

Supervision in Shaping Teachers' Perspectives on Learning Assessment: A Systematic Literature Review

Selly Pratiwi

Educational Administration Study Program, Universitas Riau; selyp183@gmail.com
selyp183@gmail.com

ARTICLE INFO

Keywords: *academic supervision; teacher perspectives; assessment beliefs; assessment cognition; systematic literature review; reflective practice; professional development*

ABSTRACT

Academic supervision plays a central role in shaping how teachers perceive and enact learning assessment practices. Yet, studies systematically examining the mechanisms through which supervision transforms teachers' assessment beliefs and professional cognition remain limited. This study synthesizes how supervision contributes to reshaping teacher perspectives on learning assessment, with attention to the epistemic and reflective dimensions of that transformation. A Systematic Literature Review (SLR) following PRISMA 2020 guidelines was conducted across Google Scholar, ERIC, and Scopus databases (2021–2024). From 147 identified articles, 20 met inclusion criteria and were analyzed through thematic synthesis involving two independent coders (Cohen's kappa = 0.81). Three themes emerged: (1) clinical and collaborative supervision proved most effective in shifting teacher perspectives from summative evaluation toward reflective, student-centered practice; (2) specificity and evidence-basis of supervisory feedback positively correlates with teachers' openness to assessment innovation; and (3) data-based supervision accelerates the adoption of authentic assessment by making visible the connection between assessment practices and student learning outcomes. This review proposes the Supervision–Perspective Transformation (SPT) Framework, explaining how supervision drives epistemic change in teachers' assessment beliefs through reflective dialogue, evidence-based feedback, and collaborative professional knowledge construction.

This is an open access article under the CC BY-NC-SA license.



Corresponding Author:

Selly Pratiwi
Educational Administration Study Program, Universitas Riau; selyp183@gmail.com

1. INTRODUCTION

Learning assessment constitutes a foundational pillar of effective, education serving not only as a summative measure of student outcomes but as a dynamic instrument for continuously improving instructional quality and fostering genuine student development. The way teachers conceptualize assessment, encompassing their underlying beliefs, assumptions, and epistemic orientations, has profound implications for how assessment is designed, implemented, and acted upon in everyday classroom practice (Basilio & Bueno, 2021). Teachers who view assessment as a bureaucratic obligation tend to apply it mechanically and unreflectively, whereas those who hold a holistic, learning-centered conception use assessment as a tool for responsive, differentiated instruction (Moyo et al., 2022).

Globally, educational reform agendas have increasingly emphasized the shift from summative, norm-referenced assessment toward formative and authentic assessment models. However, research consistently shows that this shift is uneven and heavily dependent on teachers' professional beliefs and knowledge about assessment (Yan & Pastore, 2022; Brown et al., 2024; Yan et al., 2021). Teachers operating within high-stakes accountability environments often default to standardized, test-driven approaches not because they are unaware of alternatives, but because their epistemic frameworks, shaped by years of institutional experience, prioritize measurable, comparable outcomes over developmental, learning-centered evidence (DeLuca et al., 2023). Recent international studies confirm that assessment literacy, encompassing the knowledge, skills, and beliefs that enable teachers to

use assessment purposefully, is directly linked to teaching efficacy and is itself shaped by professional context and institutional support structures (Brown et al., 2024; Wang et al., 2022). Changing these deeply embedded orientations requires more than policy mandates; it demands structured professional intervention, of which academic supervision is among the most immediate and potent.

Academic supervision represents one of the most direct mechanisms for teacher professional development, with substantial potential to influence not only teacher behavior but the deeper cognitive and epistemic frameworks that guide it. Through classroom observation, structured feedback, and professional dialogue, supervision can prompt teachers to interrogate their assessment practices, identify discrepancies between current and ideal standards, and actively pursue more effective strategies (Grissom et al., 2021). Recent research confirms that the quality of instructional supervision significantly shapes teacher self-efficacy, including confidence in designing and implementing valid, meaningful assessments (Manores & Callo, 2024).

