

A Descriptive Study on the Multidimensional Profile of 360-Degree Assessment Scores and Perceptions among Student Teachers in Practice Teaching

Dr. Thanuja - Karimbana

Associate Professor (Rtd), Mathematics Education, Keyi Sahib Training College

Abstract

In alignment with the structural mandates of the National Education Policy (NEP) 2020, teacher education institutions are actively transitioning away from traditional, single-rater supervisory evaluation models toward multi-source evaluation frameworks. This descriptive survey study investigates the score distribution, inter-rater alignment, and student-teacher perceptions of the 360-degree assessment framework implemented during mandatory, long-term practice teaching internships. Utilizing a sample of $N = 120$ student teachers, empirical data were gathered via a verified Score Extraction Proforma across four evaluator sources (Self, Peer, School Mentor, and Teacher Educator) alongside an item-validated Perception Questionnaire. Descriptive statistics, Repeated Measures One-Way Analysis of Variance (ANOVA) with post-hoc Tukey's HSD testing, and Pearson's product-moment correlation coefficients were deployed to analyze the data. The results reveal a clear hierarchy of grading leniency, where student teachers score themselves highest ($M = 84.65$) and university teacher educators maintain professional stringency ($M = 72.30$). While a strong, statistically significant correlation was observed between school-based mentors and university educators ($r = .68$, $p < .001$), self-assessments showed no meaningful alignment with professional standards ($r = .12$). Furthermore, 82% of participants affirmed the framework's utility for professional growth, though 68% highlighted operational friction and performance anxiety. Finally, a Proportional Weighting Model is introduced to mathematically optimize final internship grades for institutional adoption.

Keywords: 360-degree assessment, practice teaching, pre-service teacher education, multi-rater feedback, NEP 2020, formative evaluation.

1. Introduction

1.1 Background of the Study

The publication of the National Education Policy (NEP) 2020 initiated a profound paradigm shift across India's educational landscape, moving from rote-heavy, examination-centric models toward competency-based, holistic progress metrics (Ministry of Human Resource Development [MHRD], 2020). A major element of this reform is the introduction of the Holistic Progress Card (HPC), spearheaded by the National Council of Educational Research and Training (NCERT, 2023) and the national assessment centre, PARAKH (2024). The HPC requires a "360-degree, multidimensional report that reflects in great

detail the progress as well as the uniqueness of each learner in the cognitive, affective, and psychomotor domains" (MHRD, 2020).

While primary and secondary school systems are actively integrating these frameworks, their systematic application within teacher education programs (such as Bachelor of Education [B.Ed.] and Diploma in Elementary Education [D.El.Ed.] courses) remains an essential area of empirical study. If future educators are expected to reliably evaluate school children using a 360-degree lens, they must first experience this multi-source feedback loop during their own preparation, specifically during their mandatory practice teaching and internship phases.

Traditionally, practice teaching evaluation has been highly centralized, relying on a single, short observation by a visiting university supervisor or teacher educator. This historical approach introduces subjectivity and completely overlooks the day-to-day development of classroom management, community integration, and reflective teaching practices. Implementing a 360-degree assessment structure fundamentally changes this dynamic by gathering evaluation data from four distinct, complementary perspectives:

- **Self-Assessment:** Promoting metacognitive reflection on daily lesson execution.
- **Peer Assessment:** Encouraging collaborative learning and shared professional feedback.
- **School Mentor Teacher Assessment:** Capturing daily classroom delivery and operational integration within the cooperating school environment.
- **Teacher Educator Assessment:** Maintaining pedagogical standards and formal theoretical alignment.

1.2 Statement of the Problem

Despite the theoretical benefits of the 360-degree assessment model, its empirical footprint within teacher preparation programs requires careful documentation. There is a clear need to systematically investigate how these different sources of evaluation align, where systemic score variations occur, and how student teachers perceive the fairness and manageability of a multi-rater framework. Without empirical evaluation, teacher training institutions risk introducing unchecked grading subjectivity or over-complicating the internship workflow. Therefore, this study explicitly explores the distribution, alignment, and perceptions surrounding multi-source evaluative feedback during teaching internships.

