Published online on the page : https://journal.makwafoundation.org/index.php/intellect

I n t e l l e c t :
Indonesian Journal of Learning and Technological Innovation
| ISSN (*Online*) 2962-9233 |

Development and Preliminary Evaluation of an AHP-Based Decision Support System for Teaching Method Recommendations in a Madrasah

Rendi Kurniyawan[1,], Liza Efriyanti[2]*

[1,2]Universitas Islam Negeri Sjech M. Djamil Djambek Bukittinggi, Bukittinggi, Indonesia

Article Information

Article History:
Submitted: July 25, 2026
Revised: November 3, 2026
Accepted: November 6, 2026
Published: November 7, 2026

Keywords

Decision Support System, Analytic Hierarchy Process, Teaching Method Recommendation, Differentiated Instruction, Educational Software.

Correspondence

E-mail: rendikrnwn1606@gmail.com*

A B S T R A C T

This study developed and preliminarily evaluated SIPPERS, a web-based decision support system designed to provide teaching-method recommendations at MTs Laboratorium UIN Bukittinggi. The reported development process combined needs assessment through observation and an interview, Agile-oriented software development, and the Analytic Hierarchy Process (AHP). The decision hierarchy comprises three criteria (learning-style category, classroom participation, and learning comprehension), nine subcriteria, and five instructional alternatives. The reported criterion weights were 0.633, 0.260, and 0.106, with a criterion-level consistency ratio of 0.048. Ten documented black-box scenarios passed. Functional-suitability and compatibility assessments reported scores of 52/52 and 12/12, respectively, while seven educational staff members provided a usability score of 1,004/1,050 (95.62%). A documentation audit identified an alternative-priority vector that does not sum to one, leaving the reproducibility and pedagogical validity of the recommendations unresolved pending the original comparison matrices. The evidence supports preliminary software implementation and technical evaluation; it does not establish that the recommendations improve learning outcomes or are optimal for individual learners.

Abstrak

Penelitian ini mengembangkan dan mengevaluasi secara awal Sistem Personalisasi Pembelajaran Siswa (SIPPERS), yaitu sistem pendukung keputusan berbasis web untuk menyajikan rekomendasi metode mengajar di MTs Laboratorium UIN Bukittinggi. Pengembangan menggabungkan analisis kebutuhan melalui observasi dan wawancara, tahapan pengembangan perangkat lunak yang dilaporkan sebagai Agile, dan Analytic Hierarchy Process (AHP). Hierarki keputusan memuat tiga kriteria – gaya belajar, karakteristik keaktifan, dan daya tangkap – sembilan subkriteria, serta lima alternatif metode pembelajaran. Bobot kriteria yang dilaporkan adalah 0,633; 0,260; dan 0,106, dengan Consistency Ratio (CR) tingkat kriteria sebesar 0,048. Sepuluh skenario black-box berstatus berhasil; rekap penilaian functional suitability dan compatibility masing-masing mencapai 52/52 dan 12/12, sedangkan tujuh pengguna memberikan skor usability 1004/1050 (95,62%). Audit naskah menemukan ketidaksesuaian normalisasi pada satu vektor alternatif AHP sehingga validitas keputusan dan reproduksibilitas pembobotan masih memerlukan data sumber. Penelitian ini menunjukkan implementasi dan evaluasi teknis awal perangkat, bukan bukti ketepatan pedagogis rekomendasi ataupun peningkatan hasil belajar.

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

1. Introduction

Differentiated instruction seeks to adapt learning processes and support to students' needs, but its implementation requires ongoing assessment, instructional choices, and evaluation. Fatimatuzzahrah et al. [1] have documented challenges in implementing the Merdeka Curriculum in Indonesia. Recent reviews suggest that meaningful differentiation depends on learning objectives,

readiness, learning activities, and feedback rather than simply grouping students [2], [3]. Earlier research syntheses indicate that technology can support differentiation, although its effects depend on how interventions are integrated into teaching [4], [5].

For teachers, the practical challenge extends beyond collecting student profiles to translating that information into defensible, explainable choices of teaching methods. A review of data use in learning platforms emphasizes the need to connect student data with pedagogical reasoning and instructional decisions [6]. School data systems may improve the information available to decision-makers, but their value depends on system design and the context in which they are adopted [7]. Teachers' data literacy, confidence, and readiness to use evidence also matter [8], [9]. Data-driven decision-making frameworks further highlight the need to consider technological, pedagogical, and contextual information together [10].

Decision support systems (DSSs) provide a transparent way to organize criteria, weights, and alternatives. In education, adaptive platforms have been used to adjust learning resources and activities [11]. Recent reviews also identify distinct sources of adaptation, including learning progress, observed activity, and assessment data, which should not be treated as interchangeable [12]. Developing adaptive systems requires attention to the quality of student models, teachers' roles, and the implications of data use [13]. Other work explores personalization through natural language processing and digital feedback [14]. SIPPERS does not employ adaptive artificial intelligence; accordingly, studies of AI-based personalization provide comparative context rather than evidence of the effectiveness of the system evaluated here [15], [16], [17].

Studies more closely related to the present design include Nurjanah et al. [18], who used AHP to select instructional models for teachers; Farizi et al. [19], who applied TOPSIS to recommend school subjects; and Troussas et al. [20], who employed TOPSIS to adapt instructional strategies. Together, these studies show that multicriteria decision-making in education is not a new approach. More useful distinctions concern the unit of decision-making, input criteria, derivation of weights, system outputs, and the nature and depth of evaluation, rather than a claim that no similar systems exist.

Methodologically, AHP uses pairwise comparisons to synthesize priorities within a hierarchy [21]. Reviews of its applications demonstrate its flexibility in multicriteria decision-making [22], while methodological developments underscore the importance of consistency checks, aggregation across assessors, and sensitivity to criterion weights [23]. Mathematical consistency is therefore one property of the weighting process; it does not establish the pedagogical validity of student-profile categories or the resulting teaching-method recommendations.

The pedagogical rationale for visual, auditory, and kinesthetic categories requires particular caution. A foundational review found insufficient evidence to justify automatically matching teaching methods to students' preferred modalities [24], and subsequent criticism has warned against treating learning-style labels as instructional prescriptions [25]. A recent meta-analysis likewise found limited and inconsistent support for the matching hypothesis [26]. Other experimental and conceptual studies have identified similar limitations in the inferences that can be drawn from learning-style labels [27], [28], [29], [30], [31]. In SIPPERS, these categories are therefore treated as operational inputs to a decision model that still requires validation, not as diagnoses of ability or guarantees of instructional effectiveness.

