ACADEMIA
ISSN, 2241-1402
http://hepnet.upatras.gr

Number 41, 2025

Academia A publication of the Higher Education Policy Network

Beyond Student Outcomes: Reassessing Teaching Effectiveness through Student Evaluation

Animesh Karn[1], Pallavi Kumari[2]

Amity University, The ICFAI University Jharkhand

Abstract

This study reexamines teaching effectiveness by moving beyond traditional student outcome measures to incorporate student perceptions, offering a more comprehensive evaluation framework. Drawing from cognitive, developmental, and relational theories, we address critical gaps in existing literature that often overlook qualitative aspects of instructor-student interactions. Using a simplified 14-item questionnaire, we collected data from 319 undergraduate management students at two private universities in India. Exploratory factor analysis revealed three key dimensions of effective teaching as perceived by students: interpersonal skills, confidence in instructor ability, and confidence in guidance provided. These dimensions not only align with but also expand upon traditional measures of teacher knowledge, experience, and behavior, emphasizing the significant role of mentorship and interpersonal relationships in teaching effectiveness. The findings suggest practical implications for enhancing instructor training programs through targeted development of interpersonal competencies, pedagogical confidence, and mentorship strategies. This study advances the literature by providing a validated, student-centered framework for assessing teaching effectiveness, offering educational institutions a robust tool to promote holistic student development.

Keywords

Teaching effectiveness, student perceptions, interpersonal skills, guidance, mentorship, higher education

[1] Research Scholar, The ICFAI University Jharkhand and Assistant Professor, Amity University, Jharkhand, akarn@rnc.amity.edu

[2] Associate Professor, The ICFAI University Jharkhand, pallavikumari@iujharkhand.edu.in

Introduction

Research into effective teaching or characteristics of effective teachers has predominantly used student outcomes as the primary measure of effectiveness. This emphasis stems from the belief that every resource within an educational institution should contribute to the academic and overall success of students (Rockoff, 2004; Rivkin et al., 2005; Chetty et al., 2014). Among these resources, teachers are considered the most crucial. Consequently, enhancing teaching effectiveness has become a central focus of educational reforms, particularly in higher education. A growing body of research has linked teacher effectiveness to improvements in student learning (Goe, 2007) and in student performance percentiles (Chetty et al., 2014), as measured by standardized assessments.

However, the existing literature presents mixed results, with no consensus on the validity and impact of instructor characteristics on student outcomes (Wayne & Youngs, 2003). This ambiguity is further compounded by the contextual limitations of existing research. Much of this work originates from the United States, raising questions about its applicability to other educational systems. Moreover, a comprehensive cross-country study by Burroughs et al. (2019), using data from the International Association for the Evaluation of Educational Achievement (IEA), found significant variation in the relationship between instructor characteristics and student outcomes. This inconsistency highlights a critical gap in the literature: a lack of a universally applicable model of teaching effectiveness.

To address this gap, we propose a framework suggesting that effective teaching extends beyond the mere transmission of knowledge and involves a complex interplay of guidance, mentorship, inspiration, and personal development. We argue that objective measures of evaluation capture only a fraction of what constitutes effective teaching and propose a more comprehensive evaluation based on students' subjective perceptions of cognitive, developmental, and relational metrics.

To validate this framework, we examined the dimensions of effective teaching identified in the existing literature—knowledge, experience, and behavior—through students' perceptions of their instructors. Using exploratory factor analysis on a set of 14 statements based on our proposed metrics and rated by undergraduate students on a 5-point Likert scale, we found that students' perceptions of their professors can be distilled

into three underlying dimensions: confidence in their ability, confidence in the guidance they provide, and their interpersonal skills.

This study makes three significant contributions to the literature on teaching effectiveness. *First*, we develop a comprehensive framework that underscores the multidimensional nature of teaching effectiveness, showing that it involves not only subject matter expertise but also strong interpersonal skills and the ability to provide guidance. *Second*, by using student ratings, we provide a bottom-up perspective that captures aspects of teaching that matter to students but may not be reflected in academic outcomes. *Third*, we demonstrate the utility of a simplified 14-item questionnaire as a reliable and practical tool for evaluating teaching effectiveness, offering a quick screening method to identify areas for improvement. Based on the results, targeted interventions can be designed to enhance specific dimensions of teaching effectiveness.

1. Literature Review

1.1 Dimensions of Effective Teaching

The literature identifies three main aspects of effectiveness: (i) experience, (ii) knowledge, and (iii) behavior. In this context, experience is measured by the number of years spent working as a teacher. Knowledge implies expertise in the subject-matter, understanding of the curriculum, and mastery of educational pedagogy. Behavior, defined very narrowly in the existing literature, pertains mostly to the instructional content and the duration of student exposure to it.

1.1.1 Experience

Experience has been found to have a cumulative positive effect on student outcomes (Wiswall, 2013; Papay & Kraft, 2015; Ladd & Sorenson, 2017) but pronounced mostly at secondary school levels (Rice, 2003; Wayne & Youngs, 2003; Clotfelter et al., 2006) or during the first five years of an academic career (Rice, 2003; Rockoff, 2004; Rivkin et al., 2005; Boyd et al., 2006; Pil & Leana, 2009; Staiger & Rockoff, 2010; Papay & Kraft, 2015). But many other studies have not found any statistically significant relationship between the two (Hanushek & Luque, 2003; Wilson & Floden, 2003; Luschei & Chudgar, 2011; Gustaffsson & Nilson, 2016; Blomeke et al., 2016).

1.1.2 Knowledge

By and large, the literature is inconclusive regarding the relationship between instructors' professional knowledge and student outcomes. Many studies (Wayne & Youngs, 2003; Rice, 2003; Darling-Hammond et al., 2005; Akiba et al., 2007; Goe, 2007; Palardy & Rumberger, 2008; Blank & De Las Alas, 2009; Boyd et al., 2009; Phillips, 2010; Montt, 2011; Gustaffsson & Nilson, 2016) and one multi-country study (Woessman, 2003) found academic degrees, area specializations, and teaching certifications to be positively correlated with student outcomes. On the other hand, several studies (Wilson & Floden, 2003; Hanushek & Luque, 2003; Clotfelter et al., 2006; Aaronson et al., 2007; Wallace, 2009; Rockoff et al., 2011; Luschei & Chudgar, 2011; Harris & Sass, 2011; Blazar, 2015; Blomeke et al., 2016; Garet et al., 2016; Schmidt et al., 2017) have found a weak or non-significant association between instructors' subject-matter expertise and student academic achievements.

