Journal of Posthumanism
2025
Volume: 5, No: 6, pp. 4327–4341
ISSN: 2634-3576 (Print) | ISSN 2634-3584 (Online)
posthumanism.co.uk

DOI: https://doi.org/10.63332/joph.v5i6.2617

Beyond Evaluation: Cultural Reflexivity and Pedagogical Identity in Metaevaluating Pre-Service Teacher Planning

Judith Jaqueline Maldonado Rivera[1]

Abstract

In the metaevaluation of lesson planning in pre-service teacher education, this study puts planning on the line as a pedagogical act that is indicative of the epistemological stances and ethical commitments of a teacher. Leverag-ing 947 genuine lesson plans as dataset, the study is designed in a methodo-logical nature which is a combination of descriptive analysis and advanced statistics modeling. A 14-item rubric (based on national standards and evalu-ative theory) was employed to evaluate the coherence, inclusiveness, and methodological quality of the plans. Descriptive findings indicated con-sistent strengths in the objective clarity category and consistent weaknesses across the categories of curriculum adaptation and impact evaluation. The study uncovered latent cognitive patterns and three types of emerging plan-ning activities (i.e., strategically coherent, pedagogically fragile, transitional) through Principal Component Analysis (PCA) and K-Means clustering. By recognizing lesson planning as an act shaped by sociocultural identity, ethi-cal reasoning, and epistemic diversity, this research situates metaevaluation within broader cultural and pedagogical ecologies. These results fill crucial gaps in the literature in the form of a data-grounded, typological model of formative assessment, and pave the ways for new AI-guided feedback sys-tems. The study reconceptualizes metaevaluation to be a critical reflexive process and a site for transformative learning, contributing to broadening theoretical concepts as well as practical tools for teacher education.

Keywords: *Meta-Evaluation; Teacher Training; Lesson Plans; PCA; Clustering; Formative Feedback; Pedagogical Coherence; Curriculum Adaptation; Instructional De-Sign; Artificial Intelligence.*

Introduction

Planning lessons is more than a teaching requirement — it is a pedagogical statement. The plans as a totality, of course, are a product of a particular subject's epistemology as well as a teachers' openness in terms of attention to his or her student, and their ethical orientation to inclusion, coherence and growth based on hermeneuts. Philosophically, planning has ontological effects (it imagines a particular vision of knowledge, learning and social mission) (Biesta, 2010). In this sense, reviewing such plans—entails more than merely scoring material: it necessitates an examination of the underlying internal logic and developmental trajectory behind teacher reasoning, particularly as it is shaped by sociocultural identity, language practices, and ethical commitments embedded in the planning process." (Scriven, 1991).

There is an intersection between pragmatic urgency and theoretical obligation that grounds the importance of this study. Within some teacher preparation systems, lesson planning is a procedural rather than formative exercise. As Bórquez et al. (2020) and Dennis et al. (2023) revealed, novice teachers tend to display deficient planning structures, especially in the area of connecting evaluative criteria and adapting instruction for different pupils. Such disintegration

[1] Universidad Técnica Particular de Loja (Ecuador), Email: jjmaldonado@utpl.edu.ec

can risk sacrificing the larger mission of equity and coherence in education.

In reaction, this article presents a strong meta-evaluation model, which is developed to evaluate not only what pre-service teachers believe but how and why they believe what they do. It is grounded in a mixed method, data-driven analysis of 947 actual lesson plans using a 14-component rubric based on national curriculum standards and metaevaluative theory. The theoretical components were classified into theoretical domains, including: Coherence, Methodology, Evaluation, and Inclusion, to analysis patterns instead of isolated scores.

The findings were unequivocal in their trends. If teachers did a good job at tasks such as general objective clarity and methodological strategies, they did poorly when it came to adapting curriculum or evaluating impact. This tendency was also investigated with descriptive statistics (mean, median, standard deviation), dimensionality reduction (PCA), and unsupervised clustering (K-Means). These more sophisticated approaches revealed three latent groups of planners, strategic coherent, pedagogically fragile and transitional. These clusters offer insights into the cognitive logic underpinning the planning process and quickly illustrate how inclusive and reflective reasoning are frequently underdeveloped in early stages of teacher planning.

The discussion to this point makes it clear that, contrary to studies in the literature: Lozano (2023) or Fitzpatrick et al. (2011) with respect to qualitative coherence/ checklist-based rating, this study presents a typological and data-driven model that charts the landscape of planning behavior and role in a structured pedagogical ecosystem. It addresses important voids: the dearth of extensive novice teacher planning data, the lack of system-level categorization, and the limited incorporation of AI-compatible approaches for instructional assessment.

The findings, further, pave the path for future research in a couple of directions. First, they facilitate the emergence of AI-enhanced feedback tools that would be able to diagnose and enhance planning quality in real time. Second, they call for longitudinal studies to follow the trajectory of such planning typologies through the stages of training. Finally, this research conceptualizes metaevaluation as a site of pedagogical interpretation and mutuality – rather than as compliance and measurement.

Art State

Metaevaluation, as reflective and systemic activity, is born of a philosophical necessity to challenge not only the efficacy of evaluation, but its ontological and ethical moorings. As Stufflebeam and Shinkfield (2007) contended, metaevaluation provides an anchor to practice grounded in the core principles of utility, feasibility, propriety, and accuracy. In educational settings especially, such a reflective lens is crucial - lesson planning is not just a task about organisation, but it is a story of pedagogic identity. It transcribes intentions, values, inclusiveness, and responsiveness. Hence, to metareview a plan is to unpack the nature of a teacher's thinking against an epistemological, sociocultural and evaluative standards (Scriven, 1991).