Against this background, the issue examined in this study is how three researcher-selected student-profile inputs can be incorporated into an AHP-weighted application for teachers, and what evidence is available regarding its functionality and user experience. This is a design and decision-reporting problem in a madrasah setting, not a claim that AHP, TOPSIS, or teaching-method recommendation systems have not been studied. The contribution supported at this stage is the PHP-MySQL-based SIPPERS artifact and a demonstration of a decision workflow involving three criteria,

nine subcriteria, and five teaching methods. The pedagogical validity of the mapping and its effects on learning remain to be evaluated.

A needs assessment was conducted at MTs Laboratorium UIN Bukittinggi through observations from July to November 2025 and an interview with the school's Informatics teacher/operator on 20 October 2025. According to the available records, teaching methods were still selected manually, with a need to consider differences in students' learning behavior. The observation that students participated more actively in practical activities than in lectures indicates a local need. By itself, however, this observation cannot establish fixed learning styles or show that practical work would be the most effective method for particular students.

This study aimed to develop SIPPERS and conduct a preliminary evaluation of its capacity to help teachers organize criteria and rank alternative teaching methods. The central question was whether the criteria-processing workflow, recommendation display, and software functions could be implemented and tested in a documented manner. The study was not designed to measure the system's effects on achievement, motivation, or engagement. Accordingly, references to the 'most suitable method' denote only the ranking generated from the parameters entered into the system.

2. Methodology

2.1. Research Design

The study adopted a research-and-development approach to produce a decision support system prototype and document its preliminary evaluation. Agile describes the software development sequence reported by the authors, whereas AHP is the decision-analysis technique embedded in the application; these are distinct methodological components. The available evidence covers needs assessment, interface design, implementation, AHP calculations, and product testing. Iteration or sprint records, version-acceptance decisions, and evaluations of learning outcomes were not available to verify the complete Agile process or pedagogical effectiveness.

2.2. Research Setting and Data Sources

The study was conducted at MTs Laboratorium UIN Bukittinggi. The reported needs-assessment data came from observations between July and November 2025 and an interview with an Informatics teacher/school operator on 20 October 2025. For product evaluation, the manuscript also reports separate assessor groups for content, language, functional suitability, compatibility, and usability, the last involving seven respondents. The number of students whose profiles were recorded, criteria for selecting experts, and user recruitment procedures could not be established from the available documentation. The findings are therefore not generalized to the student population or to all teachers.

2.3. Instruments and Data Collection

System requirements were identified through classroom observations and an interview about teaching-method selection. Subsequent testing involved black-box scenarios, summaries of expert ratings, and a usability questionnaire reportedly comprising 30 items. The interview guide, profile-classification rubric, and item-level rating sheets were not provided. Verification of the indicators, scales, and measurement reliability requires the original instruments. Student information is treated here as part of the system development context, not as evidence that the learning categories have been validated.

The implemented decision structure consists of three criteria, nine subcriteria, and five alternatives. The learning-style criterion includes kinesthetic, visual, and auditory categories; participation is classified as active, moderately active, or less active; and learning comprehension is categorized as fast, moderate, or slow. The alternatives are lecture, discussion, practical work, Project-Based Learning (PjBL), and Problem-Based Learning (PBL). These categories serve as coded inputs to

the development model. The original manuscript did not provide independently testable classification rules or evidence of psychometric validation.

2.4. System Development Stages

Software development followed a sequence described as Agile: planning, design, coding, testing, and release. The results section presents the use-case design, user interfaces, database, and implementation of the web application. However, documentation of development iterations, changes in user requirements, and between-sprint testing was not available. In this study, Agile therefore refers to the reported development framework, not to a verified level of process maturity.

As shown in the development workflow in Figure 1, the identified requirements informed the AHP decision structure and application design. After implementation, software functions and user experience were evaluated using the reported instruments. Each stage requires traceable evidence, particularly the AHP matrices, software revision records, and testing documentation. The reported release stage does not, by itself, establish that the software has been evaluated for its effectiveness in teaching and learning.

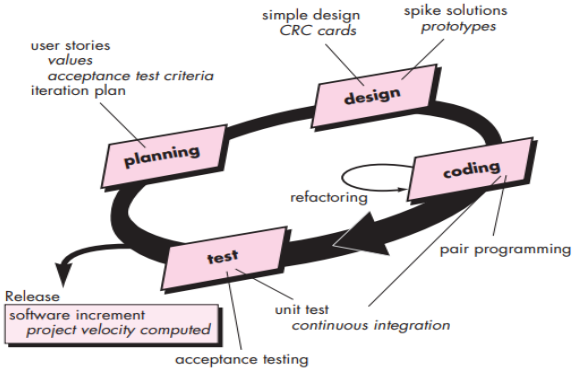


Figure 1. Agile development workflow

2.5. AHP Decision-Making Model

AHP was used to calculate relative priorities for three criteria, nine subcriteria, and five teaching-method alternatives. Pairwise judgments should be obtained from assessors competent to evaluate the decision being modeled [21]. If more than one assessor is involved, the rules for aggregating judgments and resolving disagreements should be reported [23]. The manuscript provides weight summaries but does not include the original comparison matrices, the assessors' identities and number, or the procedure for combining their judgments.

The reported criterion weights are 0.633 for learning style, 0.260 for classroom participation, and 0.106 for learning comprehension. Their sum of 0.999 reflects rounding to three decimal places and does not, by itself, indicate a normalization error. Each criterion has three subcriteria, yielding nine subcriteria in total and five teaching-method alternatives. Because learning style receives the greatest weight, sensitivity analysis is needed before using the model as the basis for consequential pedagogical decisions.

AHP requires pairwise-comparison matrices, priority vectors, the maximum eigenvalue (lambda-max), Consistency Index (CI), Random Index (RI), and Consistency Ratio (CR). The manuscript uses a CR threshold of 0.10 as a consistency check. Without the matrices and level-specific calculations, the claim that all matrices are consistent cannot be reproduced from the final weights alone [22].

The final score for each alternative is calculated as follows:

$$\text{Method Score} = \Sigma(BG \times BA)$$

Here, BG denotes the global subcriterion weight and BA the alternative weight under the corresponding subcriterion. The resulting scores form the basis for the system's teaching-method recommendations for students.

2.6. System Testing

Evaluation comprised four components: (i) scenario-based black-box functional testing, (ii) assessment of software functional suitability, (iii) compatibility testing, and (iv) user-rated usability. Expert assessments of content and language validity were distinct from these software quality characteristics. Auditing an Aiken's V coefficient requires item-level ratings, the number of experts, and the original rating scale [32]. Similarly, usability questionnaire scores reflect users' assessments of the application rather than the accuracy of its recommendations [33].

For reproducibility, the reported AHP tables need to be supplemented with input matrices, judgment aggregation and rounding rules, and CR calculations for each level. Content-validity testing requires the original scores from each item and rater. Usability reporting should identify the instrument's constructs, users' tasks, and interpretation criteria. Until these records are available, the evaluation figures should be regarded as reported results rather than independently verified measurements.

