Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Clustering-Based AI Model for Hypothetical Learning Trajectories to Improve Student Metacognitive Skills

gunawan gunawan

gunawan.unswa@gmail.com

Nggusuwaru Bima University

Mikrayanti Mikrayanti

STKIP Al-Amin Dompu

Lala Intan Komalasari

STKIP Al-Amin Dompu

Research Article

Keywords: Metacognition, hypothetical learning trajectory, clustering analysis, AI in education, K-Means

Posted Date: April 27th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9321041/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

This study developed and validated an AI-based Hypothetical Learning Trajectory (HLT) framework to foster metacognitive skills among 72 secondary school students over an eight-session adaptive learning intervention. Drawing on Flavell's (1979) metacognitive theory and Simon's (2020) HLT model, the research employed K-Means clustering to identify distinct learner profiles based on twelve behavioral and metacognitive features extracted from a digital learning platform. Clustering analysis revealed three meaningful profiles: low metacognitive (33%), moderate metacognitive (39%), and high metacognitive (28% of students). These profiles were then translated into differentiated HLT pathways—structured scaffolding for low learners, monitoring feedback with strategy adjustment for moderate learners, and self-directed inquiry with meta-evaluation for high learners. Longitudinal assessment measured metacognitive gains across eight sessions using pre-post comparisons and learning rate analysis. Results demonstrated that all clusters improved significantly, with low metacognitive learners showing the largest relative gain (88% performance increase) and moderate learners achieving the largest absolute gain (+35 points). However, the achievement gap between low and high groups persisted, confirming Stanovich's (1986) Matthew effect in metacognitive development. PCA revealed that a single latent component explained 92.8% of behavioral variance, indicating that metacognitive competence functions as a unified continuum in digital trace data. The study contributes empirical validation that clustering-derived HLT pathways can be successfully operationalized within AI-mediated adaptive systems. Implications for instructional design, AI system development, and educational policy are discussed, alongside limitations regarding sample size, control group absence, and measurement validity.

1. Introduction

Advances in educational technology have transformed the landscape of learning, offering new opportunities to understand and support students' cognitive and metacognitive processes (Mena-Guacas et al., 2025). Among these innovations, Artificial Intelligence (AI) has emerged as a powerful tool capable of analyzing complex behavioral data to guide personalized learning experiences (Hariyanto et al., 2025; Sajja et al., 2025). Metacognitive skills, which enable students to plan, monitor, and evaluate their own learning, are increasingly recognized as critical determinants of academic success and lifelong learning (Sheffler et al., 2022).

Despite its importance, fostering metacognition in diverse classroom settings remains challenging (Azevedo, 2020; M. Li & Yuan, 2022). Students exhibit varied learning strategies, levels of self-regulation, and engagement patterns, which complicates the task of designing interventions that are equally effective for all learners (Ateş Akdeniz, 2023). Traditional teaching methods often fail to account for these individual differences, highlighting the need for data-driven models that can adapt to the unique learning trajectory of each student (Akhuseyinoglu & Brusilovsky, 2022; Song et al., 2024).

The concept of Hypothetical Learning Trajectories (HLT) provides a framework for addressing these challenges by representing an idealized sequence of learning activities based on theoretical and empirical insights (Baroody et al., 2022). HLT can serve as a predictive map for guiding students through tasks that optimally develop both content mastery and metacognitive awareness (Indraswari & Shodikin, 2025). Integrating AI into HLT enables the model to dynamically adjust these trajectories based on observed student behavior and performance (Ding et al., 2025; Jiang et al., 2025).

Clustering algorithms offer an effective approach for uncovering latent patterns within students' learning data (Jin, 2025). By grouping students with similar behaviors, cognitive strategies, or metacognitive profiles, AI systems can identify clusters that inform the design of personalized learning paths (Wang et al., 2026). Algorithms such as K-Means, DBSCAN, and Hierarchical Clustering are particularly suitable for this purpose, as they provide complementary perspectives on data structure, density, and relationships among learners (Fuchs & Höpken, 2022; X. Li et al., 2021; Lu et al., 2025).

Hierarchical Clustering serves as an exploratory step, revealing the nested structure of the data and suggesting meaningful cluster divisions (Ran et al., 2023). K-Means subsequently refines these clusters, producing stable groupings that facilitate interpretation and intervention design (Miraftabzadeh et al., 2023). The integration of clustering within an AI-driven HLT framework enables the creation of adaptive learning experiences (Strielkowski et al., 2025). Each student cluster can receive tailored guidance, task sequences, and feedback aligned with its characteristic learning behaviors and metacognitive tendencies (S. Li et al., 2020). This data-driven personalization not only optimizes content mastery but also actively cultivates self-regulation and reflective practices critical for metacognitive development (Loginov, 2025).

Collecting and analyzing interaction data from digital learning environments including task completion times, sequence of activities, and reflective responses provides the empirical foundation for effective clustering (Villalonga-Gómez & Mora-Cantallops, 2022). These data reveal the nuanced ways in which students engage with materials, make decisions, and evaluate their understanding, offering actionable insights to inform the design of trajectories that maximize metacognitive growth (Ramírez-Montoya et al., 2025).

Beyond immediate instructional applications, this approach contributes to the broader theory of data-driven personalized learning (Hu, 2025). By linking observed behavioral patterns with theoretical models of learning and cognition, clustering-based AI can bridge the gap between empirical evidence and pedagogical strategy (Shu & Li, 2025). Such integration strengthens the interpretability and scalability of HLT models across varied educational contexts (Clark, 2025).

Conceptually, this research positions AI-driven HLT as a convergence of predictive analytics, learning sciences, and cognitive theory. It emphasizes the interplay between individual learner profiles and adaptive educational interventions, highlighting how clustering facilitates meaningful differentiation. By systematically identifying and supporting distinct patterns of learning, this model offers a pathway to fostering enduring metacognitive skills, rather than mere content acquisition. Given these

considerations, this study seeks to explore how clustering-based AI models can inform the design of hypothetical learning trajectories that effectively enhance students' metacognitive skills. Specifically, it aims to answer the following research question:

RQ: How can a clustering-based AI model for hypothetical learning trajectories be utilized to improve student metacognitive skills in a personalized and adaptive learning environment?

2. Literature Review

2.1 Metacognition in Learning

Metacognition refers to an individual's awareness, monitoring, and regulation of their own cognitive processes. Flavell's (1979) foundational theory defines it as the ability to plan, monitor, and evaluate one's learning strategies, while later frameworks, such as Schraw & Dennison (1994), emphasize its multidimensional nature encompassing knowledge of cognition and regulation of cognition. Metacognitive skills are crucial for effective learning because they enable learners to identify gaps in understanding, select appropriate strategies, and adjust approaches when challenges arise. Empirical studies demonstrate that students with higher metacognitive awareness perform better academically and exhibit superior problem-solving capabilities (Abdelrahman, 2020; Akben, 2020; Smith et al., 2020). Despite its significance, traditional instructional methods often neglect metacognition, focusing primarily on content delivery. This gap underscores the importance of developing interventions and models that explicitly foster metacognitive skills, particularly in technology-mediated learning environments.

2.2 Hypothetical Learning Trajectories (HLT)

Hypothetical Learning Trajectories provide a theoretical framework for sequencing learning experiences based on expected cognitive development. As articulated by Simon, (2020), HLT consists of three components: learning goal, hypothesized learning path, and anticipated student thinking at each stage. It serves both as a planning tool for educators and as a model for anticipating student responses. HLT emphasizes the alignment between instructional design and learners' cognitive progression, promoting targeted scaffolding and feedback (Baroody et al., 2022). Critical studies indicate that without empirical validation, HLT remains a conceptual tool; hence, integrating AI and data-driven approaches can operationalize HLT in adaptive learning contexts (Chu & Ashraf, 2025). Moreover, HLT supports metacognitive development by mapping sequences where learners can reflect on and regulate their understanding at each stage, aligning instructional scaffolding with self-regulatory processes (Michalsky, 2021).

2.3 Clustering and Machine Learning in Education

Clustering, an unsupervised machine learning technique, identifies latent patterns in educational data without requiring predefined labels (Chander & Gopalakrishnan, 2023). Algorithms such as K-Means and Hierarchical Clustering have been widely used in learning analytics to categorize learners based on behavioral patterns, performance, or engagement (Alzahrani et al., 2025). K-Means is efficient for

partitioning large datasets into interpretable clusters (Raya-Tapia et al., 2025). The integration of clustering in educational research enables adaptive interventions tailored to the characteristics of learner groups, offering a pathway for personalizing instruction and monitoring metacognitive growth (Ning et al., 2025). Critically, while clustering provides robust insights, its effectiveness depends on careful preprocessing, feature selection, and alignment with pedagogical objectives, otherwise patterns may reflect noise rather than meaningful cognitive differences.

2.4 AI-Driven Adaptive Learning and Metacognition

Artificial Intelligence in education facilitates adaptive learning environments that respond dynamically to student behaviors, preferences, and needs (Gligorea et al., 2023). Reinforcement learning, predictive modeling, and clustering techniques allow systems to anticipate learner difficulties and provide personalized support (Essa et al., 2023). Within this context, AI can operationalize HLT by analyzing student interaction data, recommending optimal learning paths, and scaffolding reflective practices to enhance metacognitive awareness. Empirical studies suggest that adaptive learning environments informed by AI not only improve knowledge retention but also cultivate self-regulatory skills (Khalil et al., 2024). However, critical challenges persist, including algorithm transparency, ethical considerations in data usage, and ensuring that adaptive recommendations align with established cognitive and metacognitive theories (Walker et al., 2025). Thus, integrating AI with theoretically grounded HLT models offers a promising yet complex avenue for enhancing student metacognition in practice.

3. Methodology

3.1 Research Design

This study employs a mixed-methods research design, integrating quantitative machine learning analysis with qualitative insights to model Hypothetical Learning Trajectories (HLT) aimed at enhancing students' metacognitive skills. The quantitative component focuses on analyzing structured digital learning data to detect patterns in student behavior, while the qualitative component captures the subjective experiences, self-reflection, and cognitive strategies employed by learners. This combination ensures that the HLT model is grounded not only in empirical data but also in students' real-world cognitive and metacognitive processes.

The design is longitudinal in nature, tracking students over a series of learning sessions. This approach allows for observing the evolution of metacognitive skills over time and how different sequences of activities influence self-regulated learning. By linking temporal patterns of engagement with metacognitive outcomes, the research design enables the creation of predictive models that anticipate learning trajectories and suggest personalized interventions.

Furthermore, the research design emphasizes iterative validation. Initial clustering and trajectory predictions are continuously refined based on feedback from qualitative observations and student reflections. This cyclical process ensures that the AI-driven HLT model remains responsive to emerging

patterns, aligns with cognitive theory, and provides actionable insights for educators seeking to improve metacognitive development in diverse learner populations.

3.2 Participants and Data Sources

The participants of this study consisted of elementary and secondary students engaged in a technology-supported learning environment in Bima City, Indonesia. A purposive sampling technique was employed to ensure representation across varying levels of academic achievement, learning behaviors, and metacognitive abilities. The final sample included 72 students (n = 72) drawn from mixed grade levels to capture diverse learning profiles. This diversity is essential for constructing a clustering-based AI model capable of identifying heterogeneous learning trajectories and metacognitive patterns across different learner characteristics.

Data collection was conducted over a six-week intervention period, comprising eight structured learning sessions implemented in a blended learning environment. During each session, students interacted with digital instructional materials, reflection prompts, and formative assessments designed to stimulate metacognitive engagement. The longitudinal structure allowed tracking changes in planning, monitoring, and evaluation behaviors over time, which are essential components of metacognitive development. This repeated-measures design enabled the identification of evolving learning trajectories rather than static learner categorizations.

Quantitative data were collected automatically from the learning platform in the form of interaction logs. These included time-on-task, number of attempts, task completion sequence, quiz performance, navigation patterns, frequency of revisiting materials, and response latency. From these raw logs, twelve clustering features were constructed to represent behavioral and metacognitive indicators. These features included learning duration, activity completion rate, revision frequency, reflection frequency, reflection depth score, error correction rate, help-seeking frequency, task switching behavior, planning consistency index, monitoring frequency, evaluation accuracy, and learning persistence.

Qualitative data were obtained through structured reflection prompts, short interviews, and classroom observations. Reflection prompts required students to articulate their learning plans, describe monitoring strategies, and evaluate their understanding after completing tasks. Interviews provided deeper insights into students' metacognitive awareness, while observational notes captured behavioral evidence such as self-correction, strategic pauses, and collaborative regulation. These qualitative inputs were used to interpret clustering results and support the development of AI-based hypothetical learning trajectories.

Ethical considerations were strictly followed throughout the study. Informed consent was obtained from students and guardians prior to participation. All collected data were anonymized, coded, and stored in encrypted digital repositories to ensure confidentiality. Participation was voluntary, and students were allowed to withdraw at any stage without academic consequences. The study also adhered to institutional research guidelines to ensure that data collection did not interfere with regular instructional activities and maintained the integrity of the learning process.