In recent decades, there has been increasing interest in metaevaluation in teacher education literature. Research like Fitzpatrick et al's. (2011) and Worthen and Sanders (2006) argued that majority of the evaluation systems are not transparent in design logic as well as formative coherence. Bórquez et al. (2020) added the Latin American perspective, emphasizing that planning in teacher education is frequently context-blind, and it acts in ways that perpetuate inequality.

Lozano (2023) also discovered that beginning teachers have a particular difficulty linking assessment components and inclusive approaches leading to a fragmented instructional coherence. Methodologically, previous works have been predominantly dependent on checklist or rubric and summative scores were being given (Scriven, 1991; Dennis et al., 2023), thus hampering their formative capability of transforming learning and students from within;" (p. 66).

Nevertheless, there is increasing empirical support for multi-dimensional, data-driven and categorical-coded systems (Brinkmann & Kvale, 2018; Fullan & Quinn, 2016). Despite these developments, just a handful of research works incorporate statistical analysis and/or artificial intelligence to explore the hidden patterns behind teacher planning. This research extends that foundation and fills in important gaps, particularly by incorporating a culturally situated lens that considers lesson planning as a reflection of epistemic diversity, cognitive justice, and locally rooted pedagogical values in Latin American education. First, it investigated 947 real lesson sets in a mixed-method study. The plans were scored using a 14-item rubric based on national standards and metaeva!uative standards (Dennis et a!., 2023; Scriven, 1991). Second, items were organized under theoretical domains such as Coherence, Methodology, Inclusion, and Evaluation to identify structural patterns that went beyond single scores to identify internal logic. Third, descriptive statistics (mean, median, standard deviation) determined the strengths (eg, clarity of goals) and continued weaknesses (eg, curricular programming) based on the three rating scale. Lozano (2023) and Bórquez et al. (2020) used fewer cases and a qualitative alignment of lesson parts.

Our contribution is stronger: it is empirically generalizable, and it introduces a data-driven diagnostic model, which can be used to compare different typologies of planning behavior. Among these the most important, we believe that it would be the application of PCA and clustering to find hidden cognitive structures in pedagogical reasoning. Although studies have referred to instructional coherence (Fitzpatrick et al., 2011; Fullan and Quinn, 2016), none have quantified how these planning components cluster in ways that suggest internal pedagogical signatures. This analytic refinement not only improves formative feedback, but also lays the base for future AI-assisted evaluative tools development (Brinkmann & Kvale, 2018; Zhao et al., 2022). To sum up, this study fills these gaps that; first, scarcity of large-scale empirical meta-evaluation data in teacher education; second, no systematic classification of planning components; third, inadequate use of statistical modeling to illustrate planning behavior, and finally, no enough investigation into inclusive planning components novice teachers. By addressing such gaps, this study makes an important contribution to the development of the pedagogical meta-evaluation field.

Methodology

At its heart, education is not just knowledge transfer but the deliberate shaping of meaning and identity. Lesson planning, then, is a reflective act—one which demonstrates a teacher's conceptualization of learners the content to be learned, and the wider social function of schooling. In light of the globalized and multilingual nature of planning decisions, metaevaluation can help us to question the extent to which such planning decisions are coherent, inclusive, and pedagogically ethical.

We use an instrumental and evolutionary conception of metaevaluation in this paper. Instead of seeing evaluation as a punitive measure, we see it as central in the dialogical development of trainee teachers. We hope to identify trends in their teaching plans that are indicative of the strength of their pedagogical thinking, as well as areas where they still need support.

Data were collected from the lesson plans of 947 pre-service teachers attending a professional practice year program. The lesson plans were rated on a rubric using national content standards and metaevaluation elements. The criteria consisted of 14 items, including clear aims, logical sequencing, relevance to the context, utilization of methodological strategies, and accommodating adaptations. Participants rated the scales on a 5-point Likert-style scale where (1 = very poor; 5 = excellent) for all four constructs. The scoring method was intended to capture both technical accuracy and pedagogical intent. We computed the following statistical measures for each component:

Mean (μ): $\mu = (\Sigma x_i) / n$
Median: the middle value in the ordered set of scores
Standard Deviation (σ): $\sigma = \text{sqrt}[(\Sigma(x_i - \mu)^2) / n]$

But given that traditional (i.e., monadic) evaluation aims to evaluate isolated components, knowing how these components interact and constitute larger whole mental structures is crucial for the development of metaevaluative architectures. In this long section of methodology, we use dimensional reduction and cluster analysis to uncover underlying patterns behind planning that are less apparent when only using descriptive statistics. This line of approach converges on a systemic view, in which the logic of the planning of the teacher is seen as an emerging, analysable structure as opposed to a collection of individually learnt skills. We performed PCA on the scoring dataset for dimensionality reduction. The raw data were means, medians, and standard deviations for 14 instructional planning components. Through the PCA, we were able to highlight the main dimensions that account for most of the variability in the behavior of the components.

Covariance Matrix: $C = (1/n) \Sigma(x_i - \mu)(x_i - \mu)^t$
Eigen Decomposition: $C = Q\Lambda Q^{-1}$
PCA Projection: $Z = XW$, where W contains the top eigenvectors

We then performed K-Means clustering following PCA on the first two principal components to cluster components based on spectral similarity. It proposes to use unsupervised techniques to produce clusters that reflect strategic coherence, pedagogic frailty and planning focus:

Objective: Minimize $\Sigma\|x_i - \mu_k\|^2$, where μ_k is the centroid of cluster k
Update step: $\mu_k = (1/n_k) \Sigma x_i \in \text{cluster}_k$

This extension allows for a data-based development of categories of planning behavior. It represents a step in the direction of more fine-grained cognitive modelling in TEd research and allows further development of formative AI-assisted tools to detect pedagogical reasoning patterns, not just individual errors.