3. Results and Discussion

3.1. Results

3.1.1. Needs Assessment

The reported observations and interview at MTs Laboratorium UIN Bukittinggi identified a need to help teachers organize student-profile information and compare teaching methods without calculating weights manually. The evidence comes from one madrasah and should not be treated as a comprehensive survey. The intended functions include student-data management, criterion configuration, AHP calculations, and presentation of recommendations.

The system's principal functional requirements include student-data management, criterion and subcriterion management, AHP pairwise-comparison processing, priority-weight calculations, student assessments, teaching-method recommendations, and report display and export. These requirements informed the design and implementation of the Student Learning Personalization System (SIPPERS). The decision criteria are learning style, classroom participation, and learning comprehension. Each criterion has three subcriteria, which are used to assess five teaching-method alternatives: lecture, discussion, practical work, Project-Based Learning (PjBL), and Problem-Based Learning (PBL).

3.1.2. AHP Criteria and Weighting

The SIPPERS AHP hierarchy comprises the recommendation objective, three criteria, nine subcriteria, and five alternatives. Tables 1-3 present the reported weights; the figures are retained from the source manuscript so they can be checked against the original matrices.

The criterion-level pairwise comparisons produced the priority weights shown in the following table.

Table 1. Criterion Priority Weights

Criterion	Priority Weight
Learning Style	0.633
Classroom Participation	0.260
Learning Comprehension	0.106

Table 1 reports criterion weights of 0.633, 0.260, and 0.106. The reported criterion-level CR was 0.048, below the threshold of 0.10. This indicates consistency as reported by the authors; independent verification requires the underlying matrix, lambda-max, and RI values used to obtain that CR.

Each set of subcriteria was weighted in the next step. The resulting local and global weights are shown in the following table.

Table 2. Subcriterion Priority Weights and Global Weights

Criterion	Criterion Weight	Subcriterion	Subcriterion Weight	Global Weight (Criterion Weight x Subcriterion Weight)
Learning Style	0.633	Kinesthetic	0.608	0.385
		Visual	0.272	0.172
		Auditory	0.120	0.076
Classroom Participation	0.260	Active	0.737	0.192
		Moderately Active	0.186	0.048
		Less Active	0.077	0.020
Learning Comprehension	0.106	Fast	0.715	0.076
		Moderate	0.206	0.022
		Slow	0.079	0.008

The global weights in Table 2 are calculated by multiplying each criterion weight by the local weight of its subcriterion. The original comparison matrices are needed to reassess the claim that all subcriterion-level CR values meet the 0.10 threshold. Consistent weights should not be interpreted as confirmation that student-profile classifications and instructional decisions have been validated.

Table 3 reports five alternative weights for each subcriterion. An arithmetic audit found that the five weights under 'Learning Comprehension (Fast)' sum to 0.621 (0.039 + 0.093 + 0.163 + 0.163 + 0.163), rather than approximately 1.000 as expected for a normalized priority vector. The original figures have been retained because the comparison matrices are unavailable. Any transcription or calculation error, and its possible effect on the rankings, must be checked against the initial assessment data before the model informs substantive decisions.

Table 3. Alternative Weights for Each Subcriterion

Alternative Subcriterion	Weights by	Alternative Subcriterion	Weights by	Alternative Subcriterion	Weights by
Learning Style (Kinesthetic)		Participation (Active)		Comprehension (Fast)	
Method	Priority Weight	Method	Priority Weight	Method	Priority Weight
Lecture	0.038	Lecture	0.041	Lecture	0.039
Discussion	0.073	Discussion	0.095	Discussion	0.093
Practical Work	0.409	Practical Work	0.182	Practical Work	0.163
PjBL	0.313	PjBL	0.405	PjBL	0.163
PBL	0.166	PBL	0.277	PBL	0.163
Learning Style (Visual)		Participation (Moderately Active)		Comprehension (Moderate)	
Method	Priority Weight	Method	Priority Weight	Method	Priority Weight
Lecture	0.044	Lecture	0.059	Lecture	0.060
Discussion	0.096	Discussion	0.376	Discussion	0.375
Practical Work	0.256	Practical Work	0.128	Practical Work	0.241
PjBL	0.436	PjBL	0.183	PjBL	0.128
PBL	0.168	PBL	0.253	PBL	0.195
Learning Style (Auditory)		Participation (Less Active)		Comprehension (Slow)	
Method	Priority Weight	Method	Priority Weight	Method	Priority Weight
Lecture	0.489	Lecture	0.490	Lecture	0.422

Discussion	0.233	Discussion	0.233	Discussion	0.151
Practical Work	0.052	Practical Work	0.051	Practical Work	0.252
PjBL	0.097	PjBL	0.086	PjBL	0.073
PBL	0.128	PBL	0.139	PBL	0.102

Table 3 shows that alternative weights vary across subcriteria. For example, practical work has a weight of 0.409 under kinesthetic and 0.256 under visual. These differences illustrate how the system's outputs respond to its inputs, but they do not establish that the recommended method is pedagogically optimal.

Particular caution is warranted because the learning-style criterion, weighted at 0.633, dominates the ranking. No sensitivity analysis of the weights or uncertainty in profile classification was available. Recommendations should therefore be regarded as algorithmic outputs for teachers to review, not as automatically validated instructional decisions.

To illustrate the calculations documented in the manuscript, the example of Student A is used. The calculation employs the global weight of 0.022 for the 'Moderate' comprehension category. The illustrative profile is therefore kinesthetic, moderately active, and moderate in comprehension. The original reference to 'Fast' is a discrepancy in the narrative that still requires checking against the source records.

Learning style: Kinesthetic.

Classroom participation: Moderately active.

Learning comprehension (category used in the example): Moderate.

In this illustration, which uses the 'Moderate' comprehension category, each alternative's score is obtained by summing the products of the selected subcriteria's global weights and their corresponding alternative weights. The calculations below retain the figures reported in the source tables. A valid comparison with the 'Fast' profile cannot be made until that priority vector has been verified.

$$V = (BG_1 \times BA_1) + (BG_2 \times BA_2) + (BG_3 \times BA_3)$$

Notation:

V = Method score

BG = Global subcriterion weight

BA = Alternative weight under a subcriterion

First, identify the global weight of each subcriterion corresponding to the student profile.

Global weight, Kinesthetic = 0.385

Global weight, Moderately active = 0.048

Global weight, Moderate = 0.022

Once the relevant global subcriterion weights have been identified, calculate the score for each teaching-method alternative as follows.

