

2025 International Conference on Fiscal Management, Holistic Social Studies (FMHSS 2025)

Article

Systemic Reform Driven by Educational Big Data: A Mixed Study of Multimodal Learning Analytics and Evidence-Based Decision Making

Shuo Sun^{1,*}

¹ Graduate School of Information and Communication, Ajou University, Suwon, Korea

* Correspondence: Shuo Sun, Graduate School of Information and Communication, Ajou University, Suwon, Korea

Abstract: With the rapid advancement of educational big data and learning analytics technologies, data-driven systemic transformation in education has become increasingly viable and impactful. As digital learning environments proliferate globally, the sheer volume of data generated offers unprecedented opportunities to understand and optimize the learning process. This study explores how multimodal learning analytics (MMLA) can support evidence-based decision making in educational transformation efforts. Using a comprehensive mixed-methods approach, we systematically analyze quantitative data extracted from digital learning platforms alongside in-depth qualitative insights gathered from educators, administrators, and policymakers. The integration of these diverse data sources allows for a holistic understanding of the complex educational ecosystem. The findings highlight how MMLA enables more accurate learner profiling, facilitates real-time feedback mechanisms, and promotes highly personalized instruction tailored to individual student needs. Furthermore, the empirical evidence generated through this approach plays a crucial role in informing strategic policy adjustments and driving systemic improvements across educational institutions. By bridging the gap between raw data and actionable pedagogical strategies, this research provides a robust, practical framework for integrating data-intensive tools into institutional decision-making processes. Ultimately, the proposed framework seeks to significantly enhance teaching effectiveness, boost learner engagement, and foster greater educational equity by ensuring that interventions are precisely targeted and empirically supported. This study underscores the transformative potential of combining advanced analytics with pedagogical expertise to shape the future of modern education.

Keywords: educational big data; learning analytics; systemic transformation; data-driven education; educational equity

Received: 02 May 2026

Revised: 13 June 2026

Accepted: 25 June 2026

Published: 01 July 2026



Copyright: © 2026 by the authors.

Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Research Background and Problem Statement

In the context of ongoing digital transformation, the education sector is experiencing a pivotal transition from relying solely on "big data-driven" approaches to embracing "data-governed decision-making" frameworks [1]. Historically, educational data has been predominantly confined to structured formats, such as grades and attendance records, which provide a limited perspective on the multifaceted nature of the learning process. However, advancements in technologies, including sensors, learning management systems (LMS), and speech and video recognition, have enabled the collection and analysis of educational big data in ways that are increasingly multimodal, real-time, and

comprehensive. This evolution is fostering significant advancements in the methodologies and strategies employed within educational systems, paving the way for more informed and adaptive decision-making processes.

1.2. Theoretical Foundation and Practical Significance

Multimodal Learning Analytics (MMLA) integrates information from diverse sources, including language, images, and physiological data, enabling a comprehensive visualization and modeling of learning processes [2, 3]. This approach facilitates a deeper understanding of how learners interact with educational content and environments. Evidence-Based Decision Making emphasizes the importance of systematic data analysis over reliance on empirical assumptions, ensuring that educational strategies are grounded in objective reasoning. The integration of these methodologies not only enhances the effectiveness of educational strategies but also supports the scientific advancement of management frameworks, fostering more informed and adaptive practices in the educational domain.

1.3. Research Objectives and Content

This study seeks to investigate methods for integrating multimodal learning analytics with evidence-based decision-making within the framework of educational big data, aiming to achieve substantial advancements in educational practices. The research focuses on constructing a comprehensive framework for the collection and analysis of multimodal data, examining practical pathways for implementing data-driven decision-making in both teaching and administrative contexts, and proposing a combination of qualitative and quantitative research methodologies to support systematic improvements in educational systems [3]. These efforts are intended to provide actionable insights for enhancing the effectiveness and efficiency of educational processes.