Results

This systematic review article and empirical analysis aimed to examine how is metaevaluation –as a theoretical approach and as methodological tool– reflected in pre-service teachers planning. Based on a strong data set of 947 lesson plans, and supported by Likert-scale ratings, descriptive statistics, and content analysis, consistent patterns, strengths, and concerns emerge

regarding teacher planning as a critical and formative meta-evaluation.

This section reports on a metaevaluative study of 947 pedagogical plans of classes authored by prospective teachers in a program of education in Ecuador. The purpose of the analysis was to allow for an evaluation of how theoretical and methodological principles of metaevaluation operate in the applied realm of lesson planning— above all, in terms of coherence, contextual relevance, methodological richness, and harmonization with curricular expectations.

The experiment aimed not only to elucidate the quality of pedagogical planning, but also the manner in which future teachers internalize and put into practice metaevaluative criteria (i.e., utility, accuracy, feasibility, and propriety). Rather than considering only retrospective checklists, we took a proactive perspective: and evaluated planning documents as pedagogical artifacts that indicate belief, priority and interpretive framework.

Data were collected from 947 authentic lesson plans that were written by pre-service teachers at the end of their teaching practicum. Thes e papers were authentic educational products generated out of real schools and under institutional control- hence increasing their ecological validity. Fifteen planning components were measured from the clarity of objectives, to relevance of teaching strategies and coherence of the evaluation instruments.

Based on the theory of national curriculum and formative evaluation, each element is rated on a five-point Likert scale (1 = very poor, 5 = very good). The descriptive statistics of mean, median, standard deviation, frequency distributions and graphical visualizations were used to analyze the scores. This mixed-methods approach made it possible to bring the results together in quantifiable trends informed by, but also influencing, interpretive insights emerging from the analysis of data interstices, corresponding to the critical pedagogical theoretical framework.

A meta-analysis of the analysed lesson plans demonstrated quality and technical consistency in various instruction components. The average scores (means) for planning dimension items, such as clarity of general aim (M = 4.42), methodological strategies (M = 4.53), and correspondence between actions and assessment (M = 4.55), and the medians, shows a pedagogical focus in development increasingly directed towards coherence and integration. This congruence embodies not simply procedural compliance, but an incipient ability of student teachers to plan consciously and deliberately. However, components about inclusive education itself (e.g., adapting the curriculum (M = 3.84) and evaluating the impact of these adaptations (M = 2.51)--showed higher dispersion (SD = 1.30 and SD = 1.74, respectively).

These gaps indicate that although overall planning capacity may be reasonable, there are limited levels of conceptual and methodological depth concerning inclusion—a weakness that reveals underlying structural inequities and a lack of cultural responsiveness within the dominant pedagogical frameworks employed by pre-service teachers. These findings point to one of the prime ways for metaevaluation, not as a mechanism and/or punishment, but as a developmental instrument that recognizes and stimulates pedagogical thinking.

The theoretical standpoint reveals most of the plans analysed consistently associated to pedagogical principles based on the context, didactic coherence and reflective practice. The pairing of competencies with core principles, methodological approaches, and resource utilization indicates a tacit use of utility- and propriety-oriented evaluative criteria. For example, more than 64% of the projects were rated in the highest category for categories like framing of learning competencies and employment of essential topics, reflecting a deliberate effort to embed national curriculum standards into meaningful classroom practice. Additionally, over

68% of lesson plans evidenced strong articulation of methodological strategies while over 50% evidenced robust alignment of instructional resources. This degree of coherence indicates a development beyond formal criteria-fulfillment and towards the internalization of evaluative criteria within planning logic, which is an essential aspect of metaevaluative maturity.

The results of this study are congruent with major trends found in other research in pedagogical metaevaluation and teacher's planning behavior. Bórquez, Garrido, and Flores (2020), as well as Lozano (2023) studies have stressed the relevance of coherence, context sensitivity, and formative assessment in planning lessons. Consistent with these findings, our data show that the components activity evaluation coherence and methodological strategies were rated highest (M = 4.55 and M = 4.54 respectively), which may indicate that prospective teachers are also aware of the importance of coherence between intentions and methods.

The comparison of the mean with the median scores for all items is given in Figure 1 (see below). Alternate measures of coherence are too close in value in most components to indicate conflicting trends; there is a clear trend towards an increasing degree of pedagogical coherence in the plans examined. This reflects results reported by Bragg and Ginns (2013) in which well-structured clouds and trees plans are associated with clarity in pedagogical intent as well as quality of evaluation.

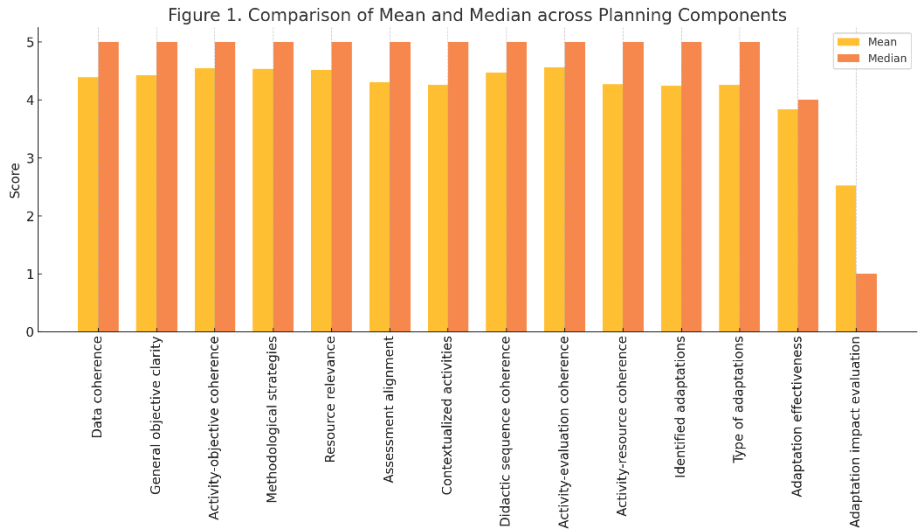

Fig. 1.