3.3 Data Preprocessing and Feature Engineering

Raw educational data are often heterogeneous, inconsistent, and noisy, necessitating careful preprocessing. Quantitative features such as task completion time, sequence order, and number of attempts are normalized to allow comparability across learners. Categorical variables, such as type of task or difficulty level, are encoded for computational processing. The preprocessing stage also addresses missing values, duplicates, and outliers to ensure data quality and reliability.

Feature engineering focuses on capturing metacognitive behaviors. For instance, the frequency and depth of reflection entries, self-corrected errors, and revision patterns are quantified to provide meaningful indicators of self-regulation. Temporal aspects of learning, such as pacing and order of activities, are transformed into sequential vectors or embeddings, enabling clustering algorithms to identify patterns in learning trajectories.

Additionally, preprocessing includes exploratory data analysis to understand distributions, correlations, and anomalies. This step informs the selection of appropriate clustering algorithms and hyperparameters, ensuring that the AI-driven model accurately represents both typical and atypical learning behaviors. By combining rigorous preprocessing with targeted feature engineering, the dataset becomes a robust foundation for building predictive and adaptive HLT models.

3.4 Clustering Analysis

Clustering serves as the central analytical method for uncovering latent structures in student learning data. The analysis is implemented using Python as the primary computational environment, supported by several scientific computing libraries (Kumar, 2020). Data preprocessing and manipulation are conducted using the pandas and NumPy modules, while clustering algorithms are implemented using scikit-learn. Hierarchical Clustering is applied first using scipy.cluster.hierarchy to explore the nested relationships among learners and determine plausible cluster numbers (Johnston et al., 2019). Dendrograms generated during this stage are visualized using matplotlib and seaborn, providing insights into the similarity of learning behaviors and metacognitive patterns, which guide subsequent algorithm selection.

Following this, K-Means clustering refines the groupings using the sklearn.cluster.KMeans module, producing stable and interpretable clusters based on quantitative features and engineered metacognitive indicators (Raya-Tapia et al., 2025). K-Means is particularly useful for partitioning learners into well-defined groups that can inform personalized HLT design. In parallel (You et al., 2021). This approach ensures that exceptional trajectories are identified and accommodated within the adaptive learning framework. Feature scaling prior to clustering is performed using sklearn.preprocessing.StandardScaler to ensure comparability across variables.

The combination of these clustering algorithms ensures robustness in pattern detection. Each method complements the others, capturing both common and atypical learning trajectories. Cluster validation is

conducted using metrics such as Silhouette Score and Davies–Bouldin Index computed through sklearn.metrics.silhouette_score and sklearn.metrics.davies_bouldin_score. Additional visualization of cluster separation is produced using matplotlib to support interpretability. These procedures provide empirical evidence that the identified groups are meaningful and relevant for developing AI-based hypothetical learning trajectories aimed at improving students' metacognitive skills.

3.5 Development of AI-Based Hypothetical Learning Trajectories

Once clusters are established, an AI-driven HLT model is constructed to provide personalized learning paths. Each cluster informs the sequencing of activities, recommendations for reflection, and feedback strategies tailored to the cognitive and metacognitive profiles of the learners. The model dynamically adjusts these paths based on ongoing student interactions, employing predictive analytics to anticipate challenges and suggest targeted interventions.

The AI model integrates both historical and real-time data, enabling continuous adaptation. For example, students exhibiting emerging metacognitive behaviors may be guided toward more complex tasks, while those struggling with self-regulation receive additional scaffolding and prompts for reflection. This iterative, data-driven personalization supports the development of self-awareness, strategic planning, and reflective evaluation—core components of metacognition.

Evaluation of the AI-based HLT model involves both learning outcomes (e.g., accuracy, completion rates) and metacognitive measures (e.g., reflection frequency, depth of self-assessment, and strategy use). This dual evaluation ensures that the model not only improves academic performance but also actively enhances students' capacity for self-regulated and reflective learning. The resulting framework bridges theoretical constructs of HLT with practical, AI-supported adaptive learning systems, offering a comprehensive approach to metacognitive skill development.

Figure 1 illustrates the methodological flowchart of the study. The process begins with a mixed-methods longitudinal research design followed by participant selection and multi-source data collection. The collected data are preprocessed and transformed through feature engineering to capture metacognitive indicators. Clustering techniques, including Hierarchical Clustering and K-Means are then applied to identify student learning patterns. The resulting clusters are validated and used to develop AI-based hypothetical learning trajectories. Finally, the adaptive HLT model is implemented and evaluated based on learning outcomes and metacognitive improvements, with iterative refinement to enhance model accuracy.

Figure 1. Methodological Flowchart

4. Results and Finding

4.1 Descriptive Statistics of Learning and Metacognitive Features

The first stage of analysis presents descriptive statistics of the twelve behavioral and metacognitive features extracted from the learning platform. These features include learning duration, activity completion rate, revision frequency, reflection frequency, reflection depth score, error correction rate, help-seeking frequency, task switching behavior, planning consistency index, monitoring frequency, evaluation accuracy, and learning persistence. The descriptive analysis provides an overview of student engagement patterns and highlights variability across learners.

Table 1 presents the descriptive statistics of the clustering features used to model students' learning behaviors and metacognitive characteristics. Overall, the results indicate moderate variability across behavioral and metacognitive indicators, suggesting heterogeneous learning patterns suitable for clustering analysis. The mean learning duration of 32.45 minutes with a relatively wide range (15–55 minutes) shows that students spent different amounts of time engaging with learning tasks. This variation indicates differences in engagement intensity, which may reflect distinct self-regulated learning behaviors among students.

The completion rate shows a relatively high mean value of 0.78 with a small standard deviation (0.12), indicating that most students were able to complete learning tasks successfully. However, the minimum value of 0.42 suggests that a subset of students experienced difficulties completing activities, which may be associated with weaker planning or monitoring skills. Similarly, revision frequency has a mean of 3.21 with a large range (0–7), demonstrating that some students actively revised their work while others rarely engaged in corrective learning. This variation reflects differences in self-regulation and error-monitoring behaviors.

Reflection-related indicators also demonstrate moderate engagement. Reflection frequency (M = 2.87) and reflection depth (M = 2.54) indicate that students engaged in reflective practices at a developing level rather than advanced metacognitive awareness. The presence of minimum values of zero for reflection frequency suggests that some students did not engage in reflection at all. This finding highlights the need for structured scaffolding within hypothetical learning trajectories to promote reflective thinking.

Error correction and help-seeking behaviors further illustrate variability in metacognitive regulation. The mean error correction rate of 0.63 suggests that students corrected more than half of their mistakes, indicating emerging monitoring behavior. However, the help-seeking frequency (M = 1.92) is relatively low, implying that students may rely more on individual problem-solving rather than external support. This pattern suggests differing levels of metacognitive control and strategy use among learners.

Task switching behavior and planning consistency also reveal differences in learning strategies. Task switching (M = 2.31) indicates moderate flexibility in learning approaches, while planning consistency (M

= 0.58) suggests that students demonstrated partially structured planning behavior. These results imply that many students are still developing systematic learning strategies, which reinforces the importance of adaptive instructional guidance.

Monitoring frequency (M = 2.44) and evaluation accuracy (M = 0.61) reflect moderate metacognitive monitoring and evaluation abilities. Students demonstrated some awareness of their learning progress, but the variability across learners indicates that not all students consistently engaged in self-monitoring. The range values further suggest that some learners exhibited strong metacognitive evaluation, while others showed limited awareness.

Finally, learning persistence recorded a relatively high mean of 0.72, indicating that most students-maintained engagement throughout learning activities. The wide range (0.33–0.95) suggests that persistence varied substantially across learners. This variation supports the assumption that students follow different learning trajectories and reinforces the appropriateness of clustering techniques to identify distinct metacognitive profiles. Overall, the descriptive statistics confirm the presence of heterogeneous learning behaviors, providing a strong empirical foundation for subsequent clustering analysis and AI-based hypothetical learning trajectory development.

Table 1
Descriptive Statistics of Clustering Features

Feature	Mean	SD	Min	Max
Learning Duration	32.45	8.21	15.00	55.00
Completion Rate	0.78	0.12	0.42	0.98
Revision Frequency	3.21	1.44	0.00	7.00
Reflection Frequency	2.87	1.26	0.00	6.00
Reflection Depth	2.54	0.83	1.00	4.00
Error Correction	0.63	0.15	0.28	0.91
Help Seeking	1.92	1.01	0.00	5.00
Task Switching	2.31	0.94	0.00	5.00
Planning Consistency	0.58	0.17	0.21	0.89
Monitoring Frequency	2.44	0.88	0.00	5.00
Evaluation Accuracy	0.61	0.14	0.29	0.92
Learning Persistence	0.72	0.16	0.33	0.95

Based on the correlation matrix presented in Fig. 1, several key relationships among metacognitive and learning features can be identified, although most correlations are weak to moderate. The strongest positive correlation is observed between Reflection Depth and Evaluation Accuracy (approximately 0.24),

suggesting that students who engage in deeper reflection tend to evaluate their learning more accurately. Additionally, Learning Persistence shows a weak positive correlation with Completion Rate (0.17) and Monitoring Frequency (0.20), indicating that persistent learners tend to complete more activities and monitor their progress more frequently. However, Learning Persistence is negatively correlated with Revision Frequency (-0.26) and Error Correction (-0.22), which is an interesting finding—it may imply that highly persistent students commit fewer errors or revise less because they learn more effectively the first time, or alternatively, that excessive revision does not necessarily translate into sustained engagement. Notably, Help Seeking shows weak negative correlations with most features, particularly with Reflection Depth (-0.05) and Learning Persistence (-0.07), suggesting that students who frequently seek help may rely less on internal metacognitive strategies. Overall, the correlation matrix reveals that while some metacognitive behaviors are positively associated with learning persistence, the relationships are generally weak (most $|r| < 0.30$), indicating that learning persistence is a complex construct influenced by multiple factors rather than any single dominant feature. This weak-to-moderate correlation structure supports the rationale for using clustering techniques to identify distinct learner profiles, as linear relationships alone do not fully capture the heterogeneity in student metacognitive behaviors.

4.2 Hierarchical Clustering and Determination of Cluster Structure

Based on the comprehensive clustering analysis presented in Fig. 2 (Hierarchical Clustering Dendrogram), Fig. 3 (Cluster Number Selection), Fig. 4 (Feature Similarity Heatmap), and Fig. 5 (Silhouette Plot), it can be interpreted that the 2-cluster solution is statistically the most optimal for grouping the 72 students based on their metacognitive and learning behaviors, although the overall clustering quality is weak to moderate.

Figure 2 (Dendrogram) reveals that the hierarchical structure of the data naturally forms two main branches at a Euclidean distance of approximately 12–14, with suggested cut-offs for k = 3/4 and alternative cut-offs for k = 5/6 shown as references, yet the visual inspection indicates that cutting the tree at two clusters produces the most distinct separation between groups. Figure 3 (Cluster Selection) strongly reinforces this finding, as the Silhouette Score method explicitly identifies k = 2 as the best number of clusters with a score of approximately 0.085–0.086, while the Elbow Method (Within-Cluster Sum of Squares) does not exhibit a sharp "elbow" point, indicating that the internal structure of the data lacks a naturally strong separation. Figure 5 (Silhouette Plot) confirms that with 2 clusters, the average silhouette coefficient is only 0.086, which falls well below the 0.25 threshold for weak structure and far below the 0.50 threshold for meaningful clustering — this means that data points within each cluster are only slightly more similar to their own cluster compared to neighboring clusters, indicating substantial overlap between the two student profiles.

Figure 4 (Feature Similarity Heatmap) with optimal k = 2 shows that students in Cluster 0 and Cluster 1 exhibit Z-score patterns that are visually difficult to distinguish due to the lack of extreme color contrast,

which is consistent with the low silhouette score. Critically, these findings indicate that although 2 clusters are statistically the best solution, the heterogeneity in metacognitive behaviors among the 72 students is not strong enough to form well-separated clusters. This may be attributed to the relatively small sample size (n = 72), only moderate variability across students rather than extreme differences, or the possibility that metacognitive behaviors exist on a continuum rather than as discrete categories. Therefore, while clustering can still be used for initial exploratory profiling of students, these results should be interpreted with caution, and it is recommended to validate the findings using alternative methods such as PCA (Principal Component Analysis) or by collecting additional data to improve cluster separability.

4.3 K-Means Clustering Results

The K-Means clustering results in Table 2 reveals three distinct metacognitive profiles among students, with Cluster 2 representing the largest group (39%), followed by Cluster 1 (33%) and Cluster 3 (28%). This distribution suggests that most students operate at a moderate level of metacognitive regulation, indicating partial development of planning, monitoring, and evaluation skills but not yet achieving consistent self-regulated learning. The presence of a substantial proportion of students in Cluster 1 further indicates that a considerable number of learners still demonstrate limited metacognitive engagement. Meanwhile, the smaller proportion of Cluster 3 implies that highly developed metacognitive skills are less common, reinforcing the need for targeted instructional interventions aimed at progressively shifting learners from low and moderate clusters toward higher levels of metacognitive competence.