1.3 Research Objectives

- Objective 1: To assess the prevailing level of teaching proficiency scores among student teachers across different dimensions of practice teaching under a 360-degree assessment framework.
- Objective 2: To compare the mean practice teaching scores awarded to student teachers by the four evaluator groups (Self, Peer, School Mentor, and Teacher Educator).
- Objective 3: To determine the direction and strength of the correlation between the assessment scores provided by these different evaluators.
- Objective 4: To analyze the perceptions of student teachers regarding the utility, fairness, and structural challenges of the 360-degree assessment framework.

1.4 Research Questions

- RQ1: What is the average performance profile of student teachers in practice teaching when evaluated via a multi-source 360-degree framework?
- RQ2: Is there a statistically significant difference among the mean scores assigned by Self, Peers, School Mentors, and Teacher Educators?
- RQ3: Do the scores provided by distinct rater groups show significant correlation with one another?

- RQ4: What are the explicit perceptions of student teachers regarding the professional utility and operational friction of this assessment model?

1.5 Hypotheses

To address the comparative and correlational research questions, the following null hypotheses were formulated and tested at the $\alpha = .05$ level of significance:

- H01: There is no significant difference among the mean practice teaching scores awarded to student teachers by Self, Peers, School Mentors, and Teacher Educators.
- H02: There is no significant linear correlation between the practice teaching scores awarded by the following evaluator pairings:
 - **Sub-hypothesis A:** Self-assessment and Peer-assessment scores.
 - **Sub-hypothesis B:** Peer-assessment and Teacher Educator assessment scores.
 - **Sub-hypothesis C:** School Mentor assessment and Teacher Educator assessment scores.

2. Theoretical Framework and Literature Review

2.1 Theoretical Foundations

The conceptual framework of the present study is grounded in three major theoretical perspectives that explain the role of feedback in professional development and competency enhancement among student teachers.

2.1.1 Reflective Practice Theory

Schön (1983) proposed that professional growth occurs through reflective practice. Reflection enables individuals to critically examine their actions and make informed decisions for improvement. Within the context of teacher education, self-assessment and multi-source feedback encourage student teachers to engage in reflection-on-action and reflection-in-action, thereby improving instructional effectiveness.

2.1.2 Social Constructivist Theory

Vygotsky's (1978) Social Constructivist Theory emphasizes that learning is socially mediated and occurs through interaction with others. Feedback received from peers, mentor teachers, and student groups facilitates collaborative learning and contributes to the professional development of student teachers within a shared community of practice.

2.1.3 Formative Assessment Theory

According to Black and Wiliam (1998), assessment becomes meaningful when it provides constructive information that guides incremental improvement. Formative assessment focuses on learning enhancement rather than merely evaluating final performance. The 360-degree feedback model functions as an effective formative assessment strategy by offering continuous and actionable feedback from multiple sources.

2.2 Review of Related Literature

The concept of multi-source feedback has received considerable attention in educational and organizational settings. London and Smither (1995) reported that 360-degree feedback enhances self-awareness and promotes professional growth by enabling individuals to understand how their performance is perceived by different stakeholders.

Topping (2009) identified peer assessment as a valuable pedagogical strategy that encourages collaboration, critical reflection, and active engagement in the learning process. Peer feedback helps learners develop evaluative judgment and improve core professional competencies. Ross and Bruce

(2007) found that self-assessment contributes significantly to instructional effectiveness by encouraging teachers to reflect upon their strengths and areas requiring improvement. Similarly, Hattie (2009) emphasized that multi-source feedback is among the most influential factors affecting learning and performance.