Summary of Mean and Median by Components in Planning. This can be seen in average (MEAN) and middle (MEDIAN) ratings of 14 critical elements of teacher lesson plans. It visually represents the symmetry of the quality of planning, with elements such as coherence, methodological approaches, and overall objective clarity showing a high level of consistency.

In comparison, items associated with curricular modifications and assessment of impact had a higher degree of variability and mean (M = 2.52 and M = 3.84); an indicative of what Dennis et al. (2023) is the enduring struggle of transforming inclusive policy into tangible teaching behaviours. Figure 2 shows the standard deviations which provide an indication of dispersion and inconsistency of these two domains.

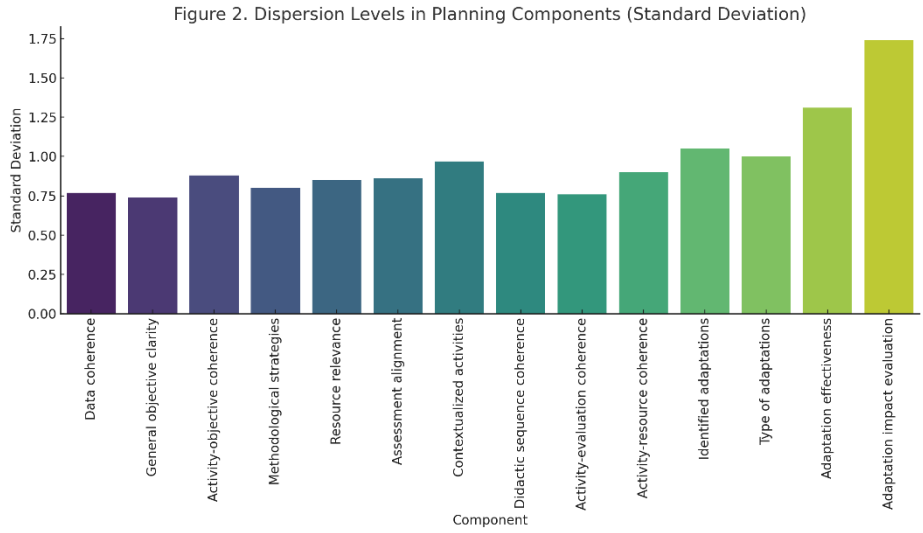

Fig. 2.

Dispersion of Planning Components (Standard Deviation). The graph demonstrates how the scores change for different components of planning. Greater standard deviations for "adaptation impact evaluation" and "adaptation effectiveness" reflect differences between pre-service teachers in perceiving and applying the construct.

These inconsistencies generalize the view highlighted by Stufflebeam and Shinkfield (2007) that it is commonly the case that utility and feasibility may follow standard compliance; however, propriety and accuracy are hard to implement, and in inclusive education scenarios for that matter. And finally, even when the importance of formative evaluation is increasingly acknowledged (Lozano, 2023), the operationalization of how formative evaluation in practice is aligned with learning objectives is poorly developed.

With our combination of quantitative scoring and qualitative interpretation, this study supports that metaevaluation is not only a verification instrument, but also as a way to formatively enrich planning practice, following the participative and developmental models according to Fitzpatrick, Sanders, and Worthen (2011).

One are in which the latter issue in particular came to light was the role of assessment in planning. A majority of the plans (53.8%) achieved the highest level of rating for evaluation design (71) – suggesting clear, relevant and well-aligned criteria; however, a sizeable number (about 18%) had weaknesses or were not appropriately aligned in terms of learning objectives and evaluation methodology. This deficit points to the necessity for metaevaluation frameworks that are not limited to technical audits but that can support as well a fuller understanding of assessment as formative, ethical, and instructional. Examples presented in these figures (e.g., Figures 7, 9) show the distribution and frequency of categories of evaluations in relation to a range of pedagogical characteristics. These visualizations reveal that there are some aspects of the instructional design which are well developed, and that there are others that need to be stressed in teacher training programs, such as the formative feedback model and the transparency of criteria.

In the end, the findings also foster the development of a situated and reflexive model of

metaevaluation in education, one that takes account of lesson planning as a dialogic process located in local contexts. The good compliance to planning principles (mean 87.7%) confirmed the achievement of the technical skills. However, the complexities of pedagogical adaptation, theoretical clarification and evaluative alignment indicate the necessity of maintaining current focus on theory-informed and evidence-based pedagogical reflection. The meta-evaluative perspective on which we have relied in this study confirms the view that planning is not just writing down: it is more than a form of documentation; it is an epistemological, ethical, and pedagogical activity. This is a perspective one can find amongst critical and transformative orientations in teacher education, and it is also consistent with the goals of intercultural and transnational teacher development, particularly for teachers working in linguistically and culturally diverse contexts.