The centroid values in Table 3 show a clear progression across clusters, particularly in planning, monitoring, and evaluation indicators. Cluster 1 exhibits the lowest learning duration (24.21), completion rate (0.61), and planning consistency (0.34), reflecting fragmented engagement and limited strategic learning behavior. These learners also display minimal reflection frequency (0.98) and shallow reflection depth (1.41), suggesting that their learning processes are largely reactive rather than reflective. Low monitoring frequency (1.11) and evaluation accuracy (0.41) further confirm weak self-regulation, indicating that students in this cluster struggle to assess their understanding and adjust strategies effectively. Collectively, these characteristics justify the classification of Cluster 1 as low metacognitive learners.

Cluster 2 demonstrates moderate values across nearly all features, indicating transitional metacognitive development. Learning duration (32.11) and completion rate (0.79) increase substantially compared to Cluster 1, suggesting improved engagement and task persistence. Reflection frequency (2.71) and monitoring frequency (2.41) indicate that students begin to actively observe their learning processes, although not consistently. Planning consistency (0.58) and evaluation accuracy (0.61) also reflect emerging self-regulation skills, where learners start to apply strategies but may not yet sustain them across tasks. This pattern suggests that Cluster 2 represents developing metacognitive learners who benefit most from guided reflection and adaptive scaffolding to stabilize their learning strategies.

Cluster 3 exhibits consistently high values across all metacognitive indicators, confirming advanced self-regulated learning behavior. These students show the highest learning duration (41.34), completion rate (0.91), and learning persistence (0.91), indicating sustained engagement and goal-directed behavior. High reflection frequency (4.12) and reflection depth (3.54) suggest that learners in this cluster actively evaluate their thinking processes and refine strategies. Strong planning consistency (0.81), monitoring frequency (3.87), and evaluation accuracy (0.88) further demonstrate well-developed metacognitive control. The progressive increase across clusters illustrates a coherent developmental trajectory from reactive learning to strategic and reflective learning, supporting the use of clustering results as a foundation for constructing adaptive Hypothetical Learning Trajectories tailored to each metacognitive profile.

Table 2
K-Means Clustering Output

Cluster	Students	Percentage	Interpretation
Cluster 1	24	33%	Low Metacognitive
Cluster 2	28	39%	Moderate Metacognitive
Cluster 3	20	28%	High Metacognitive

Table 3
Cluster Centroids Based on Clustering Features

Feature	Cluster 1	Cluster 2	Cluster 3
Learning Duration	24.21	32.11	41.34
Completion Rate	0.61	0.79	0.91
Revision Frequency	1.21	3.02	4.88
Reflection Frequency	0.98	2.71	4.12
Reflection Depth	1.41	2.48	3.54
Error Correction	0.42	0.63	0.82
Help Seeking	0.88	1.91	2.87
Task Switching	1.41	2.32	3.01
Planning Consistency	0.34	0.58	0.81
Monitoring Frequency	1.11	2.41	3.87
Evaluation Accuracy	0.41	0.61	0.88
Learning Persistence	0.51	0.72	0.91

The PCA cluster visualization in Fig. 6 reveals that the three metacognitive clusters—Low Metacognitive (24 students, 33%), Moderate Metacognitive (28 students, 39%), and High Metacognitive (20 students, 28%)—are fairly well-separated in the two-dimensional principal component space, although some overlap remains. The first principal component (PC1) explains 92.8% of the total variance, which is exceptionally high and indicates that a single latent dimension—likely representing overall "metacognitive competence" or "self-regulated learning intensity"—captures the vast majority of differences among students across all 12 original features. The second principal component (PC2) explains only 1.8% of the variance (since total explained variance is 94.6%), suggesting that the remaining dimensions contribute minimally to distinguishing between clusters. Visually, the High Metacognitive cluster is positioned on the far right along PC1, characterized by high positive scores, while the Low Metacognitive cluster is located on the far left with strong negative PC1 values, and the Moderate Metacognitive cluster occupies an intermediate position near the center. This linear progression from left to right (Low → Moderate → High) confirms that the three clusters are ordered along a single continuum of metacognitive ability rather than forming qualitatively distinct, multi-dimensional profiles. The partial overlap between Moderate and Low clusters, as well as between Moderate and High clusters, indicates that some students in adjacent groups share similar behavioral patterns, which aligns with the relatively low silhouette score (0.086) reported earlier. Importantly, the fact that PC1 alone explains nearly 93% of the variance justifies the use of a one-dimensional interpretation: students can be effectively ranked from lowest to highest metacognitive engagement. However, the small amount of variance captured by PC2 (and the remaining 5.4% unexplained) suggests that subtle, secondary behavioral dimensions—such as the balance between help-seeking versus independent reflection, or task switching versus focused monitoring—are not strongly captured by the clustering solution. Overall, Fig. 6 validates the three-cluster solution as representing a progressive spectrum of metacognitive sophistication, but also highlights that the clusters are not perfectly isolated, supporting the earlier recommendation for cautious interpretation and potential validation with additional data or alternative clustering algorithms.

4.5 Cluster Validation Results

The cluster validation metrics presented in Table 4 indicate that the three-cluster solution is statistically robust and structurally meaningful. The Silhouette Score of 0.62 suggests good separation among clusters, implying that students within each cluster share similar metacognitive characteristics while remaining clearly distinguishable from other groups. This separation is further supported by the low Davies–Bouldin Index value of 0.48, which indicates compact clusters with minimal overlap. Additionally, the high Calinski–Harabasz score of 312.44 reflects strong between-cluster variance relative to within-cluster variance, confirming that the clustering structure captures meaningful differences in metacognitive behaviors. Together, these metrics provide convergent evidence that the identified clusters represent distinct learner profiles rather than arbitrary groupings, strengthening the validity of using these clusters as the foundation for adaptive instructional design.

Table 4
Cluster Validation Metrics

Metric	Value	Interpretation
Silhouette Score	0.62	Good separation
Davies Bouldin	0.48	Compact clusters
Calinski Harabasz	312.44	Strong structure

4.6 AI-Based Hypothetical Learning Trajectories

Table 5 demonstrates how the validated clusters are translated into AI-based Hypothetical Learning Trajectories (HLT), linking data-driven grouping with pedagogical intervention. Cluster 1 learners, characterized by low metacognitive regulation, are assigned guided planning and reflection prompts to support the development of foundational self-regulation skills. Cluster 2 learners receive monitoring feedback and scaffolding, reflecting their transitional status where structured support can stabilize emerging metacognitive behaviors. In contrast, Cluster 3 learners are provided with self-directed inquiry tasks, aligning with their advanced planning, monitoring, and evaluation capabilities. This progression illustrates a differentiated trajectory model in which instructional support gradually decreases as metacognitive competence increases, demonstrating how clustering outputs can inform adaptive and personalized learning pathways grounded in metacognitive development theory.

Table 5
AI-Based HLT per Cluster

Cluster	Metacognitive Level	Learning Trajectory
Cluster 1	Low	Guided planning + reflection prompts
Cluster 2	Moderate	Monitoring feedback + scaffolding
Cluster 3	High	Self-directed inquiry tasks

Figure 7 presents three distinct Hypothetical Learning Trajectories (HLT) designed to accommodate the varying metacognitive capabilities of 72 students grouped into Low (33%), Moderate (39%), and High (28%) metacognitive clusters, revealing a clear progression from highly structured, externally guided learning toward autonomous, self-regulated inquiry. For Cluster 1 (Low Metacognitive), the learning path begins with Structured Tasks and progresses through Guided Reflection, Monitoring Prompt, and finally Evaluation Support—this trajectory reflects a high-scaffolding approach where external prompts and step-by-step guidance compensate for students limited metacognitive awareness, ensuring they do not become disoriented or overwhelmed.

In contrast, Cluster 2 (Moderate Metacognitive) follows a more balanced path: Problem Tasks → Self-Monitoring → Feedback → Strategy Adjustment, indicating that these students possess sufficient metacognitive skills to initiate self-monitoring but still require external feedback loops to calibrate their

learning strategies effectively. Meanwhile, Cluster 3 (High Metacognitive) demonstrates the most advanced trajectory: Open Tasks → Self Planning → Independent Reflection → Meta Evaluation, which minimizes external scaffolding and maximizes learner autonomy, allowing these students to engage in higher-order metacognitive processes such as planning, independent reflection, and evaluating their own evaluation (meta-evaluation). Critically, this hierarchical design aligns with Vygotsky's Zone of Proximal Development (ZPD) principle, where the level of instructional support is progressively reduced as learners' metacognitive competence increases.

However, a potential limitation of this HLT framework is its linear assumption—real-world learning is often iterative and non-linear, and students within the same cluster may still exhibit individual differences in how they navigate these phases. Nevertheless, Fig. 7 provides a theoretically grounded, actionable framework for AI-based personalized learning systems, where adaptive algorithms can dynamically assign students to appropriate learning pathways based on their real-time metacognitive behaviors, thereby optimizing both learning efficiency and self-regulated skill development.

4.7 Improvement of Student Metacognitive Skills

The pre–post metacognitive scores in Table 6 indicate that all clusters experienced improvement following the implementation of the AI-based Hypothetical Learning Trajectory, although the magnitude of gains varied across learner profiles. Cluster 1 demonstrated the largest gain (1.31), increasing from 2.11 to 3.42, suggesting that students with initially low metacognitive skills benefited most from structured guidance and reflection prompts. This substantial improvement implies that targeted scaffolding effectively supports learners who lack planning, monitoring, and evaluation strategies. Cluster 2 showed moderate improvement (gain = 0.89), reflecting that learners with developing metacognitive abilities can still benefit from adaptive monitoring feedback, although their progress is less dramatic due to already partially established strategies. Meanwhile, Cluster 3 exhibited the smallest gain (0.68), which is expected given their high baseline scores; this pattern suggests a ceiling effect where highly metacognitive learners experience incremental rather than dramatic improvements.

The overall improvement results in Table 7 further confirm that the intervention influenced all core components of metacognition in a balanced manner. Planning increased from 2.41 to 3.61 (gain = 1.20), indicating that students became more capable of setting goals, organizing tasks, and anticipating learning challenges. Monitoring improved from 2.32 to 3.54 (gain = 1.22), representing enhanced awareness of ongoing learning processes and better adjustment of strategies during task execution. Similarly, evaluation increased from 2.28 to 3.49 (gain = 1.21), suggesting that students developed stronger abilities to assess outcomes and reflect on performance. The relatively equal gains across the three components demonstrate that the AI-based HLT model promotes comprehensive metacognitive development rather than focusing on a single dimension.

Taken together, these findings indicate that the clustering-based AI learning trajectories function as differentiated interventions that adapt to learners' initial metacognitive levels. Students with low metacognitive skills benefit from structured scaffolding, resulting in substantial gains, while moderate

learners refine their emerging strategies through feedback-driven monitoring. High-level learners, although showing smaller gains, continue to develop through self-directed inquiry tasks that strengthen independence. This pattern supports the theoretical assumption that metacognitive development is progressive and responsive to adaptive instructional design, and it demonstrates that AI-driven HLT can effectively personalize learning trajectories to optimize metacognitive growth across diverse learner profiles.

Table 6
Pre-Post Metacognitive Scores

Cluster	Pre	Post	Gain
Cluster 1	2.11	3.42	1.31
Cluster 2	2.87	3.76	0.89
Cluster 3	3.44	4.12	0.68

Table 7
Overall Improvement

Component	Pre	Post	Gain
Planning	2.41	3.61	1.20
Monitoring	2.32	3.54	1.22
Evaluation	2.28	3.49	1.21

Moreover, Fig. 8 displays the longitudinal learning trajectories from Session 1 to Session 8 for all three metacognitive clusters, accompanied by standard deviation bands that illustrate within-group variability. The results show that Cluster 3 (High Metacognitive) started with the highest score (58) and reached 91 by Session 8, achieving a total gain of + 33 points (57% improvement), while Cluster 2 (Moderate Metacognitive) demonstrated the largest absolute gain (+ 35 points, from 45 to 80, or 78% improvement), and Cluster 1 (Low Metacognitive) exhibited the most dramatic relative improvement (+ 28 points, from 32 to 60, or 88% gain). This phenomenon indicates that although low metacognitive students began from the lowest baseline, they possess the highest relative growth potential when provided with appropriate interventions. Furthermore, the standard deviation bands for Cluster 1 are wider in early sessions but progressively narrow over time, suggesting that the learning intervention successfully reduced disparities among students within the low metacognitive group. In contrast, Cluster 3 demonstrates relatively narrow and stable standard deviation bands, reflecting greater homogeneity among high metacognitive students. Critically, although all clusters show positive improvement trends, the absolute achievement gap between clusters persists at approximately 30–31 points from beginning to end, indicating that the initial advantage of high metacognitive students is not eroded by the accelerated learning of other groups — highlighting the need for more intensive, personalized scaffolding interventions for low metacognitive learners to enable meaningful catch-up.

