Measure 3: Candidate Competency at Completion Dispositions Evaluation (2023-2024)

Instrument Validity and Reliability

The dispositions rubric underwent both **validity** and **reliability** studies during the 2022-2023 academic year. Relying on experts from the EPP Advisory Board as well as our Vice President for Institutional Effectiveness, we used Lawshe's (1975) method to ensure items were important, representative, and clear and to generate content validity ratios. Although all our items met Wilson et al.'s (2012) critical value for importance, we revisited items without perfect agreement; likewise, we revised items without perfect agreement for representativeness and clarity.

In each of the three areas of interest (importance, representativeness, and clarity), we used item content validity ratios to calculate a scale-level index (specifically, S-CVI/Ave for representativeness and clarity and the S-CVI/UA for importance). The S-CVI/Ave for representativeness was .984 and for clarity was .985. The S-CVI/UA for importance was .974. All CVIs were above the recommended minimum of .8 (Davis, 1992).

We then tested the reliability of the revised instrument. A sample of supervisor teachers was recruited to provide data. The pilot study was designed to test the internal consistency reliability of the rubric sub-scales. Participants (n = 32) were asked to complete the rubric based on a student teacher they had supervised in the last 5 years. The rubric contains 38 items with 8 sub-scales (see Table 1). To orient participants in completing the rubric, each supervisor was asked to describe three characteristics of the teacher they were rating. The sample included preschool (n = 1), elementary (n = 16), middle (n = 5), and high school (n = 10) teachers who supervised a student teacher between the fall of 2018 and the fall of 2022.

The seven sub-scales demonstrated strong internal consistency reliability, ranging from .82 to .96 (see Table 1) and exceed research standards for reliability coefficients. There is one additional item that measures excitement about teaching that was not evaluated for reliability.

Table 1

Internal Consistency of the Dispositions Rubric

			Scale	Cronbach Alpha
	Number Scale		Standard	Reliability
Sub-Scale	of items	Mean	Deviation	Coefficient (a)
Demonstrates sensitivity toward	7	47.47	8.65	.96
students				
Participates with others in a	5	32.72	7.53	.93
collaborative manner				
Treats others with respect	6	42.31	6.57	.92
Demonstrates lifelong learning	5	33.28	7.25	.97
Participates in professional	2	14.09	2.60	.82
development				
Demonstrates effective decision-	6	39.78	7.67	.92
making skills				
Works effectively with diverse	6	45.29	6.73	.94
learners				
Displays excitement about teaching	1	7.16		

. . . .

Instrument Administration and Use

Candidates are independently scored multiple times by their clinical instructor (CI) as well as their assigned college faculty (CF) member; administrations are characterized as weekly, midterm, or final. In addition to numerical scores, CIs and CFs may leave qualitative feedback for each subscale and at the end of the instrument. Candidates access their rubrics via Watermark, where they can see scores from their CI and CF as well as complete their own selfassessment. Easy access to their quantitative and qualitative feedback helps candidate see and respond to suggestions in a timely way—that is, helps the instrument function formatively—and provides a starting point for conferences with their CI and CF. Scores from the final administration are used to demonstrate candidate competency.

Results (2023-2024)

We used the revised dispositions rubric for the first time during the 2023-2024 academic year. We had a small group of completers in student teaching (n = 5) including two elementary, two secondary, and one special education candidate. Because of the size of the cohort, we cannot disaggregate data by licensure area or other demographic categories without compromising candidate anonymity. As we have done in previous years, our analysis of inter-rater agreement included percent of absolute and adjacent agreement, alerting us to any major disagreement between the scorers.

Although the small size of our cohort also limits the interpretations we can make about the descriptive statistics summarized in Table 2, these statistics are still useful as we move forward. As Table 2 shows, CIs and CFs are consistent in what they rate most highly (candidate excitement about teaching). Areas for improvement include working with diverse learners and participation in professional development. All sub-scale and the overall instrument means, however, exceeded the EPP's target performance level (5).

CIs consistently rate candidates higher; this is true for the instrument overall and on each sub-scale. CI scores also have a smaller standard deviation, indicating that their scores are more similar to one another than CF scores are. Percent agreement echoes these differences. While overall percent agreement is acceptable, the differences in the sub-scale means related to collaboration and participation in professional development suggest that the two groups of scorers may have different expectations. The 2024-2025 cohort is much larger than this year's, so we anticipate that analysis of their data alongside those data presented here will help us better identify strengths and weaknesses in the instrument. In the coming year, we remain focused on clear training for CIs and CFs with opportunities to view, rate, and discuss recorded teaching samples.

Table 2

2023-2024 CI and CF Mean Scores, Standard Deviations, and Percent Agreement

	CI		CF		
Sub-scale	Mean	SD	Mean	SD	% Agreement
Demonstrates sensitivity toward					
students	7.375	.558	6.7	.709	87.5%
Participates with others in a					
collaborative manner	7.325	.616	6.68	.713	70%
Treats others with respect	7.333	.519	6.695	.915	89.362%

All items	7.366	.547	6.691	.910	85.431%
teaching	7.75	.463	7.1	.738	87.5%
learners Displays excitement about	7.319	.471	6.483	1.127	76.596%
making skills Works effectively with diverse	7.354	.483	6.763	.935	89.583%
development Demonstrates effective decision-	7.188	.655	6.75	1.164	68.75%
Demonstrates lifelong learning Participates in professional	7.5	.599	6.76	.938	87.5%

References

- Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research*, 5(4), 194-197.
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology, 28*, 563-575.
- Wilson, R. F., Pan, W., Schumsky, D. A. (2012). Recalculation of the critical values for Lawshe's content validity ratio. *Measurement and Evaluation in Counseling and Development*, 45(3), 1-14. <u>https://doi.org/10.1177/0748175612440286</u>