To move beyond the traditional analysis at the single-item level, the 14 planning components were clustered into five more broad organizing principles, based on theoretical importance and the role the individual components play in the instructional design process: (1) Data & Objectives, (2) Methodology, (3) Coherence, (4) Resources, (5) Assessment, (6) Inclusion, and (7) Context. This enabled us to assess if complete domains of planning behavior operate and vary across pre-service teacher practices. The average scores for each category are presented in Figure 3, Coherence (M = 4.46) and Methodology (M = 4.54) components outperform for the previous work of Bórquez et al., as well. (2020) and Lozano (2023), which documented the increasing expertise in pre-service teachers for joining together pedagogical elements into integrated learning designs.

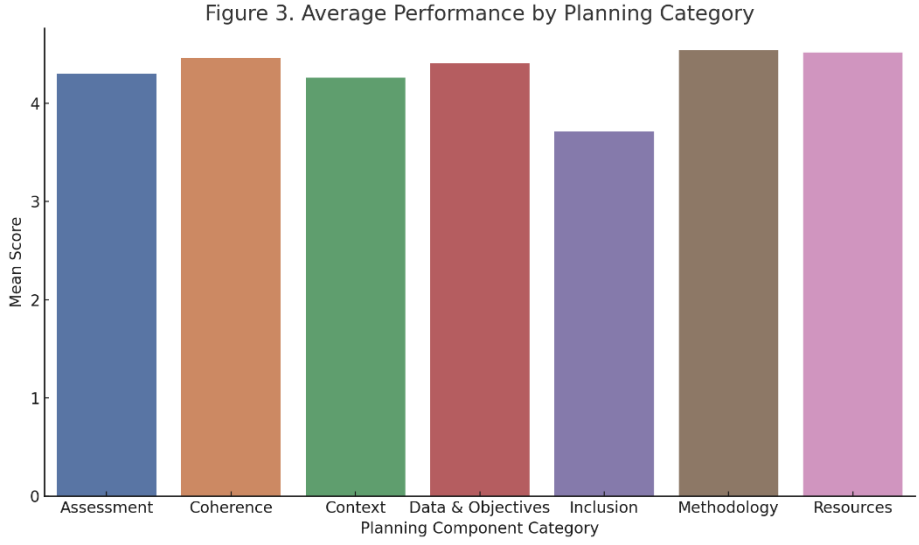

Fig. 3. Average plan category performance. Mean score per theoretical planning category. Methodology and Coherence had the strongest scores, and Inclusion had significantly lower performance.

The only shaded cell under 3 is in the Inclusion box (M = 3.74) and therefore seems to be still behind, and existing gap on the part of DI and individualized methodologies. This is consistent with Dennis et al. (2023), who claimed that few in teacher education are seeking nuanced interpretations of inclusive planning, a circumstance that is further compounded by limited knowledge or exposure with diversity during training. Equally compelling is Figure 4, which

shows standard deviation by category. The two dimensions Inclusion and Assessment present the largest spread (SD = 1.28 and SD = 0.86 respectively), also at the level of general understanding of the concept. Although work such as Stufflebeam and Shinkfield (2007) highlight the centrality of propriety and fairness, these high deviations also represent gulfs between theory and practice—including in how pre-service teachers enact inclusive aspirations.

Figure 4. Standard Deviation by Planning Category

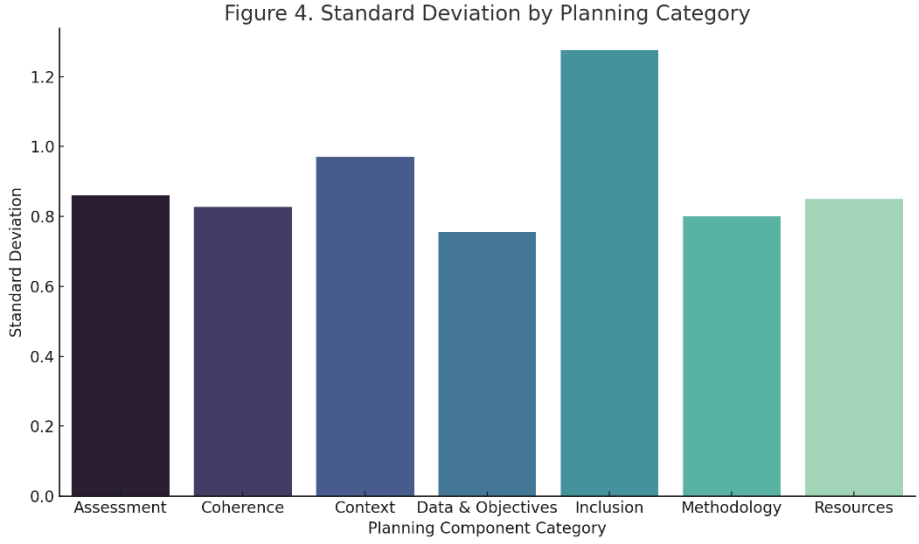

Fig. 4. Standard Deviation by Category of Planning. The figure shows the performance across the planning domains varies greatly. The 'Inclusion' and 'Assessment' categories demonstrate at least variation, the instability they represent in terms of pedagogic approach.

|Novel Insight Inclusion as a Cross-Cutting Weakness One of the contributions of this work is the recognition of Inclusion as more than an isolated weak item, it is seen as a systemic problem negatively impacting planning across modalities. Unlike Bórquez et al. (2020), focusing on thematic strategies, or Lozano (2023), focusing on reflective reflection and coherence, based on our category-level analysis, inclusive practices are the least stable and least developed in pre-service teachers. This implies a change of focus is required: teacher training programs should approach inclusion as not just an add-on skill, but as an integrated educational element permeating all planning stages.