1.1.3 Behavior

Studies on instructor behavior mainly focus on the time spent on teaching. The comprehensive IEA study (Burroughs et al., 2019) based on the 'Trends in International Mathematics and Science Study (TIMSS)' data primarily evaluates teachers' preparedness, time spent on teaching, and curricular alignment to assess behavioral aspects of teaching effectiveness. Schmidt et al., (2001; 2015) found a strong association between time spent on teaching or the opportunity to learn (OTL) and students' academic achievement across multiple educational systems. This was supported by several studies based on student reported data (Woessman, 2003; Rivkin & Schiman, 2015; Lavy, 2015; Cattaneo et al., 2016; Jerrim et al., 2017). However, Cohen (1981) cautions that higher student achievement may be due to lenient gradings rather than the degree of learning perceived by the students. Greenwald & Gillmore (1997) found that students rated teaching effectiveness higher in courses where they expected lenient grading. These findings were contested by McKeachie (1997), d'Apollonia & Abrami (1997), and Patrick (2011) who concluded that there is a weak relationship between grading leniency and effectiveness ratings. They argued that the positive correlations between the two represent only a minimal relationship and do not provide strong evidence.

1.2 Challenges with Quantitative Feedback Systems

Existing research on teaching effectiveness predominantly relies on student outcomes and quantitative measures of instructor characteristics, often overlooking the qualitative aspects of the instructor-student relationship. Several studies have shown that student evaluations based on quantitative questionnaires may contain indicators irrelevant to effective teaching or may not fully capture the nuances of the instructor-student dynamic (Schmidt et al., 2001; Burroughs et al., 2019). Furthermore, a 2007-08 OECD report found that 'relations between teachers and students' constituted only one of twelve criteria considered of 'high or moderate importance' in school self-evaluations (TALIS Board of Participating Countries, 2009), further suggesting a limited emphasis on this critical aspect.

Existing questionnaires, even those designed to be granular, often fail to adequately assess the instructors' role as a guide and mentor. Traditional evaluations primarily focus on assessing teachers' capabilities in-class instruction, neglecting the more nuanced aspects of mentorship and guidance. Although such feedback can help teachers refine their instructional strategies, it provides limited insight into their effectiveness in fostering student development beyond the classroom.

Furthermore, the literature presents conflicting perspectives on the role of student characteristics in influencing evaluations. Some studies suggest that student evaluations may be biased by factors such as academic abilities, motivation levels, gender, age, and class size (Gigliotti & Buchtel, 1990; Cashin, 1995). Others argue that students possess "metacognition" about effective teaching, minimizing the impact of these biases (Harrison et al., 1996). This ongoing debate underscores the need for a more nuanced approach.

1.3 Summary of the Literature Survey

The evidence for two of the three dimensions of effective teaching identified in the extant literature—experience and knowledge—is inconclusive. The third dimension, behavior, has been interpreted narrowly, and even for this, the evidence is mixed. Additionally, most studies have relied solely on objective measures such as certificates and degrees for knowledge, years of experience, and time spent on teaching for behavior, as indicators of teaching effectiveness. These measures lack any subjective evaluation by students.

Therefore, a more comprehensive understanding of teaching effectiveness requires incorporating qualitative feedback that directly addresses the non-instructional dimensions of teaching. This includes exploring students' perceptions of instructor knowledge and experience, as well as broadening the definition of instructor behavior to encompass mentorship and guidance aspects. Teaching effectiveness evaluations should thus (a) use subjective measures of instructor knowledge and experience and (b) incorporate aspects of instructor behavior in a more substantial, fuller sense, while controlling for student characteristics, such as gender and academic ability.

2. Theoretical Framework

To comprehensively assess teaching effectiveness, we must move beyond traditional, narrow metrics which overlook the qualitative aspects of instructor-student interactions. We draw on theories from educational economics, developmental psychology, and managerial leadership to develop an integrative framework that captures the complexity of teaching.

2.1 Cognitive Dimension

The Human Capital Theory (Becker, 1964) views education as an investment in both cognitive and non-cognitive skills. The Constructivist Learning Theory (Piaget, 1972; Vygotsky, 1978) emphasizes that students build knowledge through active engagement and interaction with their instructors and peers. Instructors play a crucial role as facilitators, fostering environments where knowledge is co-constructed through dialogue and guided enquiry.

2.2 Developmental Dimension

Students' personal and psychological development is as critical as advancing their academic growth. Self-Determination Theory (Deci & Ryan, 2000) underscores the importance of psychological needs–autonomy, competence, and relatedness–in promoting sustained engagement and motivation. Instructors should create learning spaces where students develop the confidence necessary to pursue both academic and personal goals.

2.3 Relational Dimension

Principles from Transformational Leadership Theory (Burns, 1978; Bass, 1985) can be used to reframe instructors as leaders who inspire and motivate students beyond mere academic achievement. Social Cognitive Theory (Bandura, 1986) recognizes instructors as role models whose behaviors, attitudes, and expectations significantly influence students' cognitive engagement and learning outcomes. Transformational instructors nurture empowering environments where students are valued as individuals by providing individualized attention, mentorship, and guidance.

2.4 The Integrated Framework

These theories collectively present teaching effectiveness as a fluid, adaptive, and multidimensions process where cognitive, developmental, and relational aspects continuously interact, each informing and reinforcing the other. Its assessment should capture these dynamic interactions:

1. Cognitive Measures: Assessing the quality and depth of knowledge transfer.
2. Developmental Indicators: Evaluating how instructors nurture student motivation and personal growth.
3. Relational Metrics: Assessing how well instructors mentor and provide guidance to students.

This approach ensures a holistic assessment of teaching effectiveness, balancing both measurable academic outcomes and the less tangible but equally vital aspects of personal development and mentorship. By integrating these dimensions, educational institutions can enhance their teacher evaluations, gaining a deeper understanding of how teaching influences student success. This can potentially lead to improved instructional practices and enhanced student outcomes.