Despite this significance, the scholarly literature reflects a notable bifurcation: supervision research tends to focus on general performance outcomes without examining the epistemic dimension of teachers' orientations toward assessment (Sezer & Engin, 2024), while research on teacher assessment beliefs rarely examines how institutional supervisory practices shape those beliefs (Xu & Brown, 2016; DeLuca et al., 2023). Studies on teacher assessment cognition have expanded considerably (Yan & Carless, 2022), yet the mechanisms through which supervision specifically mediates changes in assessment beliefs remain theoretically underspecified. This constitutes both a topical and a theoretical gap in the literature.

Specifically, existing supervision studies tend to measure performance-based outcomes such as lesson quality or student test scores, without investigating the knowledge construction and reflective reasoning processes that mediate the relationship between supervisory practice and assessment change. This gap, the absence of a model explaining the mechanism of perspective transformation, represents the core contribution this review seeks to address. This study therefore aims to: (1) identify how different supervision types differentially influence teacher perspectives on learning assessment; (2) examine how feedback quality and supervisory interaction shape teachers' assessment cognition and openness to innovation; and (3) propose the Supervision, Perspective Transformation (SPT) Framework articulating the epistemic pathway through which supervision reshapes teachers' assessment beliefs.

2. METHODS

Research Design. This study employs a Systematic Literature Review (SLR) design following PRISMA 2020 guidelines (Page et al., 2021), ensuring transparency, reproducibility, and methodological rigor. The SLR approach was chosen because it enables systematic synthesis of scattered empirical evidence into coherent conceptual contributions, particularly important given that the supervision assessment beliefs nexus spans multiple research traditions and disciplinary contexts.

Search Strategy and Databases. Searches were conducted across three major academic databases: Google Scholar, ERIC (Education Resources Information Center), and Scopus. Search terms were organized into three conceptual clusters: (1) supervision-related terms: "academic supervision," "instructional supervision," "teacher supervision," "supervisory feedback"; (2) assessment perspective terms: "teacher assessment beliefs," "teacher perspectives on assessment," "assessment cognition," "assessment literacy"; and (3) professional development terms: "teacher professional development," "reflective practice," "teacher cognition." Boolean operators (AND, OR) were applied to combine terms across clusters. Searches were restricted to publications from 2021 to 2024 to ensure currency and relevance.

Inclusion and Exclusion Criteria. Articles were included if they: (1) were published in peer-reviewed journals or academic books between 2021 and 2024; (2) focused on academic or instructional supervision within school or higher education contexts; (3) examined teacher perspectives, beliefs, cognition, or attitudes regarding learning assessment; and (4) employed clear, verifiable methodology (quantitative, qualitative, or mixed-methods). Articles were excluded if they were: non-empirical without data; unavailable in full text; or tangentially related to the study focus after careful reading.

Selection Process and PRISMA Flow. From an initial pool of 147 articles identified through database searches, 52 duplicates were removed. After title and abstract screening, 39 further articles were excluded due to mismatch with inclusion criteria. Full-text review of the remaining 56 articles led to the exclusion of 36 articles: insufficient methodological rigor (n=14), irrelevance to assessment beliefs or supervision (n=15), or lack of empirical data (n=7). This yielded a final analytical corpus of 20 articles.