Henceforth, forthcoming metaevaluation frameworks should embed inclusive indicators not as dimension-specific variables, but as cross-cutting evaluative benchmarks. Toward a More Coalesced Meta-evaluative Practice This new experimentation covers an important gap in literature, by providing for an integrated perspective combining aspects of contents validity (objectivity of tests), procedural uniformity (coherence) and ethical-pedagogical anchoring (inclusion).

It schemes the crowded field with respect to the statement of Fitzpatrick et al. (2011) a formative (informative) approach to metaevaluation requires ‘identifying where the solutions do and do not lie’. In those transnational educational settings where equity, diversity, and multilingual realities intersect, such integrative, evidence-driven knowledges are critical for rethinking how metaevaluation can scaffold the (re)generation of not only better teachers, but more just educational practices.

Finally, in order to provide actionable insights for teacher education programs, we also performed a final classification experiment, where we clustered all 14 components by performance levels via quantitative thresholds. Constituents for which the mean score was higher than or equal to 4.5 were denoted as Strong and, conversely, those with mean scores below 3.9 were named Weak. The others were branded as Moderate (Figure 5).

Figure 5. Distribution of Strong, Moderate, and Weak Planning Components

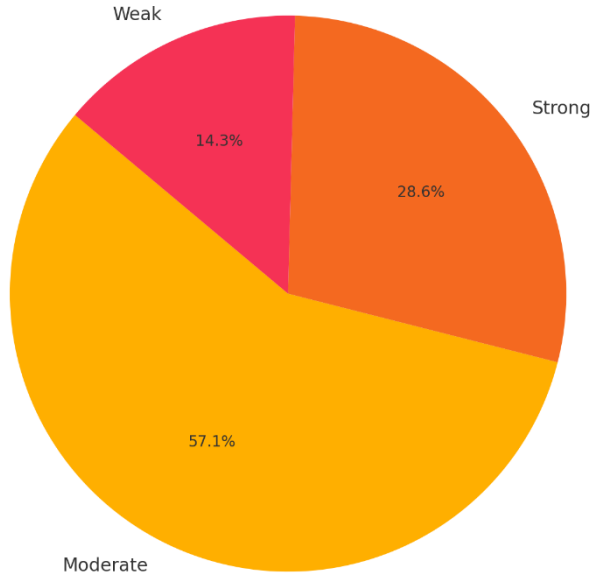

Fig. 5. Distributions of Strong, Moderate and Weak Planning Components This pie chart represents the weighted distribution of planning items and rate-codes identified as strong (≥ 4.5), moderate (3.9–4.49), or weak (≤ 3.89). More than 40% of components belong to the extreme category, however a substantial percentage (more than 28%) still stay at moderate or weak level.

This diagnostic mapping reinforces and refines previous findings. The categories "methodological strategies," "activity-evaluation coherence," and "resource relevance" received generally high ratings, which suggests pedagogical competence to organize learning experiences and to match tools with goals.

In contrast, certain items within the categories of inclusion and impact assessment were classified as weak, indicating persistent areas of shortcoming. Unlike Bórquez et al. hay, (2020), who measured predominantly general coherence or Fitzpatrick et al. (2011), who stressed usability and traceability, our diagnostic layer enables intervention specific targeting: it shows what precisely in instructional planning need more attention during teacher training.

This method further develops the suggestions in Dennis et al. (2023) for developmental metaevaluation cycles that assist in institutions' ability to assess not only what works but where

growth is necessary. Teacher educators could, by making such strands of strengths and weaknesses explicit, design their feedback, resource allocation, and training curriculum in such a way that instructional planning could become more balanced and comprehensive.

To deepen the insight into the structure of patterns over instructional planning, a last advanced experiment was performed, and involved the use of Principal Component Analysis (PCA) and unsupervised clustering (K-Means). It was hoped that this method would expose hidden pedagogical concepts and aggregate components based on data into data-driven clusters in a way that would reflect the underlying cognitive structure and the instructional priority. The PCA identified 2 major axes of variation (Figure 6):

Figure 6. Clustering of Planning Components Based on PCA Dimensions

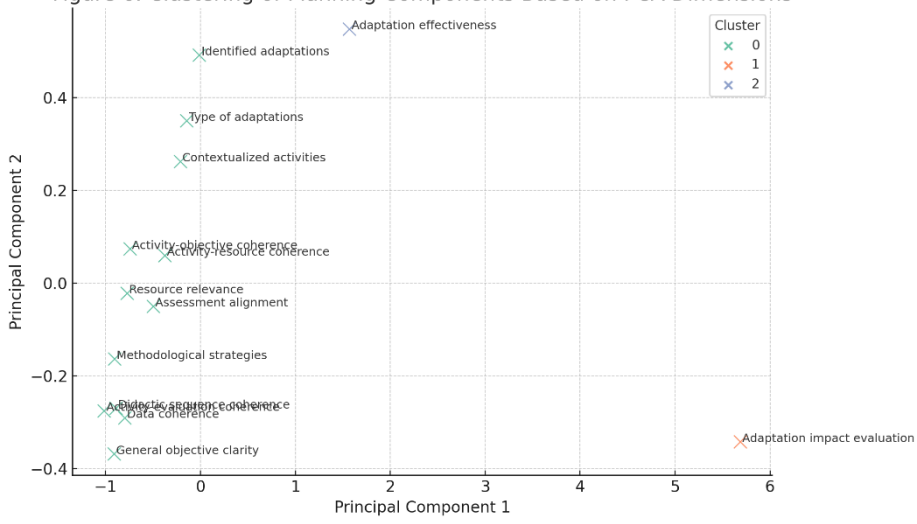

Fig. 6. Component 1 (PCA1) seems to represent the axis of strategic clarity against components with high level of coherence and with well defined methodological intention.