3. Study Aims, Objectives, and Research Questions

This study aims to explore the dimensions of effective teaching based on the proposed integrated framework. Specifically, the study aims to validate the three core aspects of effective teaching identified in the literature–knowledge, experience, and behavior–through students' subjective evaluations of the three core dimensions–cognitive, developmental, and relational–proposed in the framework. The following objectives guide the study:

1. To identify the key factors underlying students' perceptions of teaching effectiveness.
2. To examine the characteristics of these factors and the items that contribute to them.
3. To interpret the identified factors in relation to the cognitive, developmental, and relational dimensions of the proposed theoretical framework.

To this end, the research explores three primary questions:

1. What are the key factors that underlie students' perceptions of teaching effectiveness?
2. What are the characteristics of each identified factor, and which specific items contribute most strongly to each factor?
3. How can the identified factors be interpreted within the context of the cognitive, developmental, and relational dimensions of the proposed theoretical framework?

To address these questions, the study applies exploratory factor analysis to survey data collected as responses to statements describing various aspects of instructor-student interactions from undergraduate students.

4. Methods

To address these research questions, we propose a simplified questionnaire as a precursor to a more detailed one that will contain questions only about the specific focus areas indicated by the simplified feedback. This approach allows us to identify broad areas for improvement in teaching effectiveness without overwhelming the instructor with a sea of performance metrics, which can obscure the guidance and mentorship issues in plain sight.

We use a set of 14 statements (Table–1) covering various aspects of teaching, from study plans to personal problems, based on the metrics proposed in the theoretical framework. These statements are assessed on a 5-point Likert scale, with half worded positively and half negatively, but in such a way that '1' always represents the most positive attitude and '5', the least positive. This format checks for internal consistency of the ratings by asking contextually contiguous questions worded differently. The statements in the questionnaire are designed to capture the three core dimensions of our framework:

1. Cognitive Aspects: statements addressing the clarity of instruction, depth of subject knowledge, and ability to explain complex concepts.
2. Developmental Indicators: statements evaluating the instructor's ability to motivate students, nurture personal growth, and build confidence.
3. Relational Metrics: statements assessing the quality of instructor-student interactions, mentorship, and guidance.

This schedule is a modified version of the one used by Alemayehu, Bushen, and Muluneh (2009) to evaluate clinical care quality.

Table 1: *Management Students' Evaluation of Professors*

Sl. No.	Statement	Dimension
1	My teacher treats me in a friendly manner.	Relational
2	I have some doubts about the ability of my teacher.	Cognitive
3	My teacher seems cold and impersonal.	Relational
4	My teacher does his/her best to keep me from worrying.	Relational
5	My teacher assesses me as carefully as necessary.	Developmental
6	My teacher should treat me with more respect.	Developmental
7	I have some doubts about the study plan suggested by my teacher.	Developmental
8	My teacher seems very competent and well trained.	Cognitive
9	My teacher seems to have a genuine interest in me as a person.	Relational
10	My teacher leaves me with many unanswered questions about the subject matter and its understanding.	Cognitive
11	My teacher uses words that I do not understand.	Cognitive
12	I have a great deal of confidence in my teacher.	Cognitive
13	I feel I can tell my teacher about very personal problems.	Relational
14	My teacher listens to me very patiently.	Relational

The responses were analyzed using Factor Analysis to uncover the 'factors' that an instructor should focus on to become more effective. Factor analysis is an errors-in-variables class of regression analysis, distinct from standard models that assume exact regressor measurement. Standard sample size calculation formulas do not apply in this context. Research into sample size requirements (Marsh & Hau, 1999; Lee & Song, 2004; Kline, 2005; Gignac, 2015; Kyriazos, 2018) suggests that numbers of cases anywhere

between 5 to 20 times the number of variables constitutes a sufficient to very good sample size. Since our questionnaire contains 14 items, a sample comprising anywhere between 70 to 280 students would suffice. In this study, we used a sample size of 320 students, which goes beyond this sufficiency range.

The survey was conducted with undergraduate students at the two largest private universities in Jharkhand, India. Private universities in India exhibit a distinct educational dynamic, characterized by greater autonomy and flexibility in curriculum design and faculty recruitment (Altbach, 2014). These institutions typically prioritize student satisfaction, making them ideal for evaluating teaching effectiveness from the lenses of interpersonal skills and guidance provided by instructors (Varghese, 2015). In contrast, public institutions operate under more rigid bureaucratic structures, which could limit responsiveness to student evaluations and dilute the focus on mentorship—a task left for future studies (Tilak, 2015).

Both the universities surveyed are in the state capital, Ranchi, and run programs from the undergraduate to doctoral levels in humanities, social sciences, management, law, and technology. A common introductory course (albeit with slightly different syllabuses) on principles of management taught by four different members of the faculty of the two universities was selected for the survey.

The survey, administered during scheduled lecture hours, included all 320 first-year students from eight sections of BBA–MBA and BBA–LLB integrated programs, each comprising 40 students. Program coordinators informed the students in advance. Participation was full; only one submitted an incomplete survey, resulting in 319 fully completed surveys.

5. Results and Analysis

Data was fed into JASP version 0.18.3 (JASP Team, 2024) using R version 4.1 (R Core Team, 2022) and R packages ‘psych’ (Revelle, 2023), ‘lavaan’ (Rosseel, 2019), ‘semTools’ (Jorgensen et al., 2019), and ‘semPlot’ (Epskamp et al., 2019).

The questionnaire had high degree of internal consistency and reliability as indicated by high values of Cronbach’s α and the more robust, McDonald’s ω (Table–2). The Guttman’s λs, especially the more conservative λ6, are also very high, reinforcing high reliability of the scale. Greatest Lower Bound (GLB) provides a sharp lower bound to the reliability of the scale implying highest possible reliability given inter-item

correlations; its value being close (0.971) to one indicates very high reliability. Additionally, the average interitem correlation (0.506) suggests a moderate level of correlation among the items on average.