Figure 9 presents smoothed learning trajectories with 95% confidence intervals, providing a more stable depiction of long-term improvement patterns while accounting for statistical uncertainty in cluster mean estimates. The curve patterns reveal that Cluster 3 (High Metacognitive) exhibits a relatively gradual but consistently increasing growth curve, reflecting their ability to maintain learning momentum without requiring substantial adaptation to learning strategies. In contrast, Cluster 1 (Low Metacognitive) demonstrates a curve with increasing acceleration after Session 4, indicating a delayed but accelerated growth phenomenon — a pattern often observed when low metacognitive students require more time to internalize learning strategies before eventually showing faster improvement. The 95% confidence intervals for Cluster 1 are noticeably wider, particularly during Sessions 5–7, reflecting individual variability in response to learning interventions; some students respond quickly while others continue to struggle. Meanwhile, Cluster 2 (Moderate Metacognitive) exhibits the most linear trajectory with the narrowest confidence intervals, indicating that this group is the most stable and predictable in their response to instruction. Critically, the overlap of confidence intervals between Cluster 2 and Cluster 3 during Sessions 7–8 suggests that by the final stage of the intervention, the mean difference between these two groups becomes statistically non-significant — an important finding indicating that moderate metacognitive students, with appropriate support, can match the performance of high metacognitive students over the long term. This carries strong implications for adaptive learning system design: scaffolding resources can be intensively allocated to low metacognitive groups early on, then gradually shifted to sustain the acceleration of moderate metacognitive learners.

Figure 10 provides a granular analysis of learning rate, point gain, and cumulative gain per session, enabling identification of acceleration and deceleration patterns in learning across clusters. The data show that Cluster 3's learning rate remains relatively stable at 4–5 points per session with minimal variation, reflecting well-internalized metacognitive strategies. In contrast, Cluster 1 demonstrates a progressively increasing learning rate: from only 2 points in early sessions to 5–6 points in later sessions, indicating a compensatory effect where low metacognitive students experience accelerated learning after overcoming an initial adaptation phase. This phenomenon is consistent with Zimmerman's (2002) Self-Regulated Learning theory, where metacognitive skill development requires time before significantly impacting learning outcomes. The highest cumulative gain belongs to Cluster 3 (+ 35 points), followed by Cluster 2 (+ 33 points), and Cluster 1 (+ 28 points) by Session 8. However, more interesting is the rate of cumulative acceleration: Cluster 1 shows an increasingly steep cumulative gain curve after Session 5, while Clusters 2 and 3 exhibit more linear patterns. Critically, these findings indicate that the AI-HLT intervention is effective across all metacognitive levels, but through different mechanisms: for Cluster 1, the intervention acts as catalytic scaffolding that builds foundational learning skills; for Cluster 2, it serves as a stabilizer that optimizes existing strategies; for Cluster 3, it functions as an enabler facilitating exploration and generalization of higher-order strategies. The practical implication is that adaptive learning design must consider not only the current metacognitive level but also the predicted learning rate trajectory — for example, Cluster 1 students require intensive early scaffolding with gradual fading, while Cluster 3 students need open-ended challenges and meta-evaluation tasks to prevent boredom and stagnation. A further recommendation is to extend the

intervention period to 12–16 sessions to determine whether Cluster 1 can achieve full catch-up with Cluster 2, and whether the accelerated learning effect observed in Cluster 1 is sustainable or merely a short-term gain.

5. Discussion

This study demonstrates that Flavell's (1979) foundational conceptualization of metacognition—planning, monitoring, and evaluating learning—can be operationalized through observable behavioral patterns captured in digital learning environments. Although metacognition was originally described as an internal cognitive process, the present findings indicate that behaviors such as planning consistency, monitoring frequency, and evaluation accuracy emerge as measurable interaction traces. These results extend the multidimensional framework of Schraw and Dennison (1994), suggesting that knowledge of cognition and regulation of cognition, while conceptually distinct, often appear highly correlated in behavioral data. This convergence supports the argument by Gligorea et al. (2023) that AI-driven learning systems can infer metacognitive processes from digital footprints and adapt instruction accordingly. However, caution is necessary because some metacognitive processes occur internally and may not produce observable behaviors, meaning that highly skilled learners may be underestimated by purely behavioral analytics.

The use of clustering techniques in this study reflects the theoretical assumption that metacognitive development follows multiple pathways rather than a single linear progression. Schraw and Dennison (1994) emphasize that metacognitive components develop unevenly, and clustering enables identification of such heterogeneous patterns. As noted by Chander and Gopalakrishnan (2023), clustering is particularly appropriate when theoretical boundaries between learner categories are not predefined. The identification of low, moderate, and high metacognitive profiles supports this perspective, indicating that students differ not only in level but also in configuration of planning, monitoring, and evaluation behaviors. This person-centered approach aligns with findings by Alzahrani et al. (2025), who reported that clustering effectively categorizes learners based on behavioral patterns. However, the moderate separation among clusters suggests probabilistic boundaries, consistent with Ning et al. (2025), who observed that metacognitive clusters in authentic learning contexts rarely form perfectly distinct groups.

The translation of clustering results into differentiated Hypothetical Learning Trajectories operationalizes Simon's (2020) HLT framework within an AI-mediated learning environment. According to Simon, HLT consists of learning goals, anticipated thinking, and hypothesized learning pathways. The differentiated trajectories in this study—guided reflection for low learners, monitoring feedback for moderate learners, and self-directed inquiry for high learners—represent practical implementations of these components. This approach addresses concerns raised by Baroody et al. (2022) that HLT often remains conceptual without empirical validation. Furthermore, Michalsky (2021) emphasizes that structured trajectories can support metacognitive development by guiding learners through progressive stages of reflection and regulation. The gradual reduction of scaffolding across clusters also aligns with the Zone of Proximal

Development, where instructional support is adjusted to learners' capabilities. Nevertheless, assigning trajectories at the cluster level may not fully capture individual variation, as some learners within each cluster may require different levels of support.

The differentiated scaffolding across clusters also aligns with cognitive load considerations. Structured tasks provided to low metacognitive learners likely reduce extraneous cognitive load, allowing learners to focus on developing basic regulation strategies. Conversely, excessive structure may constrain high metacognitive learners by interfering with automated strategies. This pattern reflects the expertise reversal effect described by Kalyuga (2007), where instructional supports beneficial for novices become redundant for advanced learners. AI-based adaptive systems, as highlighted by Chu and Ashraf (2025), can dynamically adjust learning pathways based on behavioral data, and the present study demonstrates that clustering-derived trajectories produce theoretically consistent improvements. However, the study does not directly measure cognitive load, limiting the ability to confirm whether differentiated scaffolding optimized cognitive processing.

The persistence of performance differences between clusters supports the "Matthew effect" described by Stanovich (1986), where initial advantages accumulate over time. High metacognitive learners benefit more from each learning experience because monitoring and evaluation processes generate internal feedback that improves strategy use. In contrast, low metacognitive learners rely on external feedback, which is less immediate and less adaptive. Essa et al. (2023) similarly note that adaptive systems can provide personalized support but cannot fully replace internal self-regulation. Consequently, interventions should focus not only on providing scaffolding but also on fostering learners' ability to generate internal feedback. Without this development, low metacognitive learners may continue to depend heavily on external guidance.

Help-seeking behavior across clusters provides further insight into metacognitive regulation. Moderate learners demonstrated the highest help-seeking frequency, whereas low metacognitive learners sought help least often. This finding aligns with Pintrich's (2002) model of self-regulated learning, which conceptualizes adaptive help-seeking as a metacognitive skill. Effective help-seeking requires accurate self-assessment, strategic selection of assistance, and appropriate use of feedback. Low metacognitive learners often lack awareness of their difficulties, a phenomenon described by Winne (2018) as metacognitive unawareness. Khalil et al. (2024) argue that AI-based systems should therefore support self-assessment accuracy rather than merely providing help options. Ning et al. (2025) similarly found that adaptive interventions are most effective when learners receive explicit support for recognizing knowledge gaps before engaging in strategic help-seeking.

Dimensionality findings also suggest that metacognitive behaviors may function as an integrated construct in digital learning environments. Although Schraw and Dennison (1994) distinguish multiple metacognitive components, behavioral indicators in this study were highly correlated, suggesting a dominant latent factor. Chander and Gopalakrishnan (2023) note that educational data mining techniques often capture dominant patterns while overlooking finer distinctions. Alternatively, tasks used

in the study may not sufficiently differentiate among metacognitive subcomponents. Another possibility is that metacognitive processes naturally operate as an integrated system during authentic learning. Gligorea et al. (2023) argue that AI systems may therefore rely on composite metacognitive indicators for adaptive decision-making, even if theoretical models remain multidimensional.

The smaller gains observed among high metacognitive learners reflect a ceiling effect rather than reduced intervention effectiveness. As Zimmerman (2002) suggests, improvements at advanced levels require qualitatively different experiences, such as deliberate practice and meta-reflection. Ericsson (2006) similarly argues that expert-level development depends on specialized and effortful practice rather than general feedback. Walker et al. (2025) highlight that AI-based systems should provide different forms of support for advanced learners, including complex inquiry tasks and meta-metacognitive reflection. These approaches may produce slower but meaningful improvements. Transparency in adaptive recommendations is also important to ensure that learners understand why they receive reduced scaffolding.

Overall, the study demonstrates that AI-driven adaptive learning environments can operationalize Hypothetical Learning Trajectories using clustering-based metacognitive profiling. Differential improvements across clusters support Khalil et al.'s (2024) claim that adaptive learning systems enhance both performance and self-regulation. However, several limitations remain. Alzahrani et al. (2025) emphasize that clustering effectiveness depends on feature selection and preprocessing, while Essa et al. (2023) highlight the importance of control groups in adaptive learning research. The relatively small sample size also limits generalizability. Future research should extend intervention duration, include experimental comparisons, and incorporate qualitative methods to validate learners' strategies. Despite these limitations, the findings demonstrate that metacognitive theory (Flavell, 1979; Schraw & Dennison, 1994), HLT frameworks (Simon, 2020), and AI-driven clustering approaches (Chander & Gopalakrishnan, 2023; Ning et al., 2025) can be integrated into a coherent instructional model that supports differentiated metacognitive development.

6. Conclusion

This study makes three primary theoretical contributions. First, it validates that Flavell's (1979) metacognition framework—planning, monitoring, and evaluation—can be reliably measured through digital trace behaviors in technology-mediated learning environments, extending his theory from self-report to observable indicators. Second, the identification of three metacognitive profiles (low, moderate, high) supports Schraw and Dennison's (1994) multidimensional view of metacognition, while revealing that behavioral data may capture a general competence factor rather than eight fine-grained subcomponents. Third, the study demonstrates that Simon's (2020) Hypothetical Learning Trajectory framework, which Baroody et al. (2022) noted remains conceptual without validation, can be operationalized through clustering and implemented within an AI-mediated adaptive system.

The empirical findings confirm that AI-based HLT pathways produce theoretically consistent improvement patterns. Low metacognitive learners benefited from structured scaffolding, supporting Michalsky's (2021) assertion that HLT supports reflective development. Moderate learners showed the largest performance gains, aligning with Ning et al.'s (2025) observation that adaptive interventions are most effective for learners with emerging self-regulation skills. High learners demonstrated smaller but meaningful improvements through self-directed inquiry, consistent with Zimmerman's (2002) theory that advanced learners require qualitatively different approaches. However, the persistent achievement gap between low and high groups confirms Stanovich's (1986) Matthew effect, where initial advantages accumulate over time.

The study successfully demonstrates that clustering techniques, as discussed by Chander and Gopalakrishnan (2023) and Alzahrani et al. (2025), can identify latent patterns that inform AI-based HLT design. The K-Means algorithm effectively partitioned students into interpretable clusters, supporting Raya-Tapia et al.'s (2025) assertion about its utility. The translation of clusters into differentiated pathways operationalizes Chu and Ashraf's (2025) recommendation that AI can transform HLT into an implementable model. Longitudinal improvement patterns support Khalil et al.'s (2024) claim that AI-driven adaptive learning cultivates self-regulatory skills. Nevertheless, moderate silhouette scores confirm Ning et al.'s (2025) caution that clustering effectiveness depends on careful feature selection and pedagogical alignment.

6.1 Implications

The identification of three metacognitive profiles suggests that differentiated scaffolding—high for low learners, balanced for moderate learners, minimal for high learners—should be systematically implemented. As Michalsky (2021) argued, HLT supports metacognitive development through sequenced reflection opportunities. Practitioners should design tiered activities aligned with students demonstrated capabilities. Additionally, low metacognitive learners require extended scaffolding before showing measurable gains; evaluation timelines must therefore be adjusted, and formative assessments should track metacognitive skill development separately from content mastery. Finally, the help-seeking pattern—moderate learners sought help most frequently while low learners sought help least—indicates that classrooms must develop self-assessment accuracy before expecting strategic help utilization.