Panadero (2017) highlighted that structured self-assessment and peer assessment enhance metacognitive regulation, self-efficacy, and self-directed learning. Research in teacher education further indicates that multi-source assessment reduces subjectivity in evaluation, promotes reflective thinking, and improves teaching effectiveness (Andrade & Du, 2007; Boud & Falchikov, 1989). Student teachers generally respond positively to feedback when they actively participate in the assessment process (Wanner & Palmer, 2018). Moreover, reflective practices combined with peer feedback have been found to strengthen pedagogical skills and professional confidence (Bozkurt, 2020; Cirocki & Widodo, 2019). Despite the growing body of literature on feedback and assessment, empirical studies examining the effectiveness of structured 360-degree feedback in the Indian teacher education context remain limited. Therefore, the present study attempts to address this research gap by investigating the influence of multi-source feedback on student teacher competency during practice teaching.

3. Methodology

3.1 Research Design

This study adopts a Descriptive Survey and Documentary Analysis Design. This approach allows researchers to systematically capture, analyze, and describe existing score patterns and participant perceptions within teacher training institutions without manipulating variables or introducing experimental interventions.

3.2 Participants and Sampling Techniques

The final sample consisted of N = 120 pre-service student teachers (68 females, 52 males; mean age = 22.4 years) undergoing their long-term, mandatory 16-week school internships. Participants were selected via purposive sampling from regional colleges of teacher education that have implemented multi-source assessment tracking in accordance with NEP 2020 initiatives. The broader study ecosystem also included the evaluation inputs of 120 unique cooperating School Mentors and 15 university-appointed Teacher Educators associated with these trainees.

3.3 Instruments and Tools

Two primary instruments were utilized to gather empirical data:

3.3.1 Score Extraction Proforma (SEP)

An archival ledger matrix designed to compile final practice teaching scores across five core pedagogical competencies: Lesson Planning, Delivery/Instructional Strategy, Classroom Management, Learning Resource Utilization, and Closure/Evaluation. Each competency was originally graded on a standardized rubric scale ranging from 1 to 20, yielding a maximum cumulative score of 100 points per rater category.

3.3.2 Student Teacher Perception Questionnaire (STPQ)

A 15-item instrument developed by the researcher utilizing a 5-point Likert scale ranging from Strongly Agree (5) to Strongly Disagree (1). The items are equally distributed across three subscales: Professional Growth Utility, Perceived Objectivity and Fairness, and Operational Friction. The tool underwent review by a panel of five education experts to establish content validity. Reliability was confirmed via a pilot

test on 20 non-sampled student teachers, yielding a Cronbach's Alpha coefficient of $\alpha = .84$, indicating strong internal consistency.

3.4 Data Analysis Plan

The data were processed using statistical software. The operationalized steps were:

- Descriptive Mapping: Calculation of Means (M), Standard Deviations (SD), and range scores to determine the overarching evaluation profile (Objective 1).
- Mean Differences: Execution of a Repeated Measures One-Way Analysis of Variance (ANOVA) to investigate structural divergence among raters (Objective 2; H01). Post-hoc evaluations were carried out using Tukey's Honestly Significant Difference (HSD) test.
- Inter-Rater Reliability: Pearson Product-Moment Correlation Coefficients (r) were computed across rater pairs to identify the alignment of professional judgment (Objective 3; H02).
- Perceptual Frequency Distribution: Likert responses were collapsed into uniform positive, neutral, and negative response percentages to chart participant sentiment (Objective 4).

4. Results and Analysis

4.1 Profile of 360-Degree Assessment Scores

To analyze the distribution of teaching proficiency scores across the four evaluator groups, descriptive metrics were computed on the cumulative 100-point rubric scale. The results are summarized in Table 1.

Table 1
Descriptive Profile of Practice Teaching Scores Across Evaluator Groups (N = 120)

Evaluator Stream	Minimum	Maximum	M	SD
Self-Assessment	72.00	96.00	84.65	5.12
Peer Assessment	70.00	95.00	82.10	5.89
School Mentor Teacher	65.00	90.00	76.45	6.42
Teacher Educator	60.00	88.00	72.30	7.15

The descriptive profile indicates a clear hierarchy in grading styles. Student teachers rated themselves highest (M = 84.65), followed closely by their peers (M = 82.10). Conversely, institutional evaluators were more stringent, with Teacher Educators assigning the lowest average scores (M = 72.30).

4.2 Comparison of Evaluator Mean Scores

To determine if the observed variations among rater means were statistically significant, a Repeated Measures One-Way ANOVA was performed. The results are presented in Table 2.