Lecture (V_1)

$$V_1 = (0{,}385 \times 0{,}038) + (0{,}048 \times 0{,}059) + (0{,}022 \times 0{,}060)$$
$$= 0{,}014630 + 0{,}002832 + 0{,}001320$$
$$= 0{,}0188$$

Discussion (V_2)

$$V_2 = (0{,}385 \times 0{,}073) + (0{,}048 \times 0{,}376) + (0{,}022 \times 0{,}375)$$
$$= 0{,}028105 + 0{,}018048 + 0{,}008250$$
$$= 0{,}0544$$

Practical work (V_3)

$$V_3 = (0{,}385 \times 0{,}409) + (0{,}048 \times 0{,}128) + (0{,}022 \times 0{,}241)$$
$$= 0{,}157465 + 0{,}006144 + 0{,}005302$$
$$= 0{,}1689$$

Project Based Learning atau PjBL (V_4)

$$V_4 = (0{,}385 \times 0{,}313) + (0{,}048 \times 0{,}183) + (0{,}022 \times 0{,}128)$$
$$= 0{,}120505 + 0{,}008784 + 0{,}002816$$
$$= 0{,}1321$$

Problem Based Learning atau PBL (V_5)

$$V_5 = (0{,}385 \times 0{,}166) + (0{,}048 \times 0{,}253) + (0{,}022 \times 0{,}195)$$
$$= 0{,}063910 + 0{,}012144 + 0{,}004290$$
$$= 0{,}0803$$

The resulting scores are then ranked from highest to lowest, as shown in the following table.

Table 4. Ranking Results

Rank	Teaching-Method Alternative	Score
1.	Practical Work	0.1689
2.	Project Based Learning (PjBL)	0.1321
3.	Problem Based Learning (PBL)	0.0803
4.	Discussion	0.0544
5.	Lecture	0.0188

In the illustration using the 'Moderate' category, practical work has the highest reported score (0.1689), followed by PjBL (0.1321). This ranking reflects the relative scores generated by the specified combination of weights; it is not empirical evidence that practical work is the most appropriate teaching method for an actual student.

3.1.3. System Design

The design stage comprised three main components: system architecture, user interface, and database design.

3.1.3.1. System Architecture

The system design was represented using a use-case diagram for the Student Learning Personalization System at MTs Laboratorium UIN Bukittinggi, as shown below.

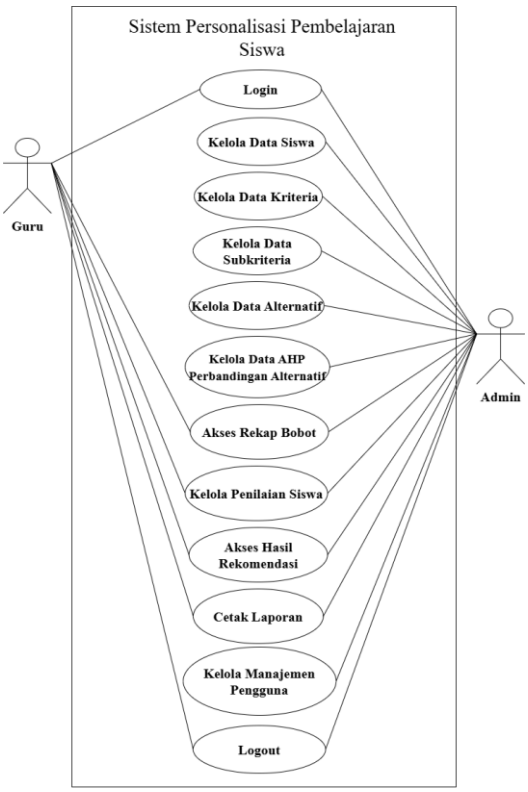


Figure 2. SIPPERS use-case diagram

3.1.3.2. User Interface Design

The user interface was designed to illustrate the intended appearance and navigation of the system. The proposed interface layouts are shown below.

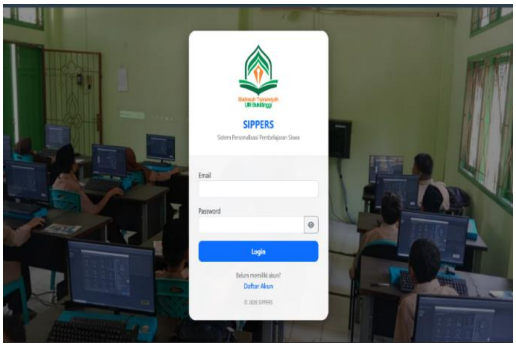


Figure 3. Login screen design

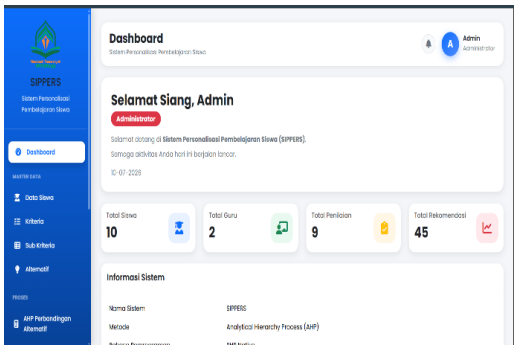


Figure 4. Dashboard design

3.1.3.3. Database Design

The database was represented using a class diagram showing the system classes, their attributes, and the relationships among the data.

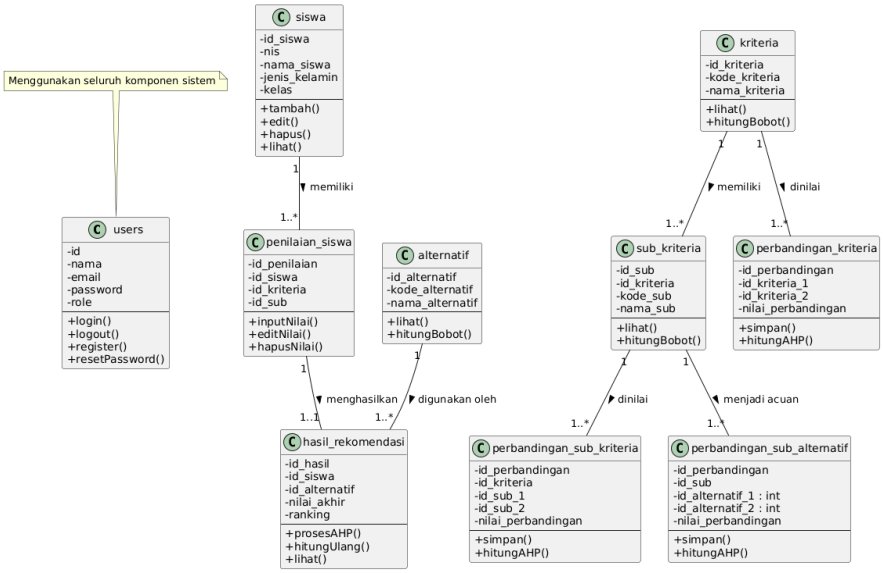


Figure 5. SIPPERS class diagram

3.1.4. System Implementation

The Student Learning Personalization System (SIPPERS) was implemented as a web-based application using PHP and MySQL, with Bootstrap supporting the user interface. The implementation includes the principal features shown below.