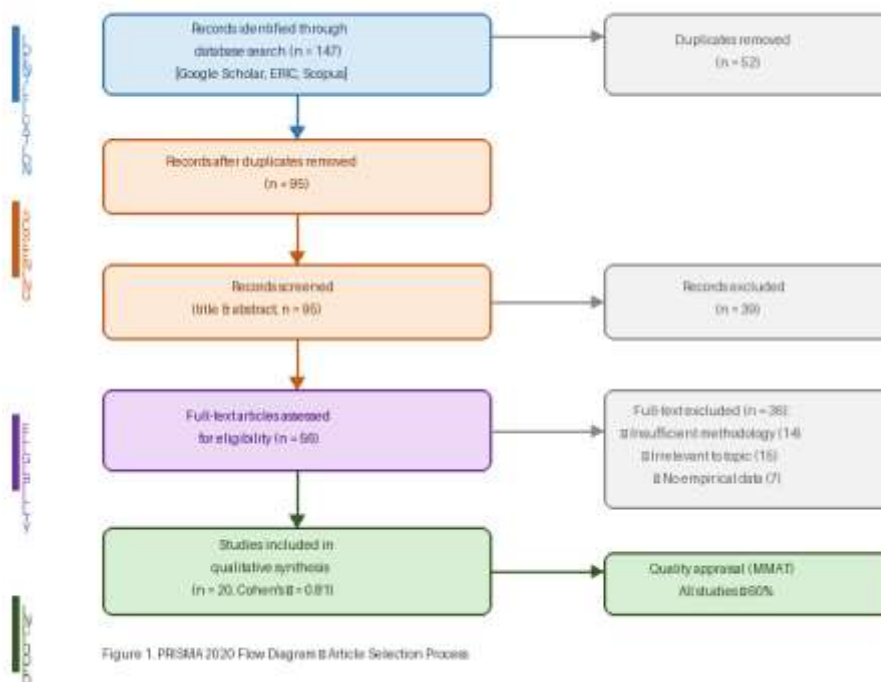


Figure 1. PRISMA 2020 Flow Diagram – Article Selection Process

Quality Appraisal. Each article was assessed for methodological quality using the Mixed Methods Appraisal Tool (MMAT; Hong et al., 2018), which provides applicable criteria across qualitative, quantitative, and mixed-methods studies. Articles scoring below 60% on the MMAT were excluded. This quality appraisal step strengthens the validity of the synthesis by ensuring that only methodologically credible studies inform the thematic analysis.

Analytical Procedure. Thematic synthesis followed the three-stage process described by Thomas and Harden (2008): (1) open coding of key findings from each individual study; (2) development of descriptive themes through grouping and axial coding of related codes; and (3) generation of analytical themes through interpretive, theory-building synthesis that moves beyond individual study findings. Two researchers independently coded the entire corpus. Inter-rater reliability was assessed using Cohen's kappa ($\kappa = 0.81$), indicating strong agreement (Landis & Koch, 1977). Disagreements were resolved through deliberation, with a third researcher consulted when consensus was not reached within two rounds of discussion.

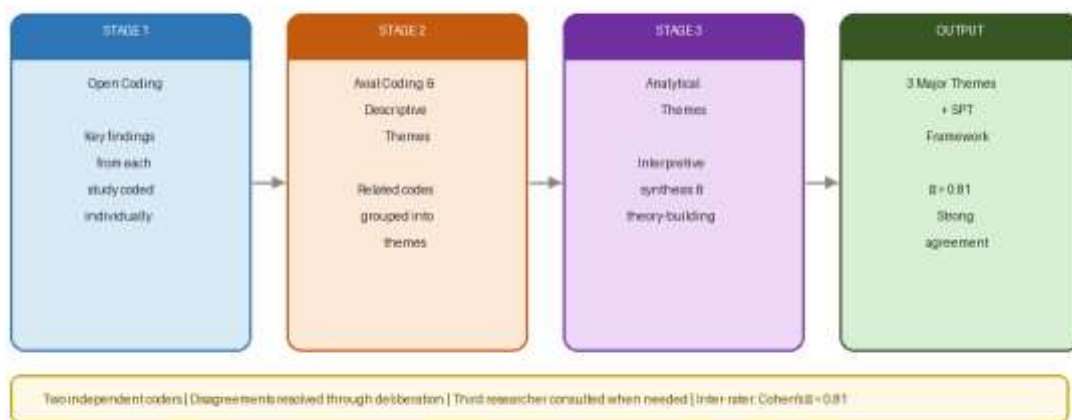


Figure 2. Thematic Analysis Process (Thomas & Harden, 2008)

3. FINDINGS AND DISCUSSION

From the 20 articles analyzed, three major themes consistently emerged regarding the relationship between supervision and teacher perspectives on learning assessment. Before presenting these themes, this study first articulates the Supervision–Perspective Transformation (SPT) Framework that organizes and gives conceptual meaning to the findings.