Table 2: *Scale Reliability Statistics*

Estimate	McDonald's ω	Cronbach's α	Guttman's λ2	Guttman's λ6	Greatest Lower Bound	Average interitem correlation
Point estimate	0.932	0.930	0.933	0.953	0.971	0.506
95% CI lower bound	0.921	0.918	0.920	0.946	0.966	0.460
95% CI upper bound	0.943	0.941	0.944	0.963	0.979	0.549

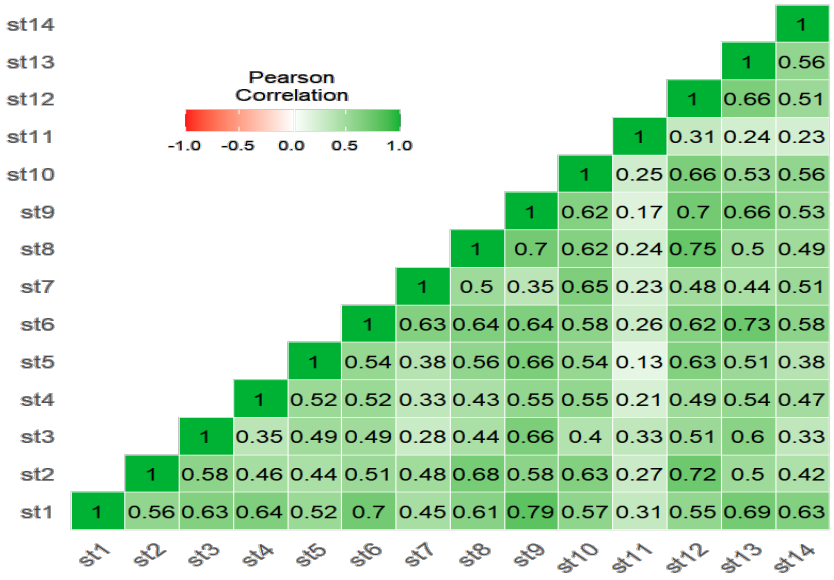
Item reliability statistics show (Table–3) that all items have good internal consistency as the α, ω, λ2, and λ6 as well as the GLB values are all above 0.9; dropping any of the items does not significantly impact the scale. Except for Statement 11, high item-rest correlation values (> 0.600) indicate that the individual items are consistent with all other items and that they relate well with the overall scale. Low item-rest correlation value of 0.320 for Statement 11 is an indication to the possibility that this item may not correlate with any of the factors extracted later.

Table 3: *Item Reliability Statistics*

	If item dropped					
Item	McDonald's ω	Cronbach's α	Guttman's λ2	Guttman's λ6	Greatest Lower Bound	Item-rest correlation
st1	0.925	0.922	0.925	0.945	0.967	0.804
st2	0.926	0.924	0.927	0.947	0.966	0.716
st3	0.929	0.927	0.930	0.948	0.965	0.628
st4	0.929	0.927	0.930	0.950	0.970	0.626
st5	0.928	0.926	0.929	0.950	0.967	0.650
st6	0.924	0.922	0.925	0.947	0.967	0.779
st7	0.930	0.928	0.931	0.950	0.968	0.595
st8	0.926	0.924	0.927	0.947	0.968	0.752
st9	0.924	0.922	0.925	0.944	0.965	0.790

Item	If item dropped					
	McDonald's ω	**Cronbach's α**	**Guttman's λ2**	**Guttman's λ6**	**Greatest Lower Bound**	**Item-rest correlation**
st10	0.925	0.923	0.925	0.947	0.967	0.754
st11	0.940	0.939	0.940	0.957	0.976	0.320
st12	0.923	0.921	0.924	0.945	0.964	0.799
st13	0.926	0.923	0.926	0.947	0.968	0.746
st14	0.928	0.926	0.929	0.952	0.972	0.645

Figure 1: *Correlation Heatmap*



Correlation heatmap of the items shows low to moderate correlation (Figure–1). None of the correlations are very high thus obviating the reason for dropping any item from the scale.

The overall Kaiser-Meyer-Olkin Measure of Sampling Adequacy is 0.881 (Table–4), which is considered “meritorious” according to Kaiser's criteria, indicating that the data is suitable for factor analysis.

Table 4: *KMO Measure of Sampling Adequacy*

	MSA
Overall MSA	0.881
Statement 1	0.880
Statement 2	0.858
Statement 3	0.819
Statement 4	0.887
Statement 5	0.908
Statement 6	0.916
Statement 7	0.841
Statement 8	0.905
Statement 9	0.864
Statement 10	0.903
Statement 11	0.720
Statement 12	0.871
Statement 13	0.893
Statement 14	0.955

To estimate the number of factors to be extracted, we use the Maximum Likelihood approach. And, then we use the principal factor analysis to estimate the loadings on the number of factors estimated through the maximum likelihood approach.

Maximum Likelihood factor analysis' goodness-of-fit test was used, successively, to assess the null hypothesis that a model with (a) one factor, (b) two factors, or (c) three factors explained the covariance of the input variables (see Table–5). The one- and two-factor models did not exhibit any significant fit to the model. For the 3-factor solution, the *p*-value is 0.058, which is not statistically significant at the 0.05 level. This suggests that the 3-factor model cannot be rejected as providing a good fit to the data.

Table 5: *Goodness-of-Fit Tests for 1-, 2-, and 3-factor Models*

Model	**Chi-Square**	***df***	***p***
1-factor	771	77	.000
2-factor	594	64	.002

3-factor	412	52	.058

But the other fit indices show a mixed result (Table–6). The RMSEA value of 0.147 is higher than the recommended cutoff of 0.08, suggesting a potentially poor fit of the model to the data. However, the SRMR value of 0.038 is below the recommended cutoff of 0.08, indicating a good fit based on this index. The TLI (0.924) and CFI (0.939) values are marginally below the recommended cutoff of 0.95, suggesting less than ideal, yet acceptable, fit of the model to the data.

Table 6: *Additional Fit Indices*

RMSEA	RMSEA 90% confidence	SRMR	TLI	CFI	BIC
0.147	0.134 — 0.161	0.038	0.924	0.939	111.736

On the balance, we consider the 3–factor model to be an adequate fit to the data. Next, we set 'principal axis factoring' as the extraction method and apply the varimax rotation to the factor loadings to obtain the 'Total Variance Explained' (Table–7).