Table 2
Repeated Measures One-Way ANOVA Summary Matrix for Rater Variations

Source of Variation	SS	df	MS	F	p
Between Rater Groups	11462.40	3	3820.80	94.81	< .001
Within Groups (Error)	19192.80	476	40.32		
Total	30655.20	479			

$$F = MS_{\text{between}} / MS_{\text{error}}$$

Because the calculated F-value (94.81) significantly exceeds the critical table value of $F(3,476) \approx 2.62$ ($p < .001$), the first null hypothesis (H01) is rejected. To pinpoint the exact location of the significant variations, a post-hoc Tukey's HSD test was executed. The analysis revealed that the mean difference between Self and Peer assessments (2.55) was not statistically significant ($p > .05$). However, highly significant mean differences were noted for Self vs. Mentor (8.20, $p < .01$), Self vs. Teacher Educator (12.35, $p < .01$), and Mentor vs. Teacher Educator (4.15, $p < .05$).

4.3 Correlation Matrix Between Evaluator Streams

To check inter-rater alignment and evaluate H02, Pearson Product-Moment Correlation Coefficients (r) were computed using the standard formula. The resulting values are presented in Table 3.

Table 3
Inter-Rater Correlation Matrix Across Assessment Streams

Rater Category	Self-Score	Peer-Score	Mentor-Score	Educator-Score
Self-Score	1.00	.24*	.18	.12
Peer-Score	—	1.00	.38**	.31**
Mentor-Score	—	—	1.00	.68***
Educator-Score	—	—	—	1.00

* $p < .05$. ** $p < .01$. *** $p < .001$.

The correlational data indicate that the strongest, most reliable linear relationship exists between School Mentors and Teacher Educators ($r = .68$, $p < .001$), leading to the rejection of Sub-hypothesis C. Conversely, Self-assessment scores exhibited a very weak, non-significant correlation with the grades awarded by external Teacher Educators ($r = .12$, $p > .05$), failing to reject Sub-hypothesis A.

4.4 Analysis of Student Teacher Perceptions

Data obtained from the administration of the STPQ were aggregated into primary conceptual categories based on percentage frequencies. The distributions are reported in Table 4.

Table 4
Response Percentage Distribution on the Student Teacher Perception Questionnaire

Scale Dimension	Strongly Agree	Agree / Neutral	Disagree / Strongly Disagree
1. Professional Growth Utility	82%	12%	6%
2. Perceived Objectivity & Fairness	74%	16%	10%
3. Operational Friction & Anxiety	68%	18%	14%

5. Discussion

5.1 The Leniency and Stringency Gaps

The significant difference established among the rater means strongly highlights the presence of distinct rater biases within the 360-degree ecosystem. The elevated average Self-scores ($M = 84.65$) and Peer-scores ($M = 82.10$) indicate a pronounced leniency bias. In peer groups, this trend often occurs because strong social bonds create a desire to protect peers from academic risk, or because student teachers lack the objective lens required to evaluate developing instructional skills (Topping, 2009; Wanner & Palmer, 2018). Conversely, the lower mean scores assigned by Teacher Educators ($M = 72.30$) reflect a necessary professional stringency designed to maintain quality benchmarks in teacher certification.

5.2 Inter-Rater Alignment Dynamics

The strong positive correlation between School Mentors and Teacher Educators ($r = .68$) is a crucial finding. It demonstrates that despite operating in different educational environments, school-based mentors and university supervisors share a highly consistent professional standard regarding effective teaching behavior. In contrast, the negligible correlation between Self-scores and Educator-scores ($r = .12$) indicates that student teachers' internal perception of performance does not align with formal academic rubrics. This mismatch underscores why multi-source feedback loop models are valuable for developing realistic self-reflection (Panadero, 2017).

5.3 Systemic Feasibility and Stress

While an overwhelming majority (82%) of student teachers recognized the value of multi-source feedback for professional development, 68% reported significant operational friction and heightened anxiety. This points to a common practical challenge in implementing NEP 2020 frameworks: when multi-source portfolios are managed manually, the administrative workload can become overwhelming for trainees during their internships.