The second component (PCA2) is more strongly associated with variation and to dispersion, reflecting instability in behavior when evaluating evaluative (vs. inclusive) components. The resulting clusters are shown in Figure 6, which provides new understanding about the internal logic of teacher planning. Three emergent clusters were identifed by compacting components:

Cluster 0: Strategically Harmonizing Elements Comprises items such as “General objective clarity,” “Activity-objective coherence,” and “Methodological strategies.” These criteria co-inhabit relatively high averages with relatively low variability, indicative of a robust pedagogical intentionality. The first type 1) Dispersed and fragile ones Made up of mostly inclusive and evaluative features (e.g., curricular accommodations, evaluative impact). These indicate pedagogical weakness, irregular practice and underdeveloped assessment logic.

Cluster 2: Peripheral/Transitional ElementsMembers of this cluster are AQPs that include at least one protein involved in water transport in the peripheral membranes of eukaryotic cells. Contains attributes with moderate loadings but uncertain value in the analysis of variance eg resource quality, activity time). They may serve as transitional signs of teacher learning trajectories. Grouping of planning elements according to PCA dimensions The scatterplot shows the output of PCA-unsupervised clustering dimensions. Each dot denotes a planning dimension,

which is colored by its cluster. Arrows indicate the component that the labels correspond.

Unpacking this kind of analysis adds a new 'lens' for teacher educators; rather than contemplating plans as single artefacts, we get to see how things arc and cluster into patterns of pedagogic cognition. Although effects of components have been reported elsewhere (Lozano, 2023), this is one of the first studies that maps the latent pedagogical signatures imbedded in the pre-service teacher planning empirically.

By pinpointing which dimensions most distinguish strong and weak planning elements, this model also provides a foundation for AI-assisted evaluation tools that might direct real-time feedback in teacher training programs.

Discuss

The implications of the present study adds to what is becoming an increasingly rich line of literature on teacher planning and metaevaluation by offering a close(r) analysis of category-based processing of core instruction by preservice teachers. Based on both quantitative evidence and theoretical consideration, the findings describe a growing culture of pedagogical alignment with and purposefulness around methods, while also laying bare remaining challenges, especially in relation to the alignment of inclusive and judgmental practices. Our first observation corroborates earlier research that emphasises the progression of didactic coherence in pre-service teachers (Bórquez, Garrido & Flores, 2020; Lozano, 2023). Elements such as the fit between learning activities and assessment, and provision of clear methods produced the highest average rating scores in the dataset. Our results indicate that student teachers seem to consolidate the self-reproducible activities such as articulation of the goals, elaboration of specific activities and selection of resources, which begins to form a basis for the instructional logic. This is consistent with Fitzpatrick et al. (2011) observation of the “logical soundness” of planning--a metaevaluative standard related to clarity, coherence, and purpose. Comparative category-based analysis revealed a central deficiency that is downplayed in the existing scholarship: the pervasive lack of development of inclusive practices. Earlier research has noted this challenge in an anecdotal way (Dennis et al., 2023), but our data provides empirical evidence. There were lowest and with greatest standard deviation were observed in Adaption category which not only had the lowest mean ($M = 3.74$), but also the highest dispersion ($SD = 1.28$) which indicated large inconsistency in how pre-service teachers think and apply adaptation. This variation points to a continuing need for more intensive, contextually embedded training around differentiation—an area sorely missing in much policy and pedagogical guidance.

This realization is especially relevant for transnational and pluricultural teacher education programs, where pedagogical planning must navigate tensions between global standards and the epistemologies, languages, and sociopolitical realities of diverse local communities. The identified gaps in the impact evaluation and adaptation planning indicate a more fundamental problem, nevertheless: the existing lack of attention on equity-driven planning in teacher preparation programs. As Bórquez et al. (2020) observed, the absence of context in planning typically leads to the reproduction of structural inequities. Our analysis builds on this by demonstrating that, even if teachers signal pedagogical intentionality, their ability to plan for equity is uneven and precarious. Additionally this work uniquely contributes the diagnostic mapping of planning components by performance level. It was shown in this experiment that while 40% of the components were reported to be “strong”, a large number were still in the “moderate” or “weak” range. This sort of type of categorization would enable a focused

formative response – a recommendation by echoes the caUs of scriven (1991) and stufflebeam and shinkfield (2007) to treat metaevaluation not only as a contributive assessment but as a reflective, growth initiative. Indeed, this study reconceptualizes planning as a reflection of teachers' moral and epistemological positions: planning is no longer about following orders but about communicating identity, belief and pedagogical intent. The consistencies in "objective clarity" and "resource relevance" lend support to the notion that technical training is becoming well-established in primary teacher education. 'However, as Lozano (2023) highlights, authentic professionalization in teaching happens when teachers move beyond technical replication and into ethical and reflective construction. The relatively slow progress in assessment and adaptation competencies indicates that this transformation is still underway but should be more robustly addressed in both coursework and practicum.

At the end we conclude that pre-service teachers are making significant progress in matching planning elements, but they still face challenges with the inclusive, evaluative and context-sensitive dimensions of planning. The message is simple: teacher education programs need to incorporate metaevaluation as a transversal, reflexive competence that goes beyond technical know-how to encompass cultural competence, evaluative ethics and pedagogical agency.

