer satisfaction percentages also require cautious interpretation because the item-level denominator was not preserved in the summary. In addition, the extremely high means suggest possible ceiling effects and social desirability or affiliation bias. Finally, the study did not include direct observations, performance assessments, graduate retention, or P-12 student outcomes; it therefore cannot establish program impact on teaching effectiveness or student learning. These constraints exemplify the design and measurement limitations that Mancenido (2024) cautions must be resolved before causal claims about teacher-preparation practices are warranted.

Future evaluation should follow cohorts from program entry through completion and into the early career years. Administration-specific reporting would permit the program to determine whether identified changes are followed by improvement. Stratified analyses could examine variation by licensure pathway, placement context, and respondent relationship to the program, provided sample sizes permit responsible reporting. Qualitative interviews or focus groups could explain why candidates rate family communication and selected clinical supports differently from other areas. Most importantly, subsequent studies should connect stakeholder perceptions with performance-based measures so that the program can examine where evidence converges and where it does not.

Conclusion

Candidate and employer feedback provides a favorable but appropriately bounded view of this university-based teacher preparation program. Employers perceived graduates as professional, adaptable, collaborative, and capable of evaluating student progress. Candidates reported strong preparation in lesson planning, instructional practice, assessment, collaboration, work with varied learners, and continued professional learning. The same data identify practical priorities involving family engagement, the design of selected clinical supports, reflective practice, responsiveness to feedback, and the quality of the evaluation system itself.

The contribution of this preliminary evaluation lies less in the uniformly high scores than in how the two stakeholder perspectives can be used together. When interpreted as local evidence rather than proof of causal effectiveness, the findings offer a credible starting point for continuous improvement. A stronger next phase will preserve complete survey metadata, validate intended scales, examine results over time, incorporate explanatory qualitative evidence, and triangulate perceptions with direct measures of candidate and graduate performance.

References

Bastian, K. C., Sun, M., & Lynn, H. (2021). What do surveys of program completers tell us about teacher preparation quality? *Journal of Teacher Education, 72*(1), 11-26. https://doi.org/10.1177/0022487119886294

Bhatnagar, R., Kim, J., & Many, J. E. (2014). Candidate surveys on program evaluation: Examining instrument reliability, validity and program effectiveness. *American Journal of Educational Research, 2*(8), 683-690. https://doi.org/10.12691/education-2-8-18

Council for the Accreditation of Educator Preparation. (2022). *CAEP revised 2022 standards for initial-level programs*. https://caepnet.org/caep-standards/

Cochran-Smith, M., Villegas, A. M., Abrams, L., Chavez-Moreno, L., Mills, T., & Stern, R. (2015). Critiquing teacher preparation research: An overview of the field, part II. *Journal of Teacher*

Education, 66(2), 109-121. https://doi.org/10.1177/0022487114558268

Darling-Hammond, L., Newton, X., & Wei, R. C. (2010). Evaluating teacher education outcomes: A study of the Stanford Teacher Education Programme. *Journal of Education for Teaching, 36*(4), 369-388. https://doi.org/10.1080/02607476.2010.513844

Mancenido, Z. (2024). Impact evaluations of teacher preparation practices: Challenges and opportunities for more rigorous research. *Review of Educational Research, 94*(2), 268-307. https://doi.org/10.3102/00346543231174413

National Academy of Education. (2024). *Evaluating and improving teacher preparation programs* (K. M. Zeichner, L. Darling-Hammond, A. I. Berman, D. Dong, & G. Sykes, Eds.). https://doi.org/10.31094/2024/2

Ronfeldt, M. (2021). *Links among teacher preparation, retention, and teaching effectiveness*. National Academy of Education. https://doi.org/10.31094/2021/3/1

Worrell, F. C., Brabeck, M. M., Dwyer, C. A., Geisinger, K. F., Marx, R. W., Noell, G. H., & Pianta, R. C. (2014). *Assessing and evaluating teacher preparation programs*. American Psychological Association. https://eric.ed.gov/?id=ED583083