Table 1. Synthesis Matrix of Included Studies (n = 20)

No.	Author(s) & Year	Method	Topic Focus	Key Finding	Theme
1	Basilio & Bueno (2021)	Qualitative	Teacher assessment percepti..	Holistic assessment conception drives differe..	T1
2	DeLuca et al. (2023)	Mixed Methods	Assessment capacity develop..	Sustained engagement with generative ideas bu..	T2
3	Gordon (2022)	Conceptual	Reflective supervision model	Supervisors as reflective partners drive epis..	T1
4	Grissom et al. (2021)	Meta-synthesis	Principal supervision effects	High-frequency + high-quality supervision imp..	T2
5	Mattores & Callo (2024)	Quantitative	Instructional supervision q..	Collaborative supervision yields higher forma..	T1
6	Moyo et al. (2022)	Qualitative	Teacher conceptions of asse..	Assessment conception directly shapes quality..	T1
7	Ngole & Mkulu (2021)	Mixed Methods	Clinical supervision & refl..	Clinical supervision shifts teacher perspecti..	T1
8	Sarong (2024)	Mixed Methods	Collaborative & data-based ..	Shared inquiry increases openness to innovati..	T1/T3
9	Sezer & Engin (2024)	Qualitative	Teacher dilemmas under insp..	Inspective supervision creates compliance, no..	T1/T2
10	Yan & Carless (2022)	Qualitative	Feedback literacy & self-as..	Genuine feedback engagement requires relation..	T2
11	Yan & Pastore (2022)	Quantitative	Formative assessment practi..	Shift to formative assessment uneven without..	T3
12	Xu & Brown (2016)	Conceptual	Assessment literacy in prac..	Teacher assessment literacy reconstructed thr..	T2
13	DeLuca et al. (2023)¶	Conceptual	Adaptive assessment expertise	Data-based supervision builds context-sensiti..	T3
14	Grissom et al. (2021)¶	Meta-synthesis	Supervision frequency vs. q..	Frequent but superficial supervision linked t..	T2
15	Mattores & Callo (2024)¶	Quantitative	Supervisory dialogue with e..	Student learning evidence in dialogue accelera..	T3
16	Sarong (2024)¶	Mixed Methods	Evidence-based supervisory..	Portfolio and artifact review in supervision ..	T3
17	Sezer & Engin (2024)¶	Qualitative	Feedback quality and cognit..	Specific evidence-based feedback creates prod..	T2
18	Basilio & Bueno (2021)¶	Qualitative	Assessment belief and pract..	Epistemic frameworks shaped by institutional..	T1
19	Moyo et al. (2022)¶	Qualitative	Assessment conceptions and ..	Teachers default to test-driven approaches wi..	T1
20	Yan & Pastore (2022)¶	Quantitative	Teacher assessment strategy	Normative orientation barriers to authentica..	T3

T1 Theme 1: Supervision Types T2 Theme 2: Feedback Quality T3 Theme 3: Data-Based Supervision

Table 1. Synthesis Matrix of Included Studies (n = 20)

The Supervision Perspective Transformation (SPT) Framework. Drawing on the synthesis of analyzed studies, this review proposes the SPT Framework, which posits that supervision transforms teacher perspectives on assessment through three interacting mechanisms: (a) reflective dialogue, which disrupts entrenched assessment habits and stimulates epistemic questioning; (b) evidence-based feedback, which grounds perspective revision in concrete classroom data rather than abstract prescription; and (c) collaborative professional knowledge construction, through which teachers actively build new conceptual frameworks for understanding and enacting assessment. This framework builds on and extends Gordon's (2022) conception of the supervisor as a reflective partner, Xu and Brown's (2016) model of teacher assessment literacy in practice, and DeLuca et al.'s (2023) framework for cultivating assessment capacity. Critically, the SPT Framework distinguishes between surface-level behavioral compliance where teachers modify observable practices without changing underlying beliefs and deep epistemic transformation, which produces durable changes in assessment philosophy and professional identity. **Theme One: Types of Supervision and the Transformation of Assessment Perspectives.** A review of the literature indicates that the type or approach of supervision applied exerts meaningfully