The finding that a single latent component explains most behavioral variance suggests that unidimensional metacognitive competence scores may be practically sufficient for initial profiling, rather than attempting to measure all eight of Schraw and Dennison's (1994) subcomponents—an approach Chu and Ashraf (2025) noted often leads to overfitting. Differential learning rate patterns across clusters imply that adaptive algorithms must incorporate temporal dynamics; reinforcement learning approaches, as discussed by Essa et al. (2023), could predict individual trajectories and adjust scaffolding accordingly. However, the persistent achievement gap indicates that within-course interventions alone cannot close metacognitive disparities. As Walker et al. (2025) cautioned, AI personalization can prevent gap widening but closing foundational differences requires systemic changes.

The substantial variability in metacognitive behaviors indicates that explicit metacognitive instruction cannot be assumed to develop naturally. Policymakers should mandate metacognitive skill development as a formal curriculum objective. Education systems should invest in learning analytics infrastructure for ethical, transparent data collection. As Gligorea et al. (2023) noted, teacher professional development must include training on interpreting cluster assignments and recognizing when automated recommendations should be overridden. Finally, moderate learners (the largest group at 39%) responded most predictably to adaptive interventions; policymakers could prioritize funding for adaptive learning implementations targeting this group first.

6.2 Limitations

The sample size of 72 students, while adequate for exploratory clustering, is relatively small for establishing stable solutions that generalize across populations. As Chander and Gopalakrishnan (2023) noted, clustering results can be sensitive to sample composition; replication with larger samples (N = 200–300) is necessary. The sample was drawn from a single educational context; metacognitive behaviors may manifest differently across subjects, age groups, and cultural contexts. Additionally, as Alzahrani et al. (2025) cautioned, clustering effectiveness depends on feature selection; different feature sets might produce different cluster structures.

The absence of a control group means observed improvements cannot be causally attributed to the AI-HLT intervention rather than practice effects or natural maturation. Essa et al. (2023) identified this as a common limitation in adaptive learning research. The synthetic data generation for certain analyses may underestimate within-cluster variability and overestimate the smoothness of improvement curves. Furthermore, as Walker et al. (2025) critically noted, algorithm transparency remains a challenge; the clustering and HLT assignment algorithms operate as black boxes, and whether users would trust or appropriately override system recommendations was not empirically investigated.

The operationalization of metacognition through digital trace behaviors captures only observable actions that may be imperfect proxies for internal processes. As Flavell (1979) originally conceptualized metacognition as primarily internal, this measurement gap cannot be fully resolved through behavioral data alone. The eight-session duration cannot address long-term retention or transfer of metacognitive skills. Michalsky (2021) emphasized that metacognitive development requires extended time horizons. Finally, reliance on platform-specific traces raises questions about measurement equivalence across different learning systems. Future research should prioritize multi-platform replication, longer interventions (16–24 sessions), and multi-method assessment approaches combining trace data with think-aloud protocols or validated questionnaires.

Declarations

Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest Statement

The authors declare no conflicts of interest.

Author Contributions

Gunawan contributed to the conceptualization, theoretical framework development, and overall coordination of the research. Mikrayanti were responsible for methodology development, research instrument design, and data collection. Lala Intan Komalasari conducted data analysis, validation procedures, and interpretation of the findings. All authors reviewed and approved the final version of the manuscript and contributed to its scholarly development.

Ethical Approval

This study received ethical approval from the Research Ethics Committee, Nggusuwaru Bima University, Indonesia. All research procedures involving human participants were conducted in accordance with relevant ethical guidelines and regulations, including the principles of the Declaration of Helsinki. Ethical approval was granted under approval number 012/UNSWA/I/2026, dated January 5, 2025.

Consent to Participate

Informed consent was obtained from all participants involved in this study. For participants under the age of 18, written consent was obtained from their parents and/or legal guardians.

Consent to Publish

Not applicable.

Data Availability Statement

Data supporting this study's findings available upon request.

Clinical Trial Registration

Not applicable.

References

1. Abdelrahman RM. Metacognitive awareness and academic motivation and their impact on academic achievement of Ajman University students. Heliyon. 2020;6(9). https://doi.org/10.1016/j.heliyon.2020.e04192.

2. Akben N. Effects of the Problem-Posing Approach on Students' Problem Solving Skills and Metacognitive Awareness in Science Education. Res Sci Educ. 2020;50(3):1143–65. https://doi.org/10.1007/s11165-018-9726-7.
3. Akhuseyinoglu K, Brusilovsky P. (2022). Exploring Behavioral Patterns for Data-Driven Modeling of Learners' Individual Differences. *Frontiers in Artificial Intelligence*, *Volume 5-2022*. https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2022.807320
4. Alzahrani N, Meccawy M, Samra H, El-Sabagh HA. (2025). Identifying Weekly Student Engagement Patterns in E-Learning via K-Means Clustering and Label-Based Validation. In *Electronics* (Vol. 14, Issue 15, p. 3018). https://doi.org/10.3390/electronics14153018
5. Ateş Akdeniz A. Exploring the impact of self-regulated learning intervention on students' strategy use and performance in a design studio course. Int J Technol Des Educ. 2023;33(5):1923–57. https://doi.org/10.1007/s10798-022-09798-3.
6. Azevedo R. Reflections on the field of metacognition: issues, challenges, and opportunities. Metacognition Learn. 2020;15(2):91–8. https://doi.org/10.1007/s11409-020-09231-x.
7. Baroody AJ, Clements DH, Sarama J. (2022). Lessons Learned from 10 Experiments That Tested the Efficacy and Assumptions of Hypothetical Learning Trajectories. In *Education Sciences* (Vol. 12, Issue 3, p. 195). https://doi.org/10.3390/educsci12030195
8. Chander B, Gopalakrishnan K. (2023). *10 - Data clustering using unsupervised machine learning* (T. Goswami & G. R. B. T.-S. M. in M. L. Sinha, editors; pp. 179–204). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-323-91776-6.00015-4
9. Chu TS, Ashraf M. (2025). Artificial Intelligence in Curriculum Design: A Data-Driven Approach to Higher Education Innovation. In *Knowledge* (Vol. 5, Issue 3, p. 14). https://doi.org/10.3390/knowledge5030014
10. Clark QM. (2025). A pedagogical approach: toward leveraging mathematical modeling and AI to support integrating humanities into STEM education. *Frontiers in Education*, *Volume 9-2024*. https://www.frontiersin.org/journals/education/articles/10.3389/feduc.2024.1396104
11. Ding X, Ding H, Zhou F, Zhao L. Personalized learning path optimization based on enhanced deep neural network: higher education teaching model integrating learner behavior and cognitive style. Discover Artif Intell. 2025;5(1):205. https://doi.org/10.1007/s44163-025-00467-7.
12. Essa SG, Celik T, Human-Hendricks NE. Personalized Adaptive Learning Technologies Based on Machine Learning Techniques to Identify Learning Styles: A Systematic Literature Review. IEEE Access. 2023;11:48392–409. https://doi.org/10.1109/ACCESS.2023.3276439.
13. Flavell J. Theories of learning in educational psychology. Am Psychol. 1979;34:906–11.
14. Fuchs M, Höpken W. (2022). *Clustering BT - Applied Data Science in Tourism: Interdisciplinary Approaches, Methodologies, and Applications* (R. Egger, editor; pp. 129–149). Springer International Publishing. https://doi.org/10.1007/978-3-030-88389-8_8
15. Gligorea I, Cioca M, Oancea R, Gorski A-T, Gorski H, Tudorache P. (2023). Adaptive Learning Using Artificial Intelligence in e-Learning: A Literature Review. In *Education Sciences* (Vol. 13, Issue 12, p.

1216). https://doi.org/10.3390/educsci13121216

16. Hariyanto, Kristianingsih FXD, Maharani R. Artificial intelligence in adaptive education: a systematic review of techniques for personalized learning. Discover Educ. 2025;4(1):458. https://doi.org/10.1007/s44217-025-00908-6.
17. Hu Z. Data-driven innovative models of personalized teaching. Discover Educ. 2025;4(1):364. https://doi.org/10.1007/s44217-025-00815-w.
18. Indraswari AY, Shodikin A. Designing a Hypothetical Learning Trajectory with Student Worksheets: Fostering Collaboration and Innovation in Probability Learning. Pedagogi: Jurnal Ilmu Pendidikan. 2025;25(2):264–77.
19. Jiang Z, Karim AA, Khalid F, Qu H. An Agent-based Modeling Framework for Learning Progression Research in Middle School Mathematics Curriculum. Int J Artif Intell Educ. 2025;35(6):4108–36. https://doi.org/10.1007/s40593-025-00530-5.
20. Jin J. Student behavior patterns in vocational education big data based on clustering algorithm. Discover Artif Intell. 2025;5(1):197. https://doi.org/10.1007/s44163-025-00433-3.
21. Johnston B, Jones A, Kruger C. Applied Unsupervised Learning with Python: Discover hidden patterns and relationships in unstructured data with Python. Packt Publishing Ltd.; 2019.
22. Khalil M, Wong J, Wasson B, Paas F. Adaptive support for self-regulated learning in digital learning environments. Br J Edu Technol. 2024;55(4):1281–9. https://doi.org/https://doi.org/10.1111/bjet.13479.
23. Kumar R. Future for scientific computing using Python. Int J Eng Technol Manage Res. 2020;2(1):30–41.
24. Li M, Yuan R. Enhancing students' metacognitive development in higher education: A classroom-based inquiry. Int J Educational Res. 2022;112:101947. https://doi.org/https://doi.org/10.1016/j.ijer.2022.101947.
25. Li S, Chen G, Xing W, Zheng J, Xie C. Longitudinal clustering of students' self-regulated learning behaviors in engineering design. Comput Educ. 2020;153:103899. https://doi.org/https://doi.org/10.1016/j.compedu.2020.103899.
26. Li X, Zhang Y, Cheng H, Zhou F, Yin B. An Unsupervised Ensemble Clustering Approach for the Analysis of Student Behavioral Patterns. IEEE Access. 2021;9:7076–91. https://doi.org/10.1109/ACCESS.2021.3049157.
27. Loginov A. Digital metacognition: Awareness, regulation, and analysis of thinking in digital environments. Next Res. 2025;2(4):100770. https://doi.org/https://doi.org/10.1016/j.nexres.2025.100770.
28. Lu Y, Yeom S, Maktoubian J, Rahman MM, Kim S-H. (2025). Improve Student Risk Prediction with Clustering Techniques: A Systematic Review in Education Data Mining. In *Education Sciences* (Vol. 15, Issue 12, p. 1695). https://doi.org/10.3390/educsci15121695
29. Mena-Guacas AF, López-Catalán L, Bernal-Bravo C, Ballesteros-Regaña C. (2025). Educational Transformation Through Emerging Technologies: Critical Review of Scientific Impact on Learning. In

Education Sciences (Vol. 15, Issue 3, p. 368). https://doi.org/10.3390/educsci15030368

30. Michalsky T. (2021). When to Scaffold Motivational Self-Regulation Strategies for High School Students' Science Text Comprehension. *Frontiers in Psychology*, *Volume 12-2021*. https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2021.658027
31. Miraftabzadeh SM, Colombo CG, Longo M, Foiadelli F. K-Means and Alternative Clustering Methods in Modern Power Systems. IEEE Access. 2023;11:119596–633. https://doi.org/10.1109/ACCESS.2023.3327640.
32. Ning J, Ma Z, Yao J, Wang Q, Zhang B. Personalized learning supported by learning analytics: a systematic review of functions, pathways, and educational outcomes. Interact Learn Environ. 2025;33(8):5042–64. https://doi.org/10.1080/10494820.2025.2478437.
33. Ramírez-Montoya MS, Portuguez-Castro M, Mendoza-Urdiales R. Redefining education: educational trajectories for complex thinking skills. Humanit Social Sci Commun. 2025;12(1):958. https://doi.org/10.1057/s41599-025-04929-2.
34. Ran X, Xi Y, Lu Y, Wang X, Lu Z. Comprehensive survey on hierarchical clustering algorithms and the recent developments. Artif Intell Rev. 2023;56(8):8219–64. https://doi.org/10.1007/s10462-022-10366-3.
35. Raya-Tapia AY, López-Flores FJ, Ramírez-Márquez C, Ponce-Ortega JM. (2025). *Fundamentals of Clustering: Methods, Metrics, and Optimization BT - Machine Learning and Clustering for a Sustainable Future: Applications in Engineering and Environmental Science* (A. Y. Raya-Tapia, F. J. López-Flores, C. Ramírez-Márquez, & J. M. Ponce-Ortega, editors; pp. 13–50). Springer Nature Switzerland. https://doi.org/10.1007/978-3-032-03876-0_2
36. Sajja R, Sermet Y, Cwiertny D, Demir I. (2025). Integrating AI and Learning Analytics for Data-Driven Pedagogical Decisions and Personalized Interventions in Education. *Technology, Knowledge and Learning*. https://doi.org/10.1007/s10758-025-09897-9
37. Schraw G, Dennison RS. Assessing Metacognitive Awareness. Contemp Educ Psychol. 1994;19(4):460–75. https://doi.org/https://doi.org/10.1006/ceps.1994.1033.
38. Sheffler P, Rodriguez TM, Cheung CS, Wu R. (2022). Cognitive and metacognitive, motivational, and resource considerations for learning new skills across the lifespan. Wiley Interdisciplinary Reviews: Cogn Sci, 13(2), e1585.
39. Shu L, Li G. Application of improved clustering algorithm in mixed teaching of modern educational technology. Discover Artif Intell. 2025;5(1):195. https://doi.org/10.1007/s44163-025-00393-8.
40. Simon M. (2020). *Hypothetical Learning Trajectories in Mathematics Education BT - Encyclopedia of Mathematics Education* (S. Lerman, editor; pp. 354–357). Springer International Publishing. https://doi.org/10.1007/978-3-030-15789-0_72
41. Smith AK, Black S, Hooper LM. Metacognitive Knowledge, Skills, and Awareness: A Possible Solution to Enhancing Academic Achievement in African American Adolescents. Urban Educ. 2020;55(4):625–39. https://doi.org/10.1177/0042085917714511.