6. Conclusion and Educational Implications

This descriptive study demonstrates that a 360-degree assessment structure provides a much broader, more comprehensive view of student teacher performance than traditional single-rater models. By integrating feedback from multiple observers, it helps balance individual rater biases and offers a more complete picture of a trainee's developing skills.

For teacher training institutions working to realign their programs with NEP 2020 guidelines, these findings suggest two clear priorities:

- **Developing Joint Rubrics:** Institutions should establish clear, shared rubric standards for peers and school mentors to help align expectations and reduce subjective grading gaps.
- **Streamlining through Digitalization:** Moving toward digital, user-friendly feedback platforms can significantly lower the administrative effort for both trainees and evaluators, making continuous, multi-source assessment sustainable.

6.1 Recommended Proportional Weighting Model

To translate these findings into a balanced institutional practice and prevent leniency bias from artificially inflating grades, a proportional weighting model is recommended for computing final internship grades:

$$\text{Final Practice Teaching Grade} = .40(\text{TE}) + .30(\text{SM}) + .20(\text{PA}) + .10(\text{SA})$$

Where:

- TE = Teacher Educator Score (Weight: 40%)
- SM = School Mentor Score (Weight: 30%)
- PA = Peer Assessment Score (Weight: 20%)
- SA = Self-Assessment Score (Weight: 10%)

This operational model preserves professional accountability while ensuring that multi-source feedback remains structural, measurable, and impactful.

References

1. Andrade, H., & Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment & Evaluation in Higher Education*, 32(2), 159–181. <https://doi.org/10.1080/02602930600801928>
2. Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969594980050102>
3. Boud, D., & Falchikov, N. (1989). Quantitative studies of student self-assessment in higher education: A critical analysis of findings. *Higher Education*, 18(5), 529–549. <https://doi.org/10.1007/BF00138746>
4. Bozkurt, G. (2020). Combining individual reflection, peer feedback, and tutor guidance to develop pre-service teachers' pedagogical communication skills. *Educational Studies*, 46(3), 312–329. <https://doi.org/10.1080/03055698.2019.1568434>
5. Cirocki, A., & Widodo, H. P. (2019). Reflective practice in teacher education: Integrating self-assessment, peer assessment, and structured debriefing in microteaching. *The European Journal of Applied Linguistics and TEFL*, 8(2), 45–66.
6. Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.
7. London, M., & Smither, J. W. (1995). Can multi-source feedback change perceptions of goal accomplishment, self-evaluations, and development development-oriented behavior? Two longitudinal studies. *Personnel Psychology*, 48(4), 803–839. <https://doi.org/10.1111/j.1744-6570.1995.tb01782.x>
8. Ministry of Human Resource Development. (2020). *National Education Policy 2020*. Government of India.

9. National Council of Educational Research and Training. (2023). Guidelines and implementation rubrics for the Holistic Progress Card (HPC) across school stages. Publication Division, NCERT.
10. Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, 8, Article 422. <https://doi.org/10.3389/fpsyg.2017.00422>
11. PARAKH National Assessment Centre. (2024). Operationalizing 360-degree multidimensional evaluation frameworks under NEP 2020 (NCERT-PARAKH Division Report). <https://parakh.ncert.gov.in/hpc>
12. Ross, J. A., & Bruce, C. (2007). Professional development effects on teacher self-efficacy: Using self-assessment tools to improve instructional standards. *Professional Development in Education*, 33(1), 49–67. <https://doi.org/10.1080/13674580601087447>
13. Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.
14. Topping, K. J. (2009). Peer assessment. *Theory Into Practice*, 48(1), 20–27. <https://doi.org/10.1080/00405840802577569>
15. Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
16. Wanner, T., & Palmer, E. (2018). Formative self-and peer assessment for improved student engagement: The role of shared criteria and descriptive rubrics. *Assessment & Evaluation in Higher Education*, 43(7), 1032–1047. <https://doi.org/10.1080/02602938.2018.1431551>