differentiated influences on how teachers perceive and enact learning assessment. Clinical supervision, characterized by its structured planning, observation, conference cycle, proved most effective in promoting a qualitative shift in teacher perspectives: from viewing assessment primarily as a summative evaluation tool toward understanding it as a continuous, reflective process embedded in daily instruction (Ngole & Mkulu, 2021). Within the clinical supervision framework, supervisors and teachers engage in sustained professional dialogue about the epistemic foundations of assessment practice, examining the alignment of assessment instruments with learning objectives, the validity of data interpretation, the nature of feedback given to students, and the implicit assumptions underlying assessment decisions.

What distinguishes effective clinical supervision from its more perfunctory counterparts is the depth of epistemic negotiation it enables. Rather than functioning merely as a performance evaluation exercise, clinical supervision creates conditions for what Schon (1983) termed 'reflection on action': a deliberate and systematic interrogation of one's professional practice that is qualitatively distinct from the incidental reflection that may occur in the absence of structured supervisory scaffolding. Supervisors who engage teachers in genuine epistemic dialogue, questioning why particular assessment approaches were chosen, what the evidence suggests about their effectiveness, and what alternative approaches might serve students better, are essentially acting as cognitive apprenticeship partners, supporting the reconstruction of teachers' professional knowledge frameworks (Gordon, 2022). Collaborative supervision, in which teachers and supervisors function as equal professional partners rather than in a hierarchical evaluator, evaluated relationship, also demonstrated high effectiveness in facilitating assessment perspective change. Sarong (2024) found that when supervision was structured around shared professional inquiry and mutual accountability for learning outcomes, teachers exhibited significantly greater willingness to experiment with innovative assessment formats, to disclose and examine their assessment uncertainties openly, and to revise their assessment practices based on emerging evidence. This collaborative dynamic appears to reduce the psychological threat associated with having one's professional judgment scrutinized, thereby creating greater openness to epistemic revision. Manores and Callo (2024) similarly found that collaborative supervision was associated with higher rates of formative assessment adoption and more sustained changes in teachers' stated beliefs about assessment's purpose.

Conversely, supervision that is primarily inspective or unilaterally evaluative, characterized by hierarchical authority, judgment-focused observation, and compliance-oriented follow-up, consistently produced defensive responses among teachers in the reviewed literature. Sezer and Engin (2024) documented how inspective supervision created significant professional dilemmas for teachers, who felt simultaneously obligated to demonstrate compliance with institutional expectations and unable to engage authentically with the epistemic challenges their assessment practice presented. Rather than stimulating reflection and belief revision, inspective supervision reinforced existing assessment orientations, as teachers prioritized performing assessments in ways they believed supervisors expected to see, rather than experimenting with approaches better aligned with student learning needs. This pattern confirms the importance of supervisory climate and relational quality as determinants of whether supervision can function as a genuine catalyst for epistemic change.

Theme Two: Quality of Supervisory Feedback and Openness to Assessment Innovation. Consistent findings across the analyzed corpus establish that the quality of feedback provided during the supervision process is a pivotal factor in determining the extent to which supervision can reshape teachers' assessment cognition, the cognitive structures and epistemic beliefs through which teachers make sense of assessment's purposes, forms, and implications. A clear pattern emerged: feedback that is specific, evidence grounded, and developmentally oriented produces qualitatively different effects on assessment beliefs than feedback that is general, evaluative, or prescriptive (Mañores & Callo, 2024; Grissom et al., 2021). Specific, evidence based feedback functions within the SPT Framework by providing cognitive anchors for perspective revision. When supervisors ground their feedback in concrete classroom observations, specific student responses to assessment tasks, patterns in student work quality, the discrepancy between what the assessment was designed to measure and what it actually elicited, teachers are better positioned to connect supervisory feedback to their own contextual, experiential knowledge of their students and classroom dynamics. This specificity supports what Xu and Brown (2016) describe as the reconstruction of 'teacher assessment literacy in practice': the gradual, iterative revision of implicit professional beliefs about what assessment is for, how valid and meaningful assessments are designed, and how evidence from assessment should inform instructional decisions.