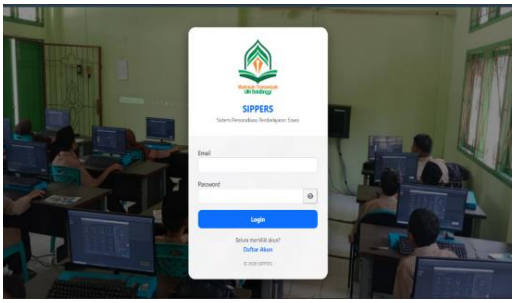


Figure 6. Login page

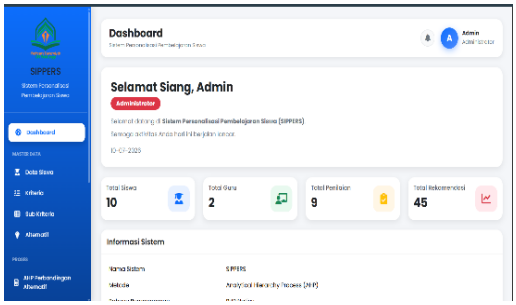


Figure 7. Dashboard page

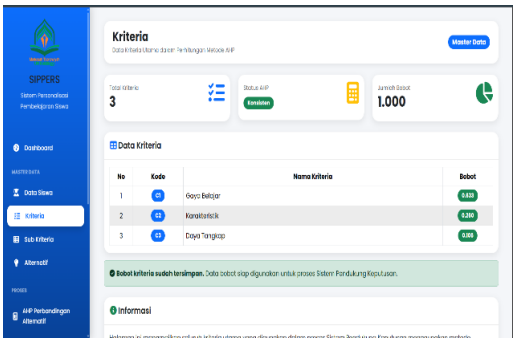


Figure 8. AHP weighting

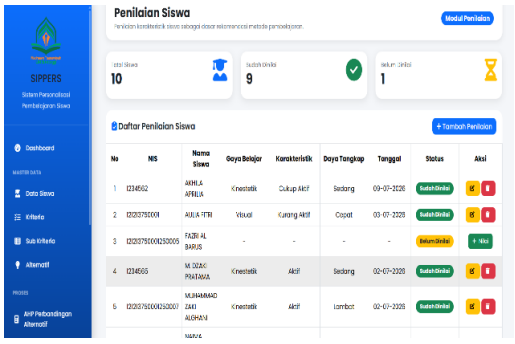


Figure 9. Student assessment

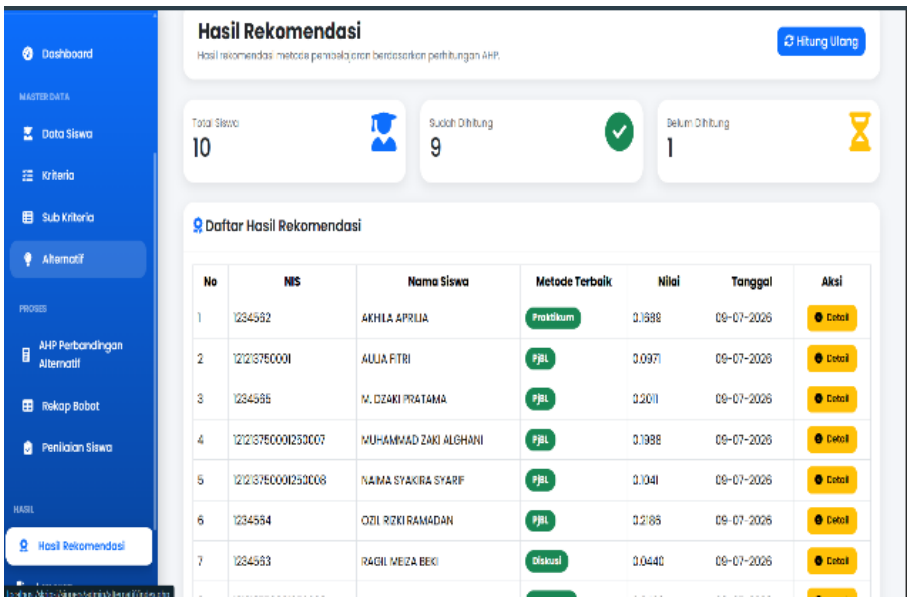


Figure 10. Recommendation results

3.1.5. Functional Testing

The black-box tests in Table 5 cover ten scenarios involving the application's main functions. All reported scenarios passed. However, the documentation does not include negative test inputs, initial conditions, or actual-output evidence sufficient to reproduce the tests in full.

The expert assessment results presented below require similarly cautious interpretation. Summary ratings cannot replace information about the items, scales, formulas, and assessor selection. Likewise, successful execution of a software function does not test the pedagogical appropriateness of its recommendations.

Table 5. Black-Box Testing

No.	Function Tested	Test Scenario	Result	Status
1.	Login	Verify email and password	Login verified; dashboard shown	Passed
2.	Student Data	Add, edit, or delete students	Student records managed	Passed
3.	Criterion	Calculate criterion weights	Weights calculated	Passed
4.	Subcriterion	Calculate subcriterion weights	Weights calculated	Passed
5.	AHP Alternatives	Calculate alternative weights	Weights calculated	Passed
6.	Weight Summary	Display all AHP weights	Weight summary displayed	Passed
7.	Student Assessment	Add, edit, or delete assessments	Assessments managed	Passed
8.	Recommendations	Display student recommendations	Recommendations generated	Passed
9.	Reports	Filter and export Excel/PDF reports	Reports filtered and exported	Passed
10.	User Management	Add, edit, or delete users	User accounts managed	Passed

As reported in the black-box test results, the tested SIPPERS functions operated as expected in the specified scenarios, and the software was judged suitable for use on that basis.

In addition to black-box testing, product quality was evaluated using reported ISO/IEC 25010-related assessments of validity, functional suitability, compatibility, and usability.

3.1.5.1. Validity Assessment

3.1.5.1.1. Content validity

Content validation was conducted by two computer science lecturers, Ahmad Zaki, M.Kom., and Dona Kurnia, M.Kom.; a school guidance and counseling teacher, Rahmiati, M.Pd.; and a homeroom teacher, Wahyuni, S.Pd. The assessment used Aiken's V, with a reported mean value of 0.97, as summarized in the following table.

Table 6. Summary of Content Validity Ratings (Source: authors' assessment data)

Validator	Score
Ahmad Zaki M.Kom	0.98
Dona Kurnia M.Kom	1
Rahmiati M.pd	0.89
Wahyuni S.pd	1
Total	3.87
Mean	0.97

3.1.5.1.2. Language validity

Language validity was assessed by Nurhamelya, S.Pd., an Indonesian-language teacher at MTs Laboratorium UIN Bukittinggi. The assessment used Aiken's V and reported a mean value of 1, as shown below.