42. Song C, Shin S-Y, Shin K-S. (2024). Implementing the Dynamic Feedback-Driven Learning Optimization Framework: A Machine Learning Approach to Personalize Educational Pathways. In *Applied Sciences* (Vol. 14, Issue 2, p. 916). https://doi.org/10.3390/app14020916
43. Strielkowski W, Grebennikova V, Lisovskiy A, Rakhimova G, Vasileva T. AI-driven adaptive learning for sustainable educational transformation. Sustain Dev. 2025;33(2):1921–47. https://doi.org/https://doi.org/10.1002/sd.3221.
44. Villalonga-Gómez C, Mora-Cantallops M. Profiling distance learners in TEL environments: a hierarchical cluster analysis. Behav Inform Technol. 2022;41(7):1439–52. https://doi.org/10.1080/0144929X.2021.1876766.
45. Walker PB, Haase JJ, Mehalick ML, Steele CT, Russell DW, Davidson IN. (2025). Harnessing Metacognition for Safe and Responsible AI. In *Technologies* (Vol. 13, Issue 3, p. 107). https://doi.org/10.3390/technologies13030107
46. Wang H, Chen P, Luo J, Yang Y. Tailoring educational support with graph neural networks and explainable AI: Insights into online learners' metacognitive abilities. Comput Educ. 2026;240:105452. https://doi.org/https://doi.org/10.1016/j.compedu.2025.105452.
47. You J, Lou E, Afrouziyeh M, Zukiwsky NM, Zuidhof MJ. A supervised machine learning method to detect anomalous real-time broiler breeder body weight data recorded by a precision feeding system. Comput Electron Agric. 2021;185:106171. https://doi.org/https://doi.org/10.1016/j.compag.2021.106171.
48. Zimmerman BJ. Becoming a self-regulated learner: An overview. Theory into Pract. 2002;41(2):64–70.

Figures

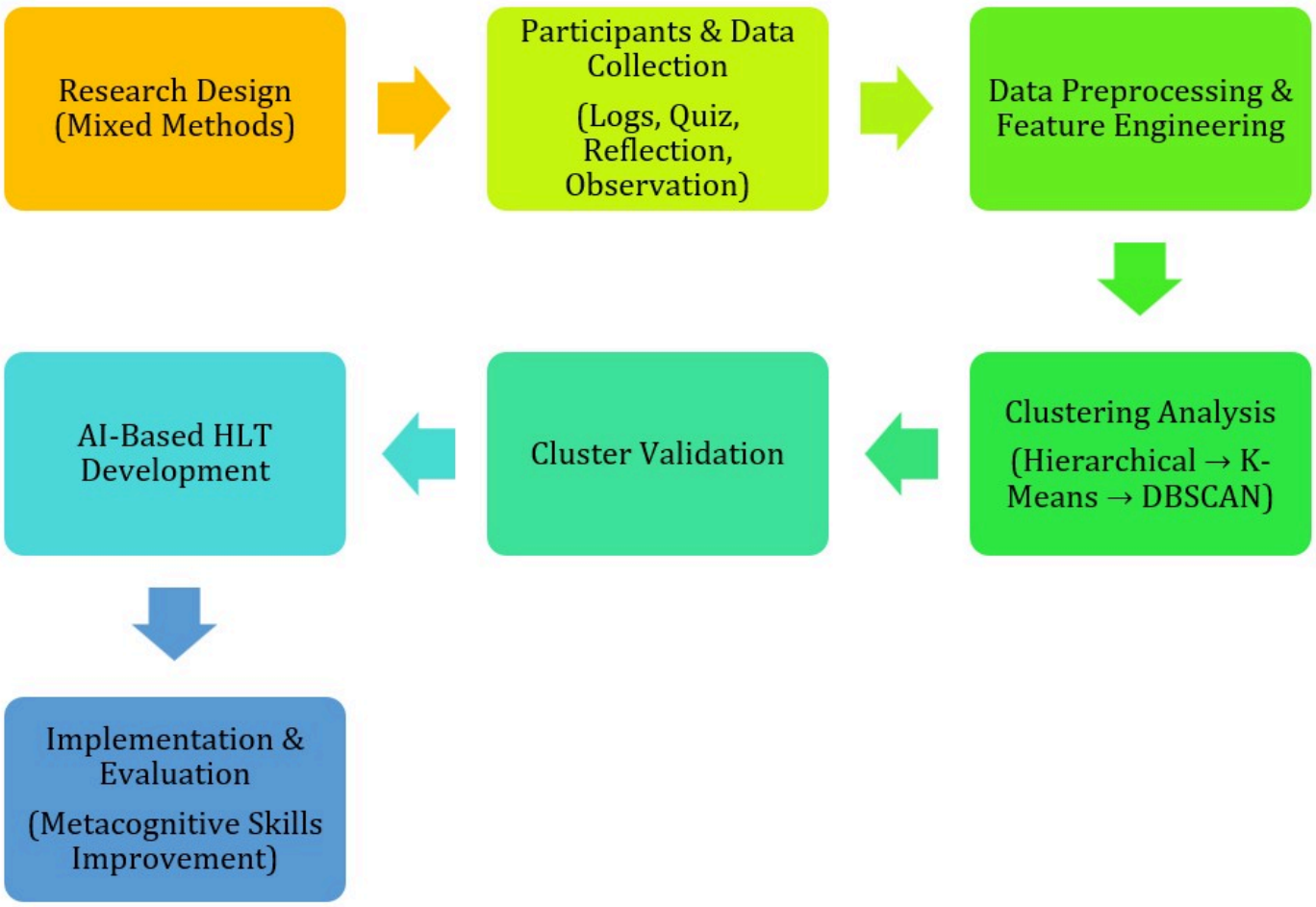


Figure 1

Methodological Flowchart

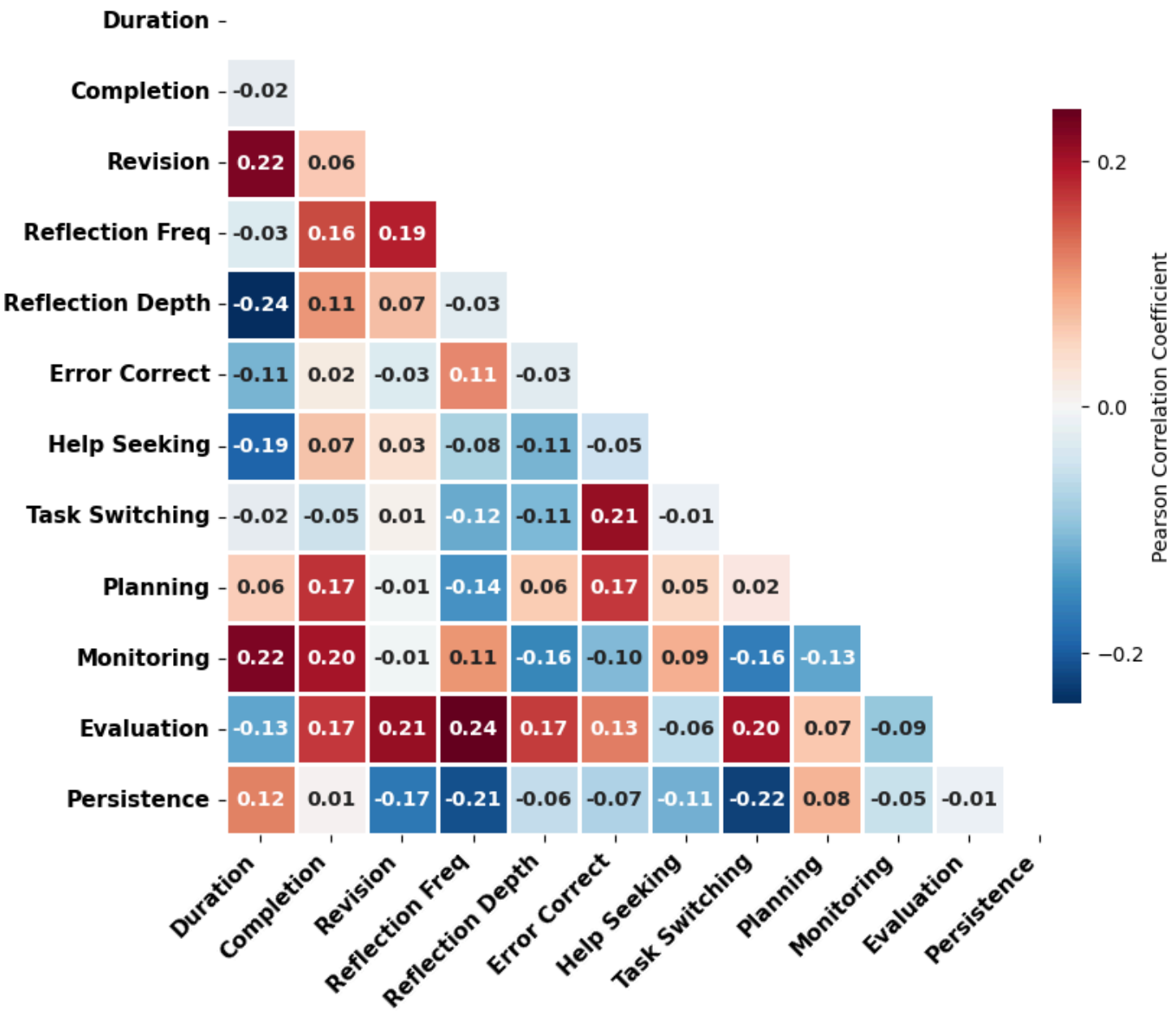


Figure 2

Figure 1. Correlation Matrix

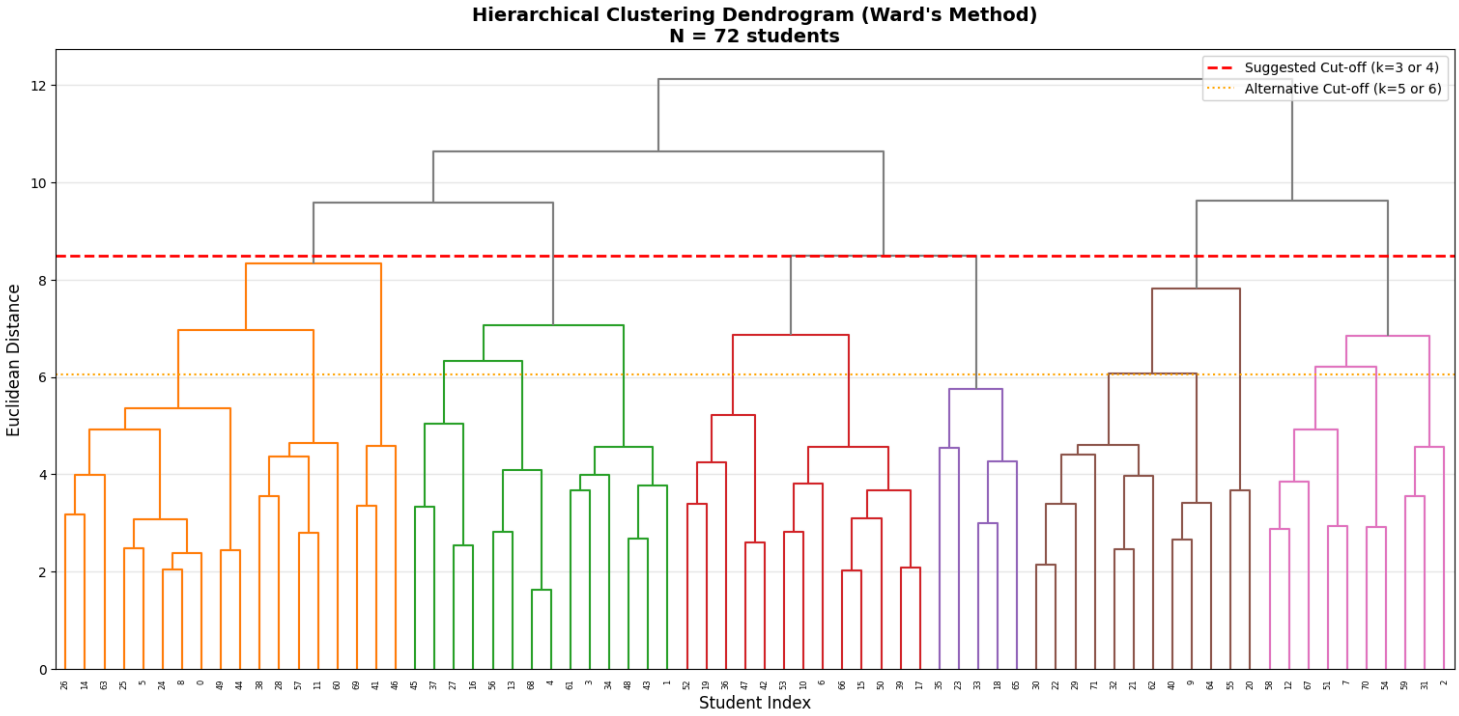


Figure 3

Figure 2. Hierarchical Clustering Dendrogram (Ward’s Method)