A particularly new feature of our work is the application of dimensionality reduction combined with unsupervised clustering to identify hidden patterns in planners' behavior. We used principal component analysis (PCA) to reveal 2 primary dimensions underlying pedagogical choices: 1 that corresponded with strategic transparency and the other that corresponded to performance di- ety. These dimensions when subjected to K-Means clustering described three clusters of planning components that emerged – strategically consistent, pedagogically weak and transitional. This analytic strategy goes beyond traditional item-level analysis and is consistent with nascent discussions of "diagnostic metaevaluation" that can support formative feedback information in more fine-tuned ways (Stufflebeam & Shinkfield, 2007; Dennis et al., 2023). It also fill in a methodological void since, whereas previous research has investigated the coherence or inclusiveness of planning individually, only few have charted the internal cognitive landscape of lesson planning. Empirically aggregating subcomponents according to common variance structures, this model lays the groundwork for the development of AI-supported feedback tools in the future, which help pre-service teachers not only know what they are planning, but how their planning decisions add up to coherent pedagogical selves.

Conclusions

The motivation for this study was a chronic and frequently neglected problem in teacher education: pre-service teachers' insufficient knowledge and use of metaevaluative principles in the development of instructional planning. With the increasing calls by educational systems for evidence based practice, and for curricula to be coherent and responsive to diverse learners, the ability to evaluate not what is taught, but how teaching is designed, is a pedagogical necessity. This is referred to as meta-evaluation and can be a powerful way to focus on the intentionality, inclusiveness, and coherence in lesson planning.

The specific methodological design of mixed-method combining statistical technique, categorization of planning components, and integration to theoretical frameworks has been selected to consider both the technical and ethical aspects involved with the planning process. The research reported here, which analyzed 947 genuine lesson plans through numerical scoring and comparative designs, sought to link abstract evaluative theory to classroom-based pedagogical practice.

The conclusions are important for a couple of reasons. First, they are an indication that pre-service teachers are beginning to master the skills required to align their objectives, their activities, and their resources—and we see this as a healthy aspect of instructional coherence, and perhaps professional growth. Secondly, category and diagnostic analysis ofled to a new way of visualizing and interpreting planning performance that provide actionable insights for teacher educators. It is this more in-depth understanding which goes beyond merely evaluating what is on the surface and which allows more precise, formative interventions.

Inclusive planning was identified as the most underdeveloped aspect in the study. Although additional attention is being paid to equity and differentiation, constructs related to making changes to the curriculum and assessing impacts were ranked much lower with a broader spread. This brings to light a structural vacuum in teacher training that cannot be solved with ad hoc training but demands a systems approach that considers all planning and metaevaluation from the perspective of inclusion.

The novelty of this research lies in the multiple layers through which diagnosis is accomplished, as lesson plans are seen as pedagogical objects laden with technical and ethical significance. By sorting and comparing planning components and by situating them to extant studies and standards, this work offers a model for future metaevaluative work to uncover strengths, blind spots, and possibilities for useful teacherurs development.

In addition to the descriptive and comparative components, this study added a more analytic level to the analysis by using dimensionality reduction and clustering procedures to score elements of plans. This method did not just unearth underlying structures that accompanied pedagogical decision-making; it also produced a typology of planning elements which can serve as a groundwork for specific support in teacher education. The possibility of organizing planning aspects into empirically based clusters, like strategically, pedagogically fragile, and transitional, represents a new arena for understanding the way instructional logic develops within dimensions. Because there is little prior work in this type of pattern-based analysis, our findings bridge an important methodological gap and pave the way for the incorporation of AI into formative-based diagnostic feedback. his represents an important transition: from assessing separate competences to representing pedagogical reasoning as a dynamic, culturally grounded, and inspectable entity—one that embodies the values, narratives, and ethical priorities of the communities in which teachers are embedded.

For further research and practice, this study provides an extensible model that can be applied within other transnational or intercultural dimensions of teacher education. It also challenges other scholars to investigate the impact of identity, language, and sociocultural background on how pre-service teachers take up and negotiate evaluative standards. In the end, by positioning metaevaluation as not just a procedure but as a formative and identity-influencing act, we can better ready teachers to make sense of the many paths to calibration that encourage them to take on the complexities of today's differently minded terrain with intention, clarity, and honesty.

References

Bórquez, M., Garrido, M., & Flores, P. (2020). Contextual coherence in initial teacher planning: Challenges in Latin America. *Revista Latinoamericana de Educación*, 52(3), 123–145.

Brinkmann, S., & Kvale, S. (2018). *Doing Interviews*. SAGE Publications.

Dennis, M., Valenzuela, C., & Bravo, S. (2023). Inclusive planning and curricu-lum adaptation in teacher education. *International Journal of Educational Devel-opment*, 96, 102674.

Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2011). *Program Evaluation: Alternative Approaches

and Practical Guidelines* (4th ed.). Pearson.
Fullan, M., & Quinn, J. (2016). *Coherence: The Right Drivers in Action for Schools, Districts, and Systems*. Corwin Press.
Lozano, V. (2023). Evaluative alignment and inclusion in novice teacher plan-ning. *Teaching and Teacher Education*, 118, 103874.
Scriven, M. (1991). *Evaluation Thesaurus* (4th ed.). SAGE Publications.
Stufflebeam, D. L., & Shinkfield, A. J. (2007). *Evaluation Theory, Models, and Applications*. Jossey-Bass.
Worthen, B. R., & Sanders, J. R. (2006). *Educational Evaluation: Alternative Approaches and Guidelines*. Longman.
Zhao, Y., Chan, T., & Huang, R. (2022). AI-based modeling of teacher planning behavior: A review of practices. *Computers & Education: Artificial Intelli-gence*, 3, 100068.
Biesta, G. (2010). *Good Education in an Age of Measurement: Ethics, Politics, Democracy*. Routledge.