Table 7. Language Validity Results

Validator	Score
Nurhamelya S.pd	1
Total	1
Mean	1

3.1.5.2. Functional suitability assessment

Four assessors evaluated functional suitability: Ahmad Zaki, M.Kom., Dona Kurnia, M.Kom., Rahmiati, M.Pd., and Dani Ramanda, S.Pd. The recorded score was 52 out of a maximum of 52, as shown in the following table.

Table 8. Results of Functional Suitability Testing

No.	Function	Expected Outcome				Total	Max. Score
		Respondent					
		1	2	3	4		
1.	Main Menu (Dashboard)	1	1	1	1	4	4
2.	Login Menu	1	1	1	1	4	4
3.	Student Data Menu	1	1	1	1	4	4
4.	AHP Calculation Menu	1	1	1	1	4	4
5.	Weight Summary Menu	1	1	1	1	4	4
6.	Student Assessment Menu	1	1	1	1	4	4
7.	Recommendation Results Menu	1	1	1	1	4	4
8.	Reports Page	1	1	1	1	4	4
9.	Logout Menu	1	1	1	1	4	4
10.	Add student records to the database	1	1	1	1	4	4
11.	Save student-assessment records to the database	1	1	1	1	4	4
12.	Completeness of program information	1	1	1	1	4	4
13.	Completion of the calculation steps	1	1	1	1	4	4
Total						52	52

Suitability percentage = (Score obtained)/(Maximum score) x 100%

$$= 52/52 \times 100\%$$
$$=100\%$$

The reported functional suitability score was 52 out of 52 (100%) for the evaluated items. This result should not be interpreted as a usability test or as evidence of pedagogically accurate recommendations.

3.1.5.3. Compatibility assessment

Compatibility was assessed by three experts: Dona Kurnia, M.Kom., Ahmad Zaki, M.Kom., and Dani Ramanda, S.Pd. The recorded score was 12 out of a maximum of 12, as shown below.

Table 9. Compatibility Test Results

No	Statement	Output			Total	Max. Score
		Respondent				
		1	2	3		
1	SIPPERS runs on Windows.	1	1	1	3	3
2	Runs on a PC with at least 4 GB RAM.	1	1	1	3	3
3	All features run in Google Chrome.	1	1	1	3	3
4	MySQL data storage, processing, and retrieval work.	1	1	1	3	3
Total					12	12

Suitability percentage = (Score obtained)/(Maximum score) x 100%

$$= 12/12 \times 100\%$$
$$=100\%$$

The reported compatibility score was 12 out of 12 (100%) in the specified test environment (Windows, Chrome, a computer with 4 GB of RAM, and a MySQL connection). It does not demonstrate compatibility across all browsers, operating systems, or devices.

3.1.5.4. Usability assessment

Usability was evaluated by 7 respondents, including homeroom teachers and other educational staff at MTs Laboratorium UIN Bukittinggi, as summarized in the following table.

Table 10. Usability Test Results

No	Responden	Usefidlness										Ease of Use										Ease of Learning										Satisfaction										Total Nilai	Skor Maksimal
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30		
1	Nurhamelya	5	5	5	5	5	5	5	5	5	4	4	4	4	5	5	5	5	5	5	5	5	5	5	4	5	5	5	5	5	5	145	150
2	Ocha Amelia	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	150	150
3	Wahyuni	5	5	5	5	5	5	5	4	4	5	5	5	5	5	5	4	5	4	5	5	4	5	5	4	5	5	5	5	5	5	144	150
4	Dani Ramanda	5	5	5	5	5	5	5	4	4	4	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	5	5	5	146	150
5	Fitri Yanis	5	5	5	5	5	5	5	4	4	5	5	4	5	5	5	4	5	5	5	5	5	5	5	4	5	5	4	5	5	5	144	150
6	Yunasri	5	5	5	5	5	5	5	5	4	5	5	5	5	5	5	4	5	5	4	5	5	5	5	5	5	5	5	5	5	5	147	150
7	Rahmiati	5	5	4	5	5	4	4	3	3	4	5	5	5	4	3	4	3	5	4	5	3	4	5	3	5	5	5	4	4	5	128	150
		Skor Total																														1004	1050

The usability results were calculated using the following formula:

$$P_Skor = (Total\ score)/(i \times r \times 5) \times 100\%$$
$$P_Skor = 1004/(30 \times 7 \times 5) \times 100\%$$
$$P_Skor = 1004/1050 \times 100\% = 95.62\%$$

Across seven respondents and 30 questionnaire items, the usability score was 1,004 out of a maximum of 1,050, or 95.62%. This is a summary of user perceptions measured with the reported instrument. The accuracy of the recommendations and their effects on learning were not evaluated.

3.2. Discussion

3.2.1. Interpretation of the Main Findings

SIPPERS demonstrates that researcher-defined student-profile inputs can be connected to an AHP workflow and a web-based recommendation interface. Successful execution of system functions must be distinguished from two further questions: whether student profiles are measured validly and whether the selected teaching methods lead to better learning experiences. Data-informed decision-making requires teachers to combine system information with assessment, experience, and pedagogical judgment [34], [35]. Learning analytics may enrich the understanding of students' needs [36], while research on teacher dashboards highlights the importance of user readiness and the usefulness of the information presented [37].

3.2.2. Comparison with Previous Studies

Compared with the AHP-based decision support system developed by Nurjanah et al. for selecting instructional models, SIPPERS adds student-profile categories, a web interface, and functional evaluation. There is, however, no evidence that its recommendations are more accurate. Unlike the TOPSIS-based subject-selection system developed by Farizi et al., SIPPERS recommends teaching methods rather than school subjects; the two systems also differ in their criteria and evaluation evidence. Troussas et al. used TOPSIS to adapt instructional strategies, whereas SIPPERS uses pairwise-weighted AHP. This difference in algorithm does not, by itself, constitute pedagogical novelty; theoretical justification and an audit of the decisions remain necessary. Research on adaptive learning and self-regulation likewise stresses the importance of well-defined inputs, system responses, and evaluations of learning experience [38], [39].

3.2.3. Research Contribution

The contribution supported by the evidence is a prototype that integrates profile assessment, AHP weighting, user management, and a recommendation display. Ten black-box scenarios reportedly passed. The software evaluation summaries show functional suitability of 52/52, compatibility of 12/12, and usability of 1,004/1,050 (95.62%). Content and language validity were reported from expert ratings, but the Aiken's V results cannot yet be audited without item-level scores and the associated assessor parameters. Reviews of AI and educational technology emphasize that system quality, teacher involvement, and pedagogical outcomes should be evaluated as distinct constructs [40], [41].