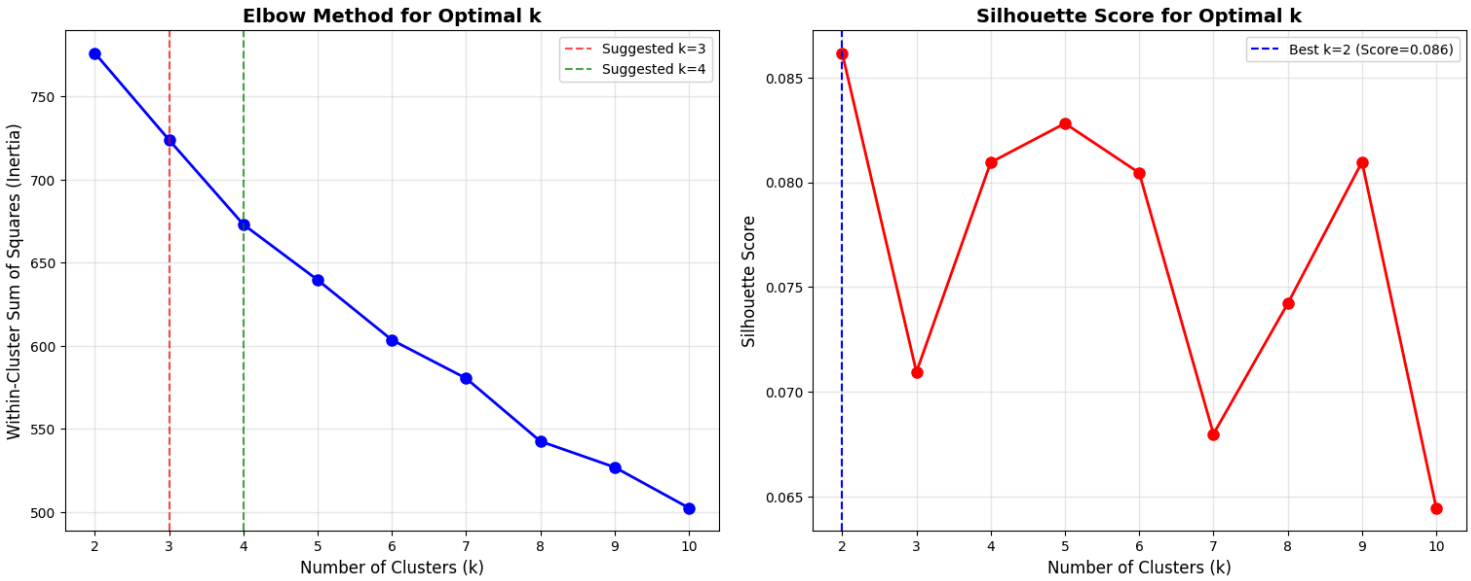


Figure 4

Figure 3. Cluster Selection

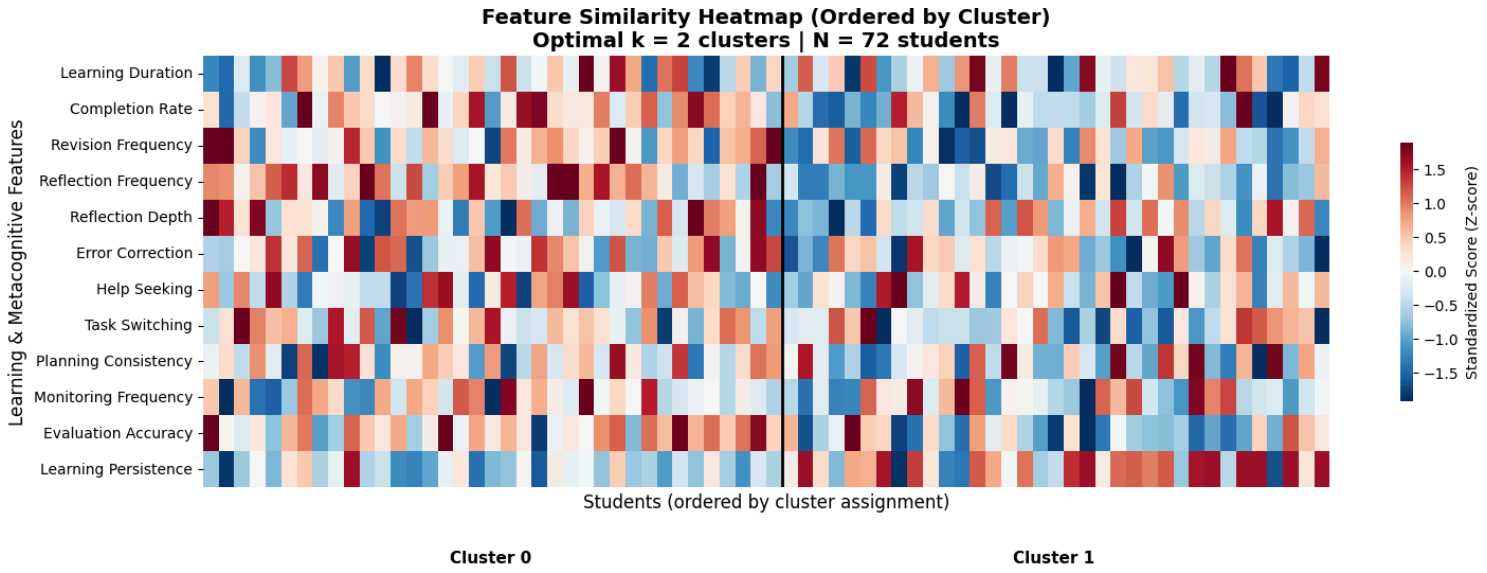


Figure 5

Figure 4. Feature Similarity Heatmap (Ordered by Cluster)