Table 7: *Factor Characteristics–Eigenvalues and Total Variance Explained*

		Unrotated solution			Rotated solution		
	Eigenvalues	SumSq. Loadings	Proportion var.	Cumulative	SumSq. Loadings	Proportion var.	Cumulative
Factor 1	7.772	7.439	0.531	0.531	3.400	0.243	0.243
Factor 2	1.007	0.676	0.048	0.580	2.880	0.206	0.449
Factor 3	0.969	0.529	0.038	0.617	2.363	0.169	0.617

The unrotated solution shows that Factor 1 accounts for 53.1% of the variance, Factor 2 accounts for 4.8%, and Factor 3 accounts for 3.8%. The cumulative variance explained by the three factors is 61.7%. The rotated solution (using varimax rotation) shows a slightly different pattern, with Factor 1 accounting for 24.3% of the variance, Factor 2 accounting for 20.6%, and Factor 3 accounting for 16.9%. The cumulative variance explained remains at 61.7%. The results also validate the choice of a 3–factor model as the first three eigenvalues generated are each greater than 0.900 satisfying Kaiser's criterion of high eigenvalues. The number of factors to be retained is also

examined through the Parallel Analysis, one of the accurate methods for determining the number of factors (Table–8).

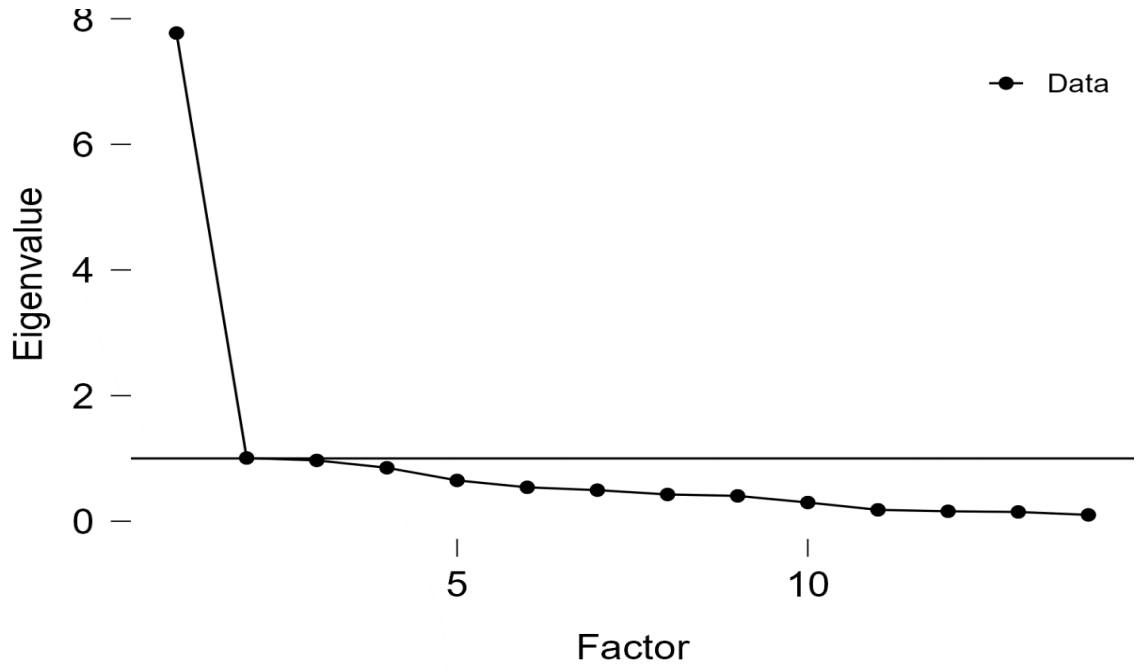
Table 8: *Parallel Analysis*

	Real data factor eigenvalues	**Simulated data mean eigenvalues**
Factor 1*	7.342	0.578
Factor 2*	0.501	0.297
Factor 3*	0.427	0.243
Factor 4	0.186	0.185
Factor 5	0.127	0.135
Factor 6	0.068	0.074
Factor 7	-0.006	0.041
Factor 8	-0.035	-0.009
Factor 9	-0.069	-0.046
Factor 10	-0.158	-0.096
Factor 11	-0.214	-0.139
Factor 12	-0.225	-0.179
Factor 13	-0.302	-0.221
Factor 14	-0.354	-0.286

Note. '*' = Factor should be retained. Results from FA-based parallel analysis.

The eigenvalues of only the first three factors exceed the corresponding eigenvalues from the random data suggesting appropriateness of the 3–factor model. This is visualized in the Scree Plot (Figure–2). The scree plot shows a clear elbow or leveling off after the third factor, suggesting that a three-factor solution may be appropriate for the data.

Figure 2: *Scree Plot*



The factor loadings (Table–9) show that items (statements) 1, 9, 13, 3, 6, 4, 5, and 14 load primarily on Factor 1, while items 12, 2, and 8 load on Factor 2, and items 7, 6, and 10 load on Factor 3.

Table 9: *Rotated Factor Loadings*

Statement	Factor 1	Factor 2	Factor 3	Uniqueness
1	0.810			0.154
9	0.708	0.543		0.178
13	0.658			0.345
3	0.594	0.423		0.465
6	0.560		0.562	0.284
4	0.512			0.559
5	0.448	0.486		0.513
14	0.448		0.529	0.479
12		0.757		0.190

Statement	Factor 1	Factor 2	Factor 3	Uniqueness
2		0.669		0.359
8		0.656		0.317
10		0.507	0.579	0.319
7			0.782	0.301
11				0.893

Note. Applied rotation method is varimax. Absolute values ≤ 0.4 are not presented in the table.

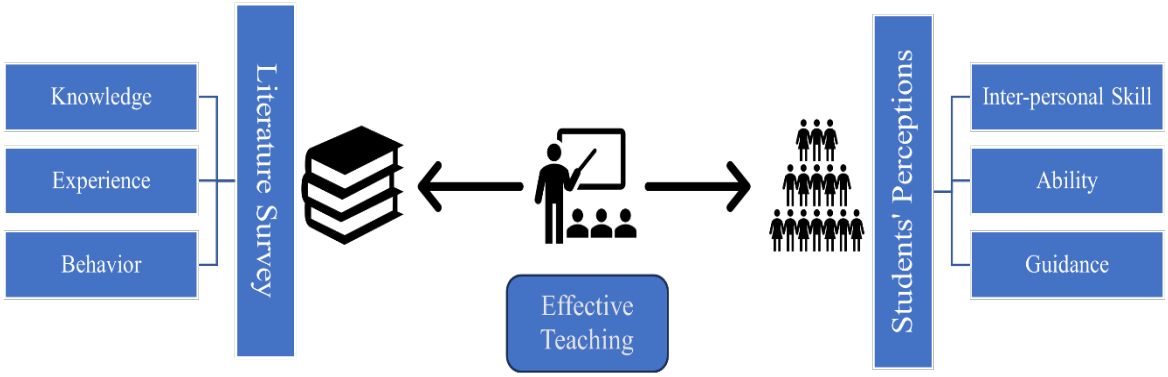
As expected from the item-rest correlation value calculated earlier, item (statement) 11 does not load substantially on any factor, suggesting it may not fit well with the rest of the items.