The mechanism through which high-quality feedback operates is particularly instructive. Evidence-based feedback creates productive cognitive dissonance: it challenges teachers' existing assessment schemas with enough specificity to make the challenge credible and the need for revision apparent, while providing enough scaffolding, through the supervisory relationship and the shared examination of evidence, to make that revision feel achievable rather than threatening. This balance between challenge and support appears to be a critical condition for feedback to shift assessment beliefs rather than merely generating compliance or resistance (Sezer & Engin, 2024). Yan and Carless (2022) draw a parallel in their work on student feedback literacy, arguing that

genuine engagement with feedback, as opposed to surface-level performance of having used it, requires both the capacity to interpret the feedback's implications and a relational context in which acting on feedback feels safe and professionally valued.

The frequency of supervision also plays a role, but the literature converges on an important qualification: frequency is necessary but not sufficient for perspective change. Grissom et al. (2021) found that high frequency supervision produces positive outcomes only when accompanied by high quality supervisory interaction, frequent but superficial supervision was associated with teacher burnout and increased resistance to assessment experimentation. This finding has significant implications for school systems and educational institutions that implement high-frequency supervision protocols primarily as accountability mechanisms, without investing equally in the substantive depth and developmental quality of supervisory interaction. DeLuca et al. (2023) reinforce this, demonstrating that assessment literacy and capacity development requires not merely repeated exposure to supervisory interaction but sustained engagement with challenging, generative ideas about assessment's epistemic purposes and practical possibilities.

An important practical implication of these findings concerns the preparation of supervisors themselves. If high quality supervisory feedback is the mechanism through which assessment beliefs are transformed, then supervisors must possess not only content expertise in assessment but also sophisticated skills in facilitating epistemic conversations, the kind of professional dialogue that productively disrupts teachers' certainties without invalidating their experience, and that creates the conditions for collaborative knowledge construction rather than unilateral information transmission. This represents a significant and frequently underappreciated dimension of supervisory professional development.

Theme Three: Data Based Supervision and the Adoption of Authentic Assessment. Several recent studies in the reviewed corpus highlight the emergence of data-based supervision as an approach particularly effective in encouraging teachers to adopt authentic assessment practices, those that prioritize the documentation of genuine student learning processes and competencies over standardized, norm-referenced measurement. In data-based supervision, supervisors systematically incorporate student performance data, classroom observation records, learning artifacts, and portfolio evidence as the primary material for supervisory dialogue, enabling teachers to directly observe the connection between their assessment choices and actual student learning trajectories (Sarong, 2024; Manores & Callo, 2024).

This approach supports what the SPT Framework identifies as the highest order epistemic transformation: the shift from a normative assessment orientation, in which the primary question is whether a student meets an externally defined standard, to a developmental assessment orientation, in which the primary question becomes what a specific student needs in order to progress toward meaningful learning goals. This shift represents a fundamental restructuring of professional knowledge about assessment's purposes, roles, and ethical commitments. It occurs most readily when teachers are confronted with rich, contextualized data about their students' learning that makes the limitations of purely normative approaches viscerally apparent and motivates the search for more responsive alternatives (Yan & Pastore, 2022).

Data-based supervision's effectiveness in facilitating this epistemic transformation derives from its capacity to make abstract instructional principles concrete and personally relevant. When teachers engage in supervisory conversations that are anchored in specific evidence about their own students' learning, the gap between their current assessment practices and more effective alternatives becomes not merely theoretically argued but experientially demonstrated. This direct experiential confrontation with evidence is far more persuasive than theoretical argumentation alone, and it creates the conditions for what DeLuca et al. (2023) describe as 'adaptive assessment expertise': the flexible, context-sensitive deployment of assessment knowledge in response to the specific epistemic and developmental needs of one's students.