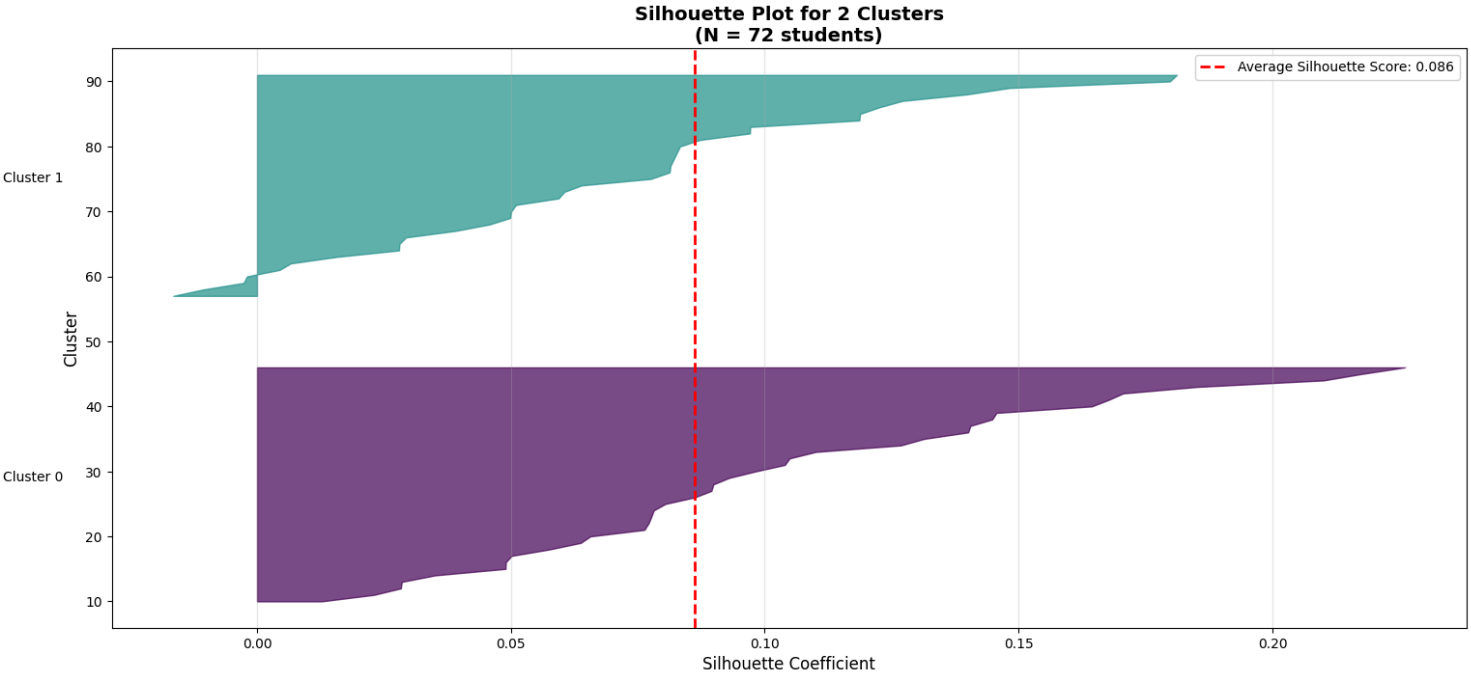


Figure 6

Figure 5. Silhouette Plot

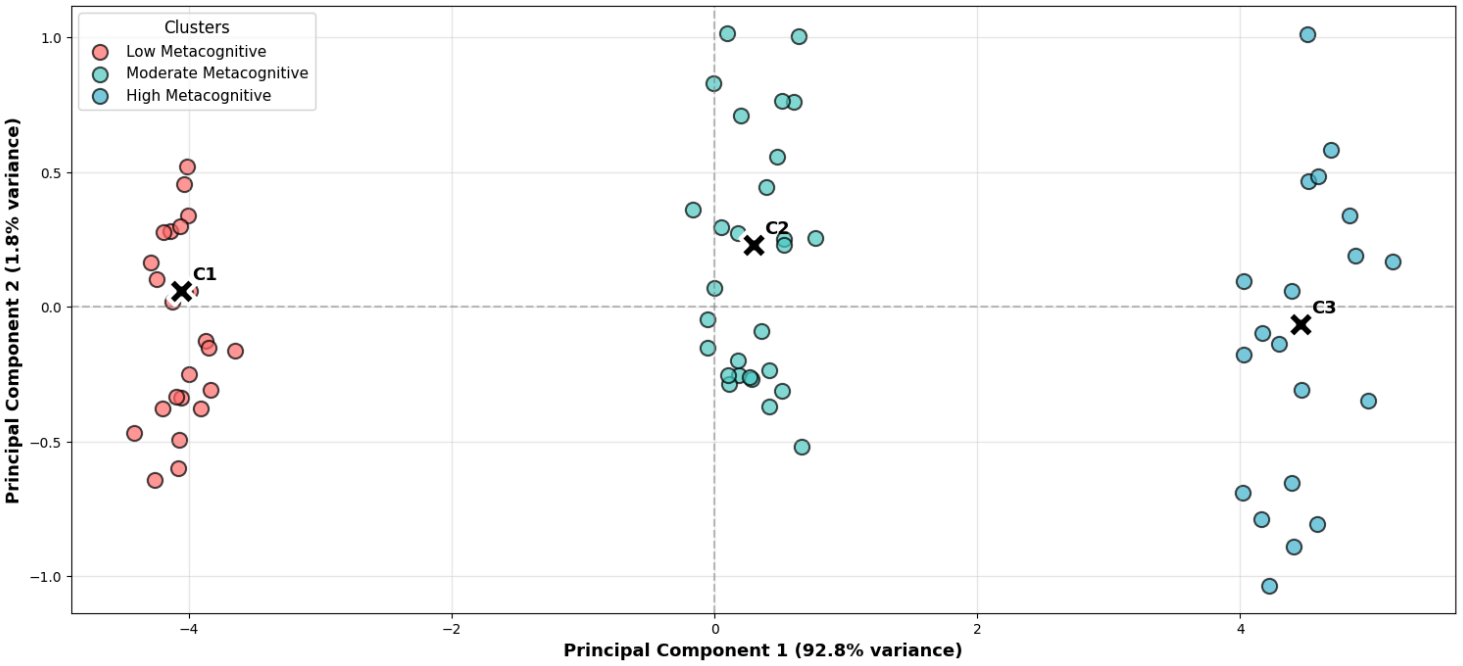


Figure 7

Figure 6. PCA Cluster

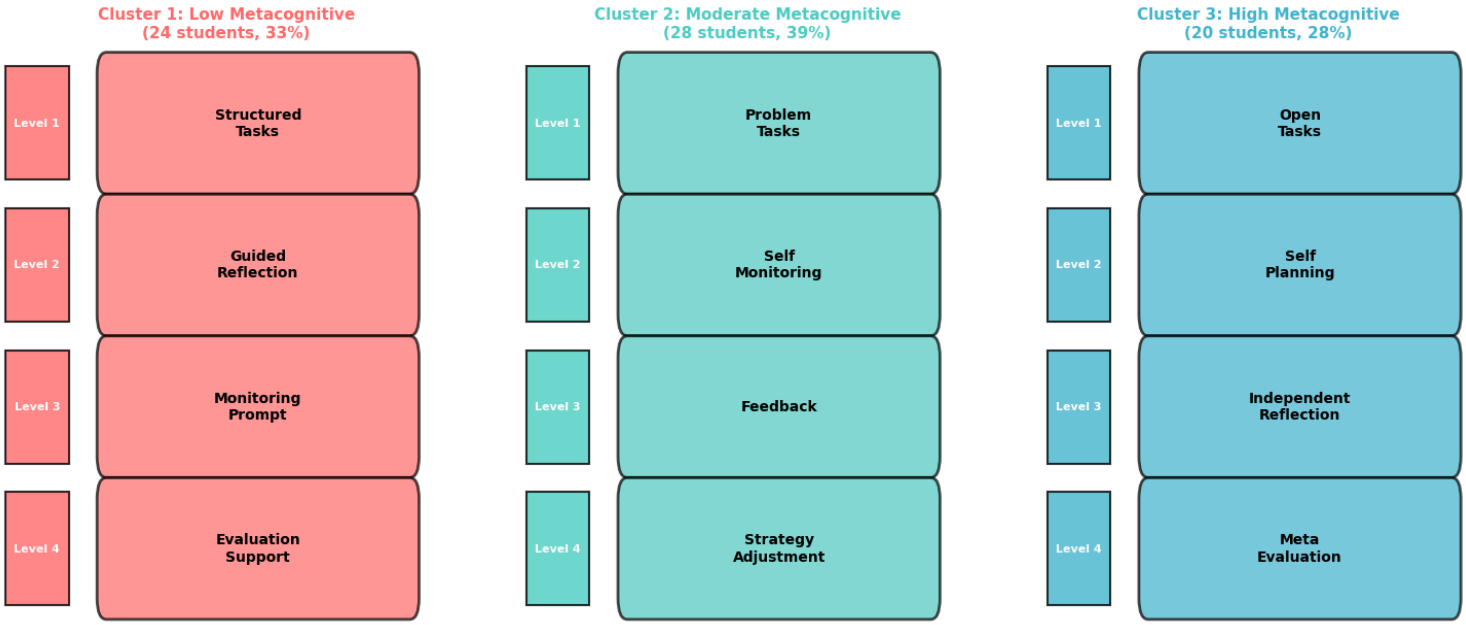


Figure 8

Figure 7. HLT Learning Path Across Metacognitive Clusters

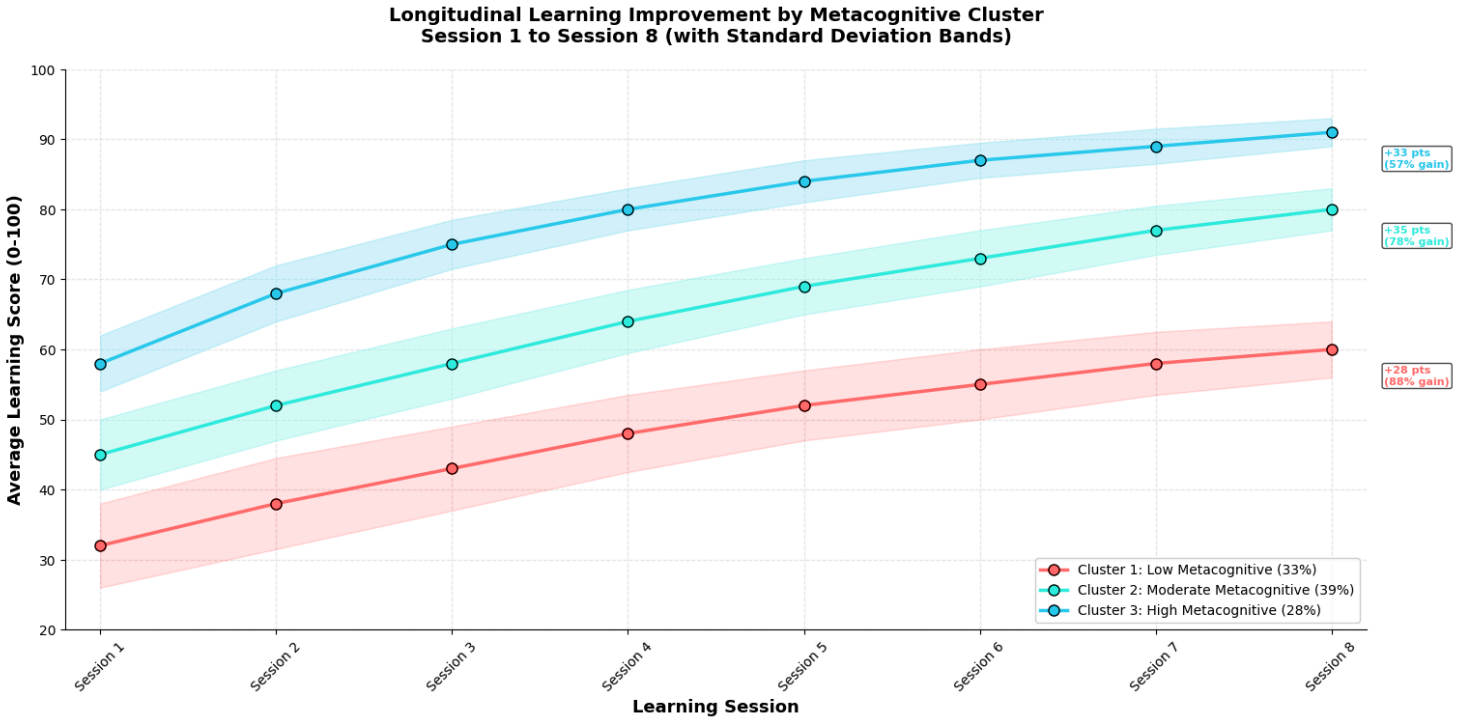


Figure 9

Figure 8. Longitudinal Learning Improvement by Metacognitive Cluster Session 1 to Session 8

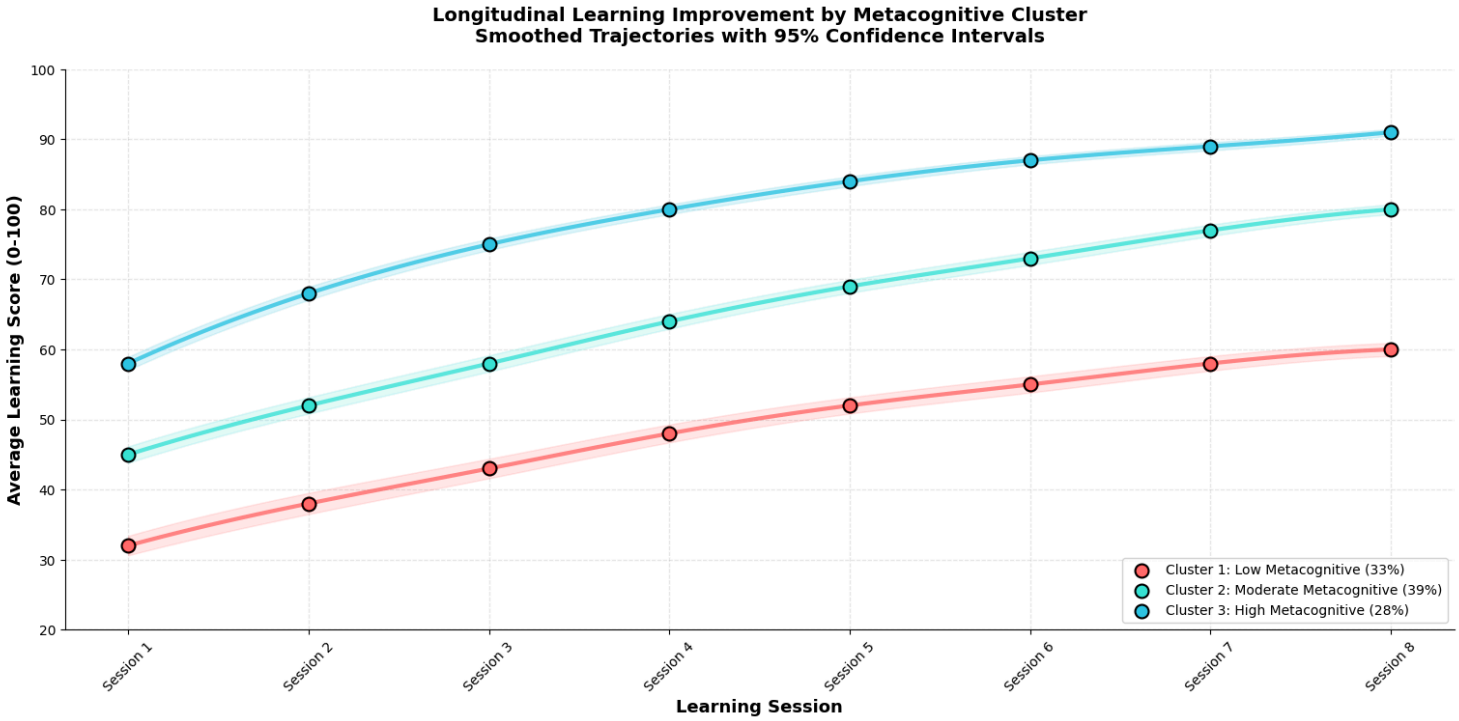


Figure 10

Figure 9. Longitudinal Learning Improvement by Metacognitive Cluster Smoothed Trajectories with 95% CI

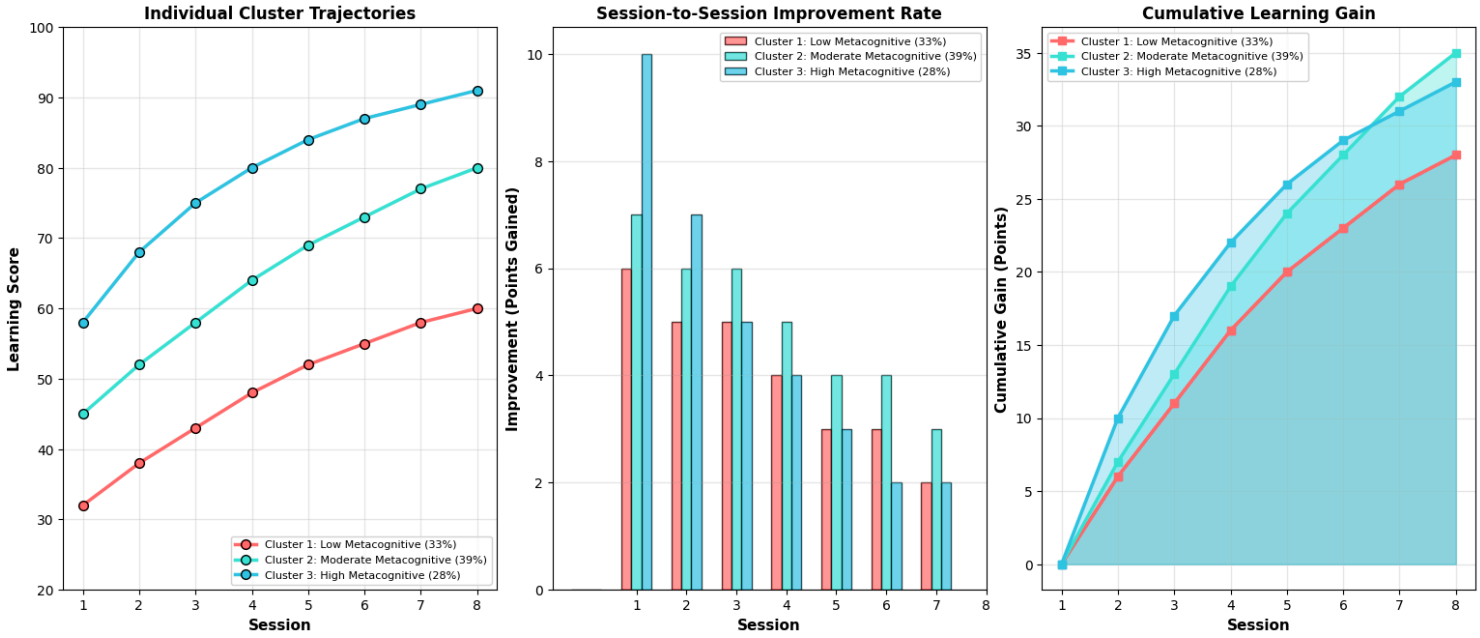


Figure 11

Figure 10. Longitudinal Learning Improvement Analysis