The rescaled rotated factor matrix, with absolute values less than 0.4 suppressed, is used for interpretation and naming the factors. The three largest loadings in **Factor 1** are 'treats me in a friendly manner', 'genuine interest in me', and 'tell… personal problems', it may be labelled as **'Evaluation of Teacher's Interpersonal Skills'**. Similarly, for **Factor 2**, the three largest loadings are 'confidence', 'doubts about the ability', and 'competent and well trained', it may be labelled as **'Confidence in Teacher's Ability'**. Again, for **Factor 3**, the two largest loadings are 'doubt about the study plan' and 'unanswered questions', it may be labelled as **'Confidence in Guidance Provided'**.

6. Discussion

This study aimed to validate the dimensions of effective teaching by analyzing students' perceptions of their professors using a simplified 14-item questionnaire rated on a 5-point Likert scale. Exploratory Factor Analysis of student ratings identified three underlying factors corresponding to previously established aspects of effective teaching: instructor knowledge, experience, and behavior. Our findings, however, offer a fresh perspective by evaluating these dimensions from the students' point of view, revealing significant alignments with and empirical support for the proposed integrated theoretical framework of teaching effectiveness (Figure–3).

Figure 3: *Dimensions of Effective Teaching—Literature Survey vs. Students' Perceptions*



The first factor, 'Evaluation of Teacher's Interpersonal Skills', captures the relational aspect of teaching effectiveness. High loadings on this factor pertain to professors' friendliness, interest in students, and approachability for discussing personal problems. This suggests that, from students' perspective, professors' interpersonal and mentoring skills are crucial aspects of effective teaching, extending beyond lecture hall instruction. These interpersonal skills foster a positive learning environment where students feel valued and supported, aligning with the mentorship and role modeling aspects emphasized by Transformational Leadership (Burns, 1978; Bass, 1985) and Social Cognitive (Bandura, 1986) Theories. Prior research has often narrowly defined instructor behavior in terms of time spent on teaching and curricular alignment (Schmidt et al., 2001; Burroughs et al., 2019). Our findings highlight the need to broaden this definition to include the guidance and mentorship role played by professors.

The second factor, 'Confidence in Teacher's Ability', represents the cognitive dimension of teaching effectiveness. High loadings on items related to students' confidence in professors' competence and training resonates with previous studies that found these to be positively associated with student outcomes (Wayne & Youngs, 2003; Darling-Hammond et al., 2005; Palardy & Rumberger, 2008). This aligns with the principles of Human Capital Theory (Becker, 1964) and Constructivist Learning Theory (Piaget, 1972; Vygotsky, 1978), which emphasize the instructors' role in facilitating knowledge transfer and construction. By directly measuring students' perceptions, our study shows how instructor knowledge builds student confidence in their professors. This confidence likely stems from the students' belief that their instructors possess the necessary expertise to effectively guide their learning process.

The third factor, 'Confidence in Guidance Provided', relates to the developmental aspect of effective teaching. Items loading on this factor highlight students' doubts about the study plan and their unanswered questions, suggesting that they perceive experienced professors as better at providing clear guidance and addressing concerns. This guidance can contribute to students' sense of competence and autonomy, key elements of Self-Determination Theory (Deci & Ryan, 2000). Previous studies have linked teaching experience to student achievement (Wiswall, 2013; Papay & Kraft, 2015). Our findings offer insight into this connection: students view experienced professors as more effective guides.

Contrasting with the OECD report (TALIS Board of Participating Countries, 2009), where 'teacher-student interaction' was one of 12 performance metrics, with the other 11 focused on instructional ability, this study revealed that two out of the three key factors extend beyond instruction: professors' interpersonal skills and their ability to guide students well. These findings underscore the importance of holistic approach to evaluating teaching effectiveness, recognizing the interplay between cognitive, developmental, and relational factors.

6.1 Practical Implications

The rise of the knowledge economy and technological advancements demand that education systems prepare students not only with cognitive skills but also with emotional and interpersonal competencies (Schmidt et al., 2017). Mentorship and instructor-student relationships address these needs directly. Effective teaching must encompass the interpersonal and emotional skills required for thriving in a VUCA (Volatile, Uncertain, Complex, and Ambiguous) world (Benett & Lemoine, 2014).

In India, the study's relevance is particularly timely given the introduction of the National Education Policy (NEP) 2020 (Ministry of Human Resource Development, 2020). NEP 2020 calls for reforming teacher education and enhancing the quality of teaching by focusing on mentorship and emotional intelligence, in addition to knowledge-based competencies. Moreover, in India's socio-economic landscape, marked by educational inequality, students from disadvantaged backgrounds often rely on their teachers for both academic instruction and emotional support (Gustafsson & Nilson, 2016). Enhancing instructor-student relationships and mentorship can significantly improve outcomes for marginalized and first-generation learners (Papay & Kraft, 2015).

6.2 Recommendations

To improve instructor training programs, especially in the context of Indian education, several key recommendations emerge. First, institutions should prioritize the enhancing instructors' interpersonal competencies through regular development workshops that incorporate role-playing, peer feedback, and emotional intelligence training. Establishing mentorship programs that pair seasoned educators with novice or less experienced members of the faculty can provide invaluable real-time guidance and help develop effective student interaction skills.

Second, institutions should focus on strengthening instructors' pedagogical confidence through continuous professional development. This should include advanced subject-matter training, exposure to innovative teaching methodologies and technology integration. Additionally, institutions should sponsor higher education opportunities to deepen instructors' expertise in their respective fields.

Third, institutions must improve instructors' guidance capabilities through structured feedback mechanisms and specialized training in advising and counseling techniques. Institutional policies should be reformed to recognize soft skills development and formally incorporate student feedback in faculty evaluations. These changes will create an environment that supports holistic student development while enriching the teaching experience.