3.2.4. Study Limitations

This study was conducted in one madrasah, and its needs assessment relied primarily on one interview. The usability sample was also small (seven respondents). The rules for classifying student profiles, sources of AHP judgments, comparison matrices, and item-level ratings were not reported in sufficient detail for full replication. In Table 3, the alternative-priority vector for the 'Fast' category sums to 0.621; its accuracy, any resulting changes in rankings, and the stability of recommendations must therefore be reassessed using the original data. Learning-style categories should not be treated as fixed diagnoses that determine optimal teaching methods. Broader implementation would require field evaluation, assessment of teacher acceptance and learning outcomes, and attention to research ethics and student-data protection. Research on learning-progress monitoring further suggests that instructional decisions should be informed by students' actual development over time [42].

4. Conclusion

The study produced SIPPERS, a web-based prototype that uses AHP to rank five teaching-method alternatives on the basis of three criteria and nine subcriteria. The reported criterion weights were 0.633, 0.260, and 0.106, with a criterion-level CR of 0.048. The test documentation records ten

successful black-box scenarios, functional suitability of 52/52, compatibility of 12/12, and usability of 1,004/1,050 (95.62%) from seven respondents. These findings support the software's operation in the reported scenarios, not the pedagogical accuracy of its recommendations. Because the 'Fast' alternative vector in Table 3 is not normalized and the original AHP matrices are unavailable, the validity of the full set of decision outputs cannot be independently confirmed.

SIPPERS may help teachers organize information and consider alternative teaching methods, but professional judgment remains essential to the final decision. The study did not test improvements in achievement, motivation, or student engagement. Future work should provide the original AHP matrices and evaluation instruments, examine weight sensitivity and agreement between system recommendations and teacher judgments, assess student-data security, and then conduct controlled field evaluations aligned with educational objectives.

References

[1] Fatimatuzzahrah, L. Sakinah, and S. A. Alyasari, "Problematika Implementasi Kurikulum Merdeka Di Sekolah: Tantangan Membangun Kualitas Pendidikan," Jurnal Bintang Pendidikan Indonesia, vol. 2, no. 1, pp. 43–53, 2024, doi: 10.55606/jubpi.v2i1.2339

[2] Fatimatuzzahrah, L. Sakinah, and S. A. Alyasari, "Problematika Implementasi Kurikulum Merdeka Di Sekolah: Tantangan Membangun Kualitas Pendidikan," Jurnal Bintang Pendidikan Indonesia, vol. 2, no. 1, pp. 43–53, 2024, doi: 10.55606/jubpi.v2i1.2339

[3] S. Dell'Anna, A. Frizzarin, and G. Biagiotti, "Practice-oriented research on differentiation strategies in primary and secondary education: A systematic review," Teaching and Teacher Education, vol. 177, art. no. 105559, 2026, doi: 10.1016/j.tate.2026.105559

[4] B. N. Langelaan, L. Gaikhorst, W. Smets, and R. J. Oostdam, "Differentiating instruction: Understanding the key elements for successful teacher preparation and development," Teaching and Teacher Education, vol. 140, art. no. 104464, 2024, doi: 10.1016/j.tate.2023.104464

[5] M. I. Deunk, A. E. Smale-Jacobse, H. de Boer, S. Doolaard, and R. J. Bosker, "Effective differentiation practices: A systematic review and meta-analysis of studies on the cognitive effects of differentiation practices in primary education," Educational Research Review, vol. 24, pp. 31–54, 2018, doi: 10.1016/j.edurev.2018.02.002

[6] A. E. Smale-Jacobse, A. Meijer, M. Helms-Lorenz, and R. Maulana, "Differentiated Instruction in Secondary Education: A Systematic Review of Research Evidence," Frontiers in Psychology, vol. 10, art. no. 2366, 2019, doi: 10.3389/fpsyg.2019.02366

[7] A. Hase and P. Kuhl, "Teachers' use of data from digital learning platforms for instructional design: A systematic review," Educational Technology Research and Development, vol. 72, pp. 1925–1945, 2024, doi: 10.1007/s11423-024-10356-y

[8] D. Alonzo, V. Quimno, G. Townend, and C. Z. Oo, "Using information and communication technology (ICT)-based data systems to support teacher data-driven decision-making: Insights from the literature (2013–2023)," Educational Assessment, Evaluation and Accountability, vol. 36, pp. 433–451, 2024, doi: 10.1007/s11092-024-09443-8

[9] J. Vick et al., "Exploring Elementary Teachers' Confidence and Anxiety in Data-Based Decision-Making," TechTrends, vol. 69, pp. 1288–1303, 2025, doi: 10.1007/s11528-025-01113-w

[10] P. Hawlitschek and S. Henschel, "Promoting teacher knowledge and self-efficacy for data-based decision-making: Dimensional structure, sensitivity to professional development, and relationships with teachers' intended data use," Zeitschrift für Erziehungswissenschaft, vol. 28, pp. 1577–1603, 2025, doi: 10.1007/s11618-025-01321-3

[11] S. Evenstein Sigalov, A. Hershkovitz, A. Cohen, and R. Nachmias, "Trusting the data: An updated framework for teachers' data-driven decision-making (DDDM) in higher education," Education and Information Technologies, vol. 31, pp. 953–981, 2026, doi: 10.1007/s10639-025-13819-8

[12] F. Herliana, T. Hafinda, and R. Firmayanto, "Differentiated Instruction through Adaptive Learning Platform in Science Education: A Systematic Literature Review," Jurnal Pendidikan MIPA, vol. 25, no. 2, pp. 914–931, 2024, doi: 10.23960/jpmipa/v25i2.pp914-931

[13] K. M. Bach, S. I. Hofer, and S. Bichler, "Adaptive learning, instruction, and teaching in schools: Unraveling context, sources, implementation, and goals in a systematic review," Learning and Individual Differences, vol. 124, art. no. 102781, 2025, doi: 10.1016/j.lindif.2025.102781

[14] R. Pelánek, "Adaptive Learning is Hard: Challenges, Nuances, and Trade-offs in Modeling," International Journal of Artificial Intelligence in Education, vol. 35, pp. 304–329, 2025, doi: 10.1007/s40593-024-00400-6

[15] M. Mejeh and M. Rehm, "Taking adaptive learning in educational settings to the next level: Leveraging natural language processing for improved personalization," Educational Technology Research and Development, vol. 72, pp. 1597–1621, 2024, doi: 10.1007/s11423-024-10345-1

[16] H. Farhood, M. Nyden, A. Beheshti, and S. Muller, "Artificial intelligence-based personalised learning in education: A systematic literature review," Discover Artificial Intelligence, vol. 5, art. no. 331, 2025, doi: 10.1007/s44163-025-00598-x

[17] Hariyanto, F. X. D. Kristianingsih, and R. Maharani, "Artificial intelligence in adaptive education: A systematic review of techniques for personalized learning," Discover Education, vol. 4, art. no. 458, 2025, doi: 10.1007/s44217-025-00908-6