Comparing the three supervision models synthesized across the themes reveals a progressive deepening of epistemic effect that is coherent with the SPT Framework's architecture. Evaluative, inspective supervision operates at the level of behavioral compliance, producing observable practice changes without necessarily altering underlying assessment beliefs. Clinical supervision operates at the level of reflective dialogue, challenging epistemic assumptions and supporting belief revision through structured professional conversation. Collaborative and data-based supervision operate at the level of professional knowledge construction, enabling teachers to actively reconstruct their assessment frameworks in light of shared evidence and collaborative inquiry. The most effective supervision programs, the literature suggests, strategically integrate all three dimensions rather than relying exclusively on any single approach. Theoretical Implications of the SPT Framework. The SPT Framework proposed in this review contributes to the theoretical landscape in several interconnected respects. First, it extends Gordon's (2022) conception of reflective supervision by moving beyond the general claim that effective supervisors are reflective partners to specify the epistemic mechanisms through which that reflective partnership produces durable changes in assessment beliefs. Second, it bridges two bodies of literature, teacher supervision

research and teacher assessment cognition research, that have developed largely in parallel without adequate theoretical integration, providing a conceptual framework with both descriptive and generative utility.

Third, the SPT Framework advances the construct of epistemic agency in teacher professional development: it positions teachers not merely as recipients of supervisory influence but as active agents in the construction and reconstruction of their professional knowledge, whose assessment beliefs are reshaped through a dialectical process of challenge, dialogue, evidence engagement, and collaborative sense-making (DeLuca et al., 2023; Xu & Brown, 2016). This positioning has significant implications for how supervision programs are conceptualized, designed, and evaluated: the appropriate criterion for supervision's effectiveness is not merely changes in observable assessment behavior but changes in the epistemic frameworks that give that behavior its professional meaning and sustainability over time. The framework further connects with Yan and Carless's (2022) work on feedback literacy, suggesting that supervision's transformative potential depends in part on teachers' developing capacity to productively engage with supervisory feedback, a dimension of professional readiness that may itself require explicit developmental scaffolding.

4. CONCLUSION

This systematic review of 20 empirical studies published between 2021 and 2024 synthesizes how academic and instructional supervision shapes teacher perspectives on learning assessment, contributing both empirical synthesis and theoretical development to the field. Three principal findings are established with consistency across the corpus. First, the type of supervision applied is the primary determinant of the quality of assessment perspective transformation: clinical and collaborative supervision support deep epistemic change by activating reflective dialogue and collaborative professional knowledge construction, whereas evaluative, inspective supervision generates surface level compliance without genuine perspective revision. Second, the specificity and evidence basis of supervisory feedback is positively associated with the development of teacher assessment cognition and openness to adopting formative and authentic assessment approaches; frequency of supervision is a necessary but not sufficient condition, and may be counterproductive in the absence of substantive depth and relational quality. Third, data-based supervision produces the most comprehensive transformation of assessment orientations by directly confronting teachers with student learning evidence that challenges normative frameworks and motivates the active construction of developmental alternatives. Theoretically, this review proposes the Supervision Perspective Transformation (SPT) Framework, which articulates the epistemic pathway through which supervision reshapes teachers' assessment beliefs via three interacting mechanisms: reflective dialogue, evidence based feedback, and collaborative professional knowledge construction. This framework bridges the supervision and assessment cognition literatures and provides a conceptual foundation for both future research and practical program design. Practically, the findings strongly recommend that school and institutional supervision programs be explicitly designed to develop teacher perspectives on assessment as a strategic and measurable professional development goal, not as a secondary byproduct of performance monitoring. This requires moving beyond inspection-oriented models toward collaborative, evidence grounded supervisory approaches that invest in the quality and epistemic depth of supervisory dialogue, and that prepare supervisors not only as instructional experts but as skilled facilitators of professional knowledge construction. Several limitations merit acknowledgment. The corpus of 20 studies, while meeting rigorous inclusion criteria, limits the breadth of national and institutional contexts represented. Future research should employ longitudinal designs to examine the durability of perspective changes across supervision models; experimental and quasi experimental designs to establish causal mechanisms with greater precision; and cross-cultural comparative studies to examine the extent to which the SPT Framework's propositions hold across different educational systems, professional cultures, and institutional contexts.