6.3 Limitations

The study has some limitations that future research could address. First, the item 'My teacher uses words that I do not understand' did not load significantly on any of the three factors extracted. This might be because the use of unfamiliar words is more related to the specific subject matter than the professor's knowledge, experience, or interpersonal skills. Technical subjects inherently involve more jargon, which can affect comprehension. While introducing new vocabulary is a crucial part of learning, particularly in technical subjects, the ambiguous nature of this item–whether it measures necessary technical terminology or poor communication–might explain its weak factor loading. Future studies could address this by including multiple items that distinguish between necessary technical vocabulary and unclear communication, possibly examining variations across different disciplines.

Second, the sample was limited to management students from two universities in one city in India. While this allowed control for student characteristics, it constrains the generalizability of the findings. Cross-cultural studies could examine whether the factor structure of teaching effectiveness varies across different educational systems and disciplines.

Finally, the study's cross-sectional design captured student perceptions at a single point in time. Longitudinal studies would be valuable to investigate how these perceptions evolve over the course of a program and how they relate to objective measures of student learning and subjective measures of teaching effectiveness.

Beyond these methodological limitations, it is important to acknowledge broader concerns regarding the use of student evaluations themselves. Some studies (Gigliotti & Buchtel, 1990; d'Apollonia & Abrami, 1997) highlight instructors' negative attitudes towards student evaluations, noting that these biased perceptions limit their usefulness as tools for improvement. Future research should explore these attitudes and develop strategies to address them, ensuring that student feedback is used constructively to enhance teaching effectiveness.

References

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics, 25*(1), 95–135.

Akiba, M., LeTendre, G., & Scribner, J. (2007). Teacher quality, opportunity gap, and national achievement in 46 countries. *Educational Researcher, 36*(7), 369–387.

Alemayehu, Y. K., Bushen, O. Y., & Muluneh, A. T. (2009). Evaluation of HIV/AIDS clinical care quality: the case of a referral hospital in North West Ethiopia. *International Journal for Quality in Health Care, 21*(5), 356-62.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Prentice-Hall.

Bass, B. M. (1985). *Leadership and performance beyond expectations.* Free Press.

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education.* University of Chicago Press.

Benett, N., & Lemoine, G. J. (2014). What VUCA really means for you. *Harvard Business Review, 92*(1/2), 27.

Blank, R. K., & De Las Alas, N. (2009). *The effects of teacher professional development on gains in student achievement: How meta-analysis provides scientific evidence useful to education leaders*. Council of Chief State School Officers. https://files.eric.ed.gov/fulltext/ED544700.pdf

Blazar, D. (2015). Effective teaching in elementary mathematics: Identifying classroom practices that support student achievement. *Economics of Education Review, 48*, 16–29.

Blomeke, S., Olsen, R., & Suhl, U. (2016). Relation of student achievement to the quality of their teachers and instructional quality. In T. Nilson & J. Gustafsson (Eds.), *Teacher quality, instructional quality and student outcomes*. IEA Research for Education (Vol. 2, pp. 21–50). Springer. https://link.springer.com/chapter/10.1007/978-3-319-41252-8_2

Boyd, D. J., Grossman, P. L., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis, 31*(4), 416–440.

Boyd, D., Grossman, P., Lankford, H., & Loeb, S. (2006). How changes in entry requirements alter the teacher workforce and affect student achievement. *Education Finance and Policy, 1*(2), 176–216.

Burns, J. M. (1978). *Leadership.* Harper & Row.

Burroughs, N., Gardner, J., Lee, Y., Guo, S., Touitou, I., Jansen, K., & Schmidt, W. (2019). *Teaching for excellence and equity: Analyzing teacher characteristics, behaviors and student outcomes with TIMSS* (Vol. 6). The International Association for the Evaluation of Educational Achievement. https://doi.org/10.1007/978-3-030-16151-4

Cashin, W. E. (1992). Student ratings: The need for comparative data. *Evaluation and Faculty Development, 12*, 1–6.

Cattaneo, M., Oggenfuss, C., & Wolter, S. (2016). *The more, the better? The impact of instructional time on student performance*. (IZA Discussion Paper No. 9797). Institute for the Study of Labor. ftp.iza.org/dp9797.pdf

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review, 104*(9), 2633–2679.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of Human Resources, 41*(4), 778–820.

Cohen, P. A. (1981). Student ratings of instruction and student achievement: A meta-analysis of multisection validity studies. *Research in Higher Education, 51*, 281-309.

d'Apollonia, S., & Abrami, P. C. (1997). Navigating student ratings of instruction. *American Psychologist, 52*(11), 1198-1208. http://dx.doi.org/10.1037/0003-066X.52.11.1198

Darling-Hammond, L., Holtzman, D. J., Gatlin, S. J., & Vasquez Heilig, J. (2005). Does teacher preparation matter? Evidence about teacher certification, Teach for America, and teacher effectiveness. *Education Policy Analysis Archives, 13*, 42. https://epaa.asu.edu/ojs/article/view/147

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Epskamp, S. (2015). semPlot: Unified Visualizations of Structural Equation Models. *Structural Equation Modeling: A Multidisciplinary Journal, 22*(3), 474–483. https://doi.org/10.1080/10705511.2014.93784712

Garet, M. S., Heppen, J. B., Walters, K., Parkinson, J., Smith, T. M., Song, M., et al. (2016). *Focusing on mathematical knowledge: The impact of content-intensive teacher professional development* (NCEE 2016-4010). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, US Department of Education. https://ies.ed.gov/ncee/pubs/20094010/

Gigliotti, R. J., & Buchtel, F. S. (1990). Attributional bias and course evaluations. *Journal of Educational Psychology, 82*(2), 341-351. http://dx.doi.org/10.1037/0022-0663.82.2.341

Gignac, G. E. (2015). Estimating the strength of a general factor: Coefficient omega hierarchical. *Industrial and Organizational Psychology, 8*(13).

Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis*. National Comprehensive Center for Teacher Quality. http://www.gtlcenter.org/sites/default/files/docs/LinkBetweenTQandStudentOutcomes.pdf

Greenwald, A. G., & Gilmore, G. (1997). Grading leniency is a removable contaminant of student ratings. *American Psychologist, 52*(11), 1209-1217. http://dx.doi.org/10.1037/0003-066X.52.11.1209

Gustafsson, J., & Nilson, T. (2016). The impact of school climate and teacher quality on mathematics achievement: A difference-in-differences approach. In T. Nilson & J. Gustafsson (Eds.), *Teacher quality, instructional quality and student outcomes* (Vol. 2, pp. 81–95). Springer. https://link.springer.com/chapter/10.1007/978-3-319-41252-8_4

Hanushek, E., & Luque, J. (2003). Efficiency and equity in schools around the world. *Economics of Education Review, 22*(5), 481–502.

Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of Public Economics, 95*(7–8), 798–812.

Harrison, P. D., Ryan, J. M., & Moore, P. (1996). College students' self-insight and common implicit theories of rating of teaching effectiveness. *Journal of Educational Psychology, 88*(4), 775-782. http://dx.doi.org/10.1037/0022-0663.88.4.775

JASP Team. (2024). *JASP* (Version 0.18.3) [Computer software].

Jerrim, J., Lopez-Agudo, L., Marcenaro-Gutierrez, O., & Shure, N. (2017). What happens when econometrics and psychometrics collide? An example using the PISA data. *Economics of Education Review, 61*, 51–58.

Kline, R. B. (2005). *Principles and practice of structural equation modeling*. The Guilford Press.

Kyriazos, T. A. (2018). Applied psychometrics: Sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology, 9*(8).

Ladd, H. F., & Sorenson, L. C. (2017). Returns to teacher experience: Student achievement and motivation in middle school. *Education Finance and Policy, 12*(2), 241–279. https://www.mitpressjournals.org/doi/10.1162/EDFP_a_00194

Lavy, V. (2015). Do differences in schools' instruction time explain international achievement gaps? Evidence from developed and developing countries. *The Economic Journal, 125*(11), 397–424.

Lee, S.-Y., & Song, X.-Y. (2004). Evaluation of the Bayesian and maximum likelihood approaches in analyzing structural equation models with small sample sizes. *Multivariate Behavioral Research, 39*(4).

Luschei, T., & Chudgar, A. (2011). Teachers, student achievement, and national income: A cross-national examination of relationships and interactions. *Prospects, 41*, 507–533.

Marsh, H. W., & Hau, K. T. (1999). Confirmatory factor analysis: Strategies for small sample sizes. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 251–284). Sage.

McKeachie, W. J. (1997). Student ratings: The validity of use. *American Psychologist, 52*(11), 1218-1225. http://dx.doi.org/10.1037/0003-066X.52.11.1218

Ministry of Human Resource Development. (2020). *National Education Policy 2020.* Government of India.

Montt, G. (2011). Cross-national differences in educational achievement inequality. *Sociology of Education, 84*(1), 49–68.

Palardy, G. J., & Rumberger, R. W. (2008). Teacher effectiveness in first grade: The importance of background qualifications, attitudes, and instructional practices for student learning. *Educational Evaluation and Policy Analysis, 30*(2), 111–140.

Papay, J., & Kraft, M. (2015). Productivity returns to experience in the teacher labor market: Methodological challenges and new evidence on long-term career improvement. *Journal of Public Economics, 130*, 105–119.

Patrick, C. L. (2011). Student evaluations of teaching: Effects of the Big Five personality traits, grades and the validity hypothesis. *Assessment & Evaluation in Higher Education, 36*(2), 239–249. http://dx.doi.org/10.1080/02602930903308258

Phillips, K. J. (2010). What does 'highly qualified' mean for student achievement? Evaluating the relationships between teacher quality indicators and at-risk students' mathematics and reading achievement gains in first grade. *The Elementary School Journal, 110*(4), 464–493.

Piaget, J. (1972). *The psychology of the child.* Basic Books.

Pil, F. K., & Leana, C. (2009). Applying organizational research to public school reform: The effects of teacher human and social capital on student performance. *Academy of Management Journal, 52*(6), 1101–1124.

R Core Team. (2022). R: A language and environment for statistical computing. (Version 4.1) [Computer software]. https://cran.r-project.org

Revelle, W. (2023). *psych: Procedures for psychological, psychometric, and personality research*. [R package]. https://cran.r-project.org/package=psych

Rice, J. K. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes*. Economic Policy Institute.

Rivkin, S., & Schiman, J. (2015). Instruction time, classroom quality, and academic achievement. *The Economic Journal, 125*(11), 425–448.

Rivkin, S., Hanushek, E., & Kain, J. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417–458.

Rockoff, J. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review, 94*(2), 247–252.

Rockoff, J. E., Jacob, B. A., Kane, T. J., & Staiger, D. O. (2011). Can you recognize an effective teacher when you recruit one? *Education Finance and Policy, 6*(1), 43–74.

Rosseel, Y. (2019). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36.

Schmidt, W. H., Burroughs, N. A., Zoido, P., & Houang, R. T. (2015). The role of schooling in perpetuating educational inequality: An international perspective. *Education Researcher, 44*(4), 371–386.

Schmidt, W. H., Burroughs, N. A., Zoido, P., & Houang, R. T. (2017). The role of subject-matter content in teacher preparation: An international perspective for mathematics. *Journal of Curriculum Studies, 49*(2), 111–131. https://doi.org/10.1080/00220272.2016.1153153

Schmidt, W. H., McKnight, C., Houang, R. T., Wang, H., Wiley, D., Cogan, L., & Wolfe, R. G. (2001). *Why schools matter: A cross-national comparison of curriculum and learning*. Jossey-Bass.

Staiger, D., & Rockoff, J. (2010). Searching for effective teachers with imperfect information. *Journal of Economic Perspectives, 24*(3), 97–118.

TALIS Board of Participating Countries. (2009). *Creating effective teaching and learning environments: First results from TALIS*. OECD Directorate for Education.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.

Wallace, M. R. (2009). Making sense of the links: Professional development, teacher practices, and student achievement. *Teachers College Record, 111*(2), 573–596.

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research, 73*(1), 89–122.

Wilson, S. M., & Floden, R. E. (2003). *Creating effective teachers: Concise answers for hard questions. An addendum to the report "Teacher preparation research: Current knowledge, gaps, and recommendations*". AACTE Publications.

Wiswall, M. (2013). The dynamics of teacher quality. *Journal of Public Economics, 100*, 61–78.

Woessman, L. (2003). Schooling resources, educational institutions, and student performance: The international evidence. *Oxford Bulletin of Economics and Statistics, 65*(2), 117–170.