[18] N. Kerimbayev, K. Adamova, R. Shadiev, and Z. Altinay, "Intelligent educational technologies in individual learning: A systematic literature review," Smart Learning Environments, vol. 12, art. no. 1, 2025, doi: 10.1186/s40561-024-00360-3

[19] A. Nurjanah, A. Solehudin, and A. Primajaya, "Sistem Pendukung Keputusan Penentuan Model Pembelajaran Untuk Guru Dengan Metode Analytical Hierarchy Process (Studi Kasus: SMK PGRI Telagasari)," Jurnal Ilmiah Wahana Pendidikan, vol. 8, no. 17, pp. 32–45, 2022, doi: 10.5281/zenodo.7069500

[20] H. A. Farizi, H. K. Saputra, Y. Hendriyani, and K. Budayawan, "Implementasi Sistem Rekomendasi Pemilihan Mata Pelajaran Pada Kurikulim Merdeka Menggunakan Metode TOPSIS," Jurnal Pendidikan Tambusai, vol. 9, no. 2, pp. 27627–27636, 2025, doi: 10.31004/jptam.v9i2.31314

[21] C. Troussas, A. Krouska, P. Mylonas, and C. Sgouropoulou, "Personalized Instructional Strategy Adaptation Using TOPSIS: A Multi-Criteria Decision-Making Approach for Adaptive Learning Systems," Information, vol. 16, no. 5, art. no. 409, 2025, doi: 10.3390/info16050409

[22] T. L. Saaty, "Decision making with the analytic hierarchy process," International Journal of Services Sciences, vol. 1, no. 1, pp. 83–98, 2008, doi: 10.1504/IJSSCI.2008.017590

[23] O. S. Vaidya and S. Kumar, "Analytic hierarchy process: An overview of applications," European Journal of Operational Research, vol. 169, no. 1, pp. 1–29, 2006, doi: 10.1016/j.ejor.2004.04.028

[24] A. Ishizaka and A. Labib, "Review of the main developments in the analytic hierarchy process," Expert Systems with Applications, vol. 38, no. 11, pp. 14336–14345, 2011, doi: 10.1016/j.eswa.2011.04.143

[25] H. Pashler, M. McDaniel, D. Rohrer, and R. Bjork, "Learning styles: Concepts and evidence," Psychological Science in the Public Interest, vol. 9, no. 3, pp. 105–119, 2008, doi: 10.1111/j.1539-6053.2009.01038.x

[26] P. A. Kirschner, "Stop propagating the learning styles myth," Computers & Education, vol. 106, pp. 166–171, 2017, doi: 10.1016/j.compedu.2016.12.006

[27] V. Clinton-Lisell and C. Litzinger, "Is it really a neuromyth? A meta-analysis of the learning styles matching hypothesis," Frontiers in Psychology, vol. 15, art. no. 1428732, 2024, doi: 10.3389/fpsyg.2024.1428732

[28] A. Touloumakos et al., "Visual Type? Not My Type": A Systematic Study on the Learning Styles Neuromyth Employing Frequentist and Bayesian Statistics," Mind, Brain, and Education, vol. 17, no. 3, pp. 197–208, 2023, doi: 10.1111/mbe.12380

[29] C. Hattan, E. Grossnickle Peterson, and K. Miller, "Revising teacher candidates' beliefs and knowledge of the learning styles neuromyth," Contemporary Educational Psychology, vol. 77, art. no. 102269, 2024, doi: 10.1016/j.cedpsych.2024.102269

[30] B. A. Rogowsky, B. M. Calhoun, and P. Tallal, "Matching Learning Style to Instructional Method: Effects on Comprehension," Journal of Educational Psychology, vol. 107, no. 1, pp. 64–78, 2015, doi: 10.1037/a0037478

[31] D. T. Willingham, E. M. Hughes, and D. G. Dobolyi, "The Scientific Status of Learning Styles Theories," Teaching of Psychology, vol. 42, no. 3, pp. 266–271, 2015, doi: 10.1177/0098628315589505

[32] P. M. Newton, "The Learning Styles Myth is Thriving in Higher Education," Frontiers in Psychology, vol. 6, art. no. 1908, 2015, doi: 10.3389/fpsyg.2015.01908

[33] L. R. Aiken, "Three Coefficients for Analyzing the Reliability and Validity of Ratings," Educational and Psychological Measurement, vol. 45, no. 1, pp. 131–142, 1985, doi: 10.1177/0013164485451012

[34] J. R. Lewis, "IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use," International Journal of Human–Computer Interaction, vol. 7, no. 1, pp. 57–78, 1995, doi: 10.1080/10447319509526110

[35] M. Bond et al., "A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour," International Journal of Educational Technology in Higher Education, vol. 21, art. no. 4, 2024, doi: 10.1186/s41239-023-00436-z

[36] S. K. Banihashem et al., "A systematic mapping review at the intersection of artificial intelligence and self-regulated learning," International Journal of Educational Technology in Higher Education, vol. 22, art. no. 50, 2025, doi: 10.1186/s41239-025-00548-8

[37] M. Roski, R. Sebastian, R. Ewerth, A. Hoppe, and A. Nehring, "Learning analytics and the Universal Design for Learning (UDL): A clustering approach," Computers & Education, vol. 214, art. no. 105028, 2024, doi: 10.1016/j.compedu.2024.105028

[38] Y. Liu, X. Hu, J. T. D. Ng, Z. Ma, and X. Lai, "Ready or not? Investigating in-service teachers' integration of learning analytics dashboard for assessing students' collaborative problem solving in K–12 classrooms," Education and Information Technologies, vol. 30, pp. 1745–1776, 2025, doi: 10.1007/s10639-024-12842-5

[39] A. Riedmann, P. Schaper, and B. Lugrin, "Reinforcement Learning in Education: A Systematic Literature Review," International Journal of Artificial Intelligence in Education, vol. 35, pp. 2669–2723, 2025, doi: 10.1007/s40593-025-00494-6

[40] M. Mejeh, L. Sarbach, and T. Hascher, "Effects of adaptive feedback through a digital tool – a mixed-methods study on the course of self-regulated learning," Education and Information Technologies, vol. 29, 2024, doi: 10.1007/s10639-024-12510-8

[41] F. Almasri, "Exploring the Impact of Artificial Intelligence in Teaching and Learning of Science: A Systematic Review of Empirical Research," Research in Science Education, vol. 54, pp. 977–997, 2024, doi: 10.1007/s11165-024-10176-3

[42] O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education – where are the educators?," International Journal of Educational Technology in Higher Education, vol. 16, art. no. 39, 2019, doi: 10.1186/s41239-019-0171-0

[43] A. Fuchs, A. Radkowitsch, and D. Sommerhoff, "Using learning progress monitoring to promote academic performance? A meta-analysis of the effectiveness," Educational Research Review, vol. 46, art. no. 100648, 2025, doi: 10.1016/j.edurev.2024.100648