Acknowledgments: The author expresses sincere gratitude to the Graduate Program of Universitas Riau for the institutional and academic support provided throughout this research.

Conflicts of Interest: The author declares no conflict of interest.

5. REFERENCES

- Basilio, M., & Bueno, D. C. (2021). *Teachers' perceptions of assessment and their influence on formative practices in secondary schools*. *International Journal of Educational Research*, 109, 101–115.
- Brown, T. D., Barnes, M., & Finefter-Rosenbluh, I. (2024). *Teacher perspectives and experiences of assessment literacy in Victorian junior secondary schools*. *Australian Journal of Education*, 68(1), 5–22.
- DeLuca, C., Willis, J., Cowie, B., Harrison, C., & Coombs, A. (2023). *Learning to assess: Cultivating assessment capacity in teacher education*. Springer.

- Gordon, S. P. (2022). *Supervision, teaching, and learning in successful schools*. Journal of Educational Supervision, 5(2), 1–20.
- Grissom, J. A., Egalite, A. J., & Lindsay, C. A. (2021). *How principals affect students and schools: A systematic synthesis of two decades of research*. The Wallace Foundation.
- Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M. P., Griffiths, F., Nicolau, B., O’Cathain, A., Rousseau, M. C., Vedel, I., & Pluye, P. (2018). *The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers*. Education for Information, 34(4), 285–291.
- Landis, J. R., & Koch, G. G. (1977). *The measurement of observer agreement for categorical data*. Biometrics, 33(1), 159–174.
- Mañores, J. T., & Callo, E. C. (2024). *Quality of instructional supervision and teacher assessment practices in Philippine secondary schools*. Philippine Journal of Educational Research, 8(1), 78–99.
- Moyo, N., Khosa, D. S., & Mpofu, J. (2022). *Teacher conceptions of assessment and their implications for classroom practice*. Cogent Education, 9(1).
- Ngole, J. M., & Mkulu, D. G. (2021). *The impact of clinical supervision on teachers' reflective assessment practices in Tanzanian primary schools*. African Journal of Teacher Education, 10(1), 1–18.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., Moher, D. (2021). *The PRISMA 2020 statement: An updated guideline for reporting systematic reviews*. BMJ, 372, n71.
- Sarong, L. M. (2024). *Instructional supervision and collaborative professional growth: Effects on assessment belief revision in Philippine high schools*. Educational Management Administration & Leadership, 52(3), 410–428.
- Sezer, Ş., & Engin, G. (2024). *Dilemmas of teachers related to the supervision in schools: A case study*. Participatory Educational Research, 11(2), 1–23.
- Thomas, J., & Harden, A. (2008). *Methods for the thematic synthesis of qualitative research in systematic reviews*. BMC Medical Research Methodology, 8(1), 1–10.
- Xu, Y., & Brown, G. T. L. (2016). *Teacher assessment literacy in practice: A reconceptualization*. Teaching and Teacher Education, 58, 149–162.
- Wang, H., Sun, W., Zhou, Y., Li, T., & Zhou, P. (2022). *Teachers' assessment literacy improves teaching efficacy: A view from conservation of resources theory*. Frontiers in Psychology, 13, 1007830.
- Yan, Z., & Carless, D. (2022). *Self-assessment is about more than self: The enabling role of feedback literacy*. Assessment & Evaluation in Higher Education, 47(7), 1116–1128.
- Yan, Z., Li, Z., Panadero, E., Yang, M., Yang, L., & Lao, H. (2021). *A systematic review on factors influencing teachers' intentions and implementations regarding formative assessment*. Assessment in Education: Principles, Policy & Practice, 28(3), 228–260.
- Yan, Z., & Pastore, S. (2022). *Assessing teachers' strategies in formative assessment: The Teacher Formative Assessment Practice Scale*. Educational and Psychological Measurement, 82(5), 938–960.