2. Theoretical and Practical Framework of Multimodal Learning Analytics

2.1. Definition and Classification of Multimodal Data

Multimodal data encompasses diverse information sets derived from multiple sources and presented in various forms [2]. Within the domain of educational research and learning analytics, such data is collected from a wide range of inputs, including digital activity logs from online learning platforms, eye-tracking metrics recorded during classroom activities, video analyses of facial expressions, variations in speech patterns, handwriting movements, and operational trajectories. Additionally, multimodal data incorporates formats such as images, textual content, audio recordings, and physiological signals. These data types collectively provide insights into learners' evolving cognitive states, such as levels of understanding or confusion, emotional responses like interest or unease, and behavioral attributes, including focus and engagement. Based on the nature and purpose of data collection, multimodal data can be broadly categorized into perception-based types, such as eye movements and facial expressions; behavior-based types, including clickstream data and handwriting patterns; language-based types, such as spoken and written communication; and physiological types, including metrics like heart rate and brain activity. This classification enables a comprehensive exploration of the learning process, offering valuable perspectives for understanding and enhancing educational outcomes.

2.2. Techniques for Multimodal Data Fusion and Analysis

The fusion processing of multimodal data represents a critical step in advancing effective learning analytics [4]. Mainstream fusion techniques are generally categorized into three primary types. Early Fusion, also known as Feature-level Fusion, involves combining raw features from various modalities during the preprocessing phase. This approach is particularly effective for fine-grained data modeling, such as integrating eye movement trajectories with mouse movement paths to analyze and predict learning challenges. Late Fusion, or Decision-level Fusion, focuses on merging results derived from independently modeled modalities. This is typically achieved through methods like

weighted averaging or voting, making it more suitable for scenarios where maintaining the distinct characteristics of each modality is essential. Intermediate Fusion, often referred to as Hybrid Fusion, operates between the stages of feature extraction and decision-making. This method facilitates deeper semantic alignment while preserving the unique attributes of each modality. With the rapid advancements in artificial intelligence technologies, the application of deep learning models, including convolutional neural networks and recurrent neural networks, has become increasingly prominent in educational data analysis. Additionally, Graph Neural Networks (GNNs) are being utilized to capture complex structural and temporal relationships within multimodal data. These advanced methodologies significantly enhance the ability to model intricate learning behaviors, offering robust technical support for tasks such as analyzing learning pathways and forecasting learning outcomes.

2.3. Applications of Multimodal Learning Analytics in Educational Scenarios

Multimodal learning analytics technology demonstrates significant potential for application across diverse educational scenarios. For instance, in the analysis of classroom interactions, data such as students' speech intonation, facial expressions, and participation frequency are integrated to identify areas of active engagement and zones of minimal interaction within the classroom environment. In the context of monitoring immersive learning experiences, data collected from virtual reality devices, including eye movement patterns and head motion trajectories, are combined with physiological responses to evaluate levels of immersion and cognitive load [5]. Similarly, in studies focusing on writing and reading behaviors, features such as typing rhythm, mouse movement trajectories, and eye fixation durations are utilized to gain insights into learners' writing strategies and information processing methods. Moreover, multimodal analysis serves as a valuable tool for supporting instructional design and tailoring personalized learning recommendations. By quantifying engagement levels, mapping attention distribution, and analyzing decision-making pathways, this approach enables educators to develop more precise and effective teaching strategies. As advancements in multimodal data collection and processing technologies continue, their applications in areas such as educational assessment, learning diagnostics, and intelligent tutoring systems are expected to expand, offering even greater potential for enhancing educational outcomes.

3. Evolution of Evidence-Based Decision-Making Mechanisms in Educational Management

3.1. Basic Logic of Evidence-Based Decision Making

Evidence-Based Decision Making is a methodology that emphasizes the use of scientific and systematic approaches to enhance decision-making processes. At its core, this approach relies on verified data and research outcomes as the foundation for decisions, moving away from subjective reliance on personal experience, intuition, or isolated incidents. In the context of educational policy formulation and instructional adjustments, this methodology introduces quantifiable and verifiable data to ensure that decisions are more rational, scientific, and traceable. Particularly in complex and dynamic educational environments, evidence-based thinking enhances the relevance and effectiveness of proposed measures. For instance, prior to implementing a specific teaching strategy, conducting experimental studies or empirical analyses to validate its effectiveness can help prevent unnecessary resource expenditure and potential inefficiencies [6]. Furthermore, evidence-based decision-making incorporates a "continuous feedback" mechanism, which involves ongoing evaluation of implementation outcomes through real-time data collection. This iterative process allows for the optimization of policy details, creating a sustainable cycle of "data---practice---revalidation," thereby fostering continuous improvement in educational management practices.

3.2. Data Barriers and Trust Crisis in Educational Decision Making

Although the concept of evidence-based decision-making is gaining traction among educators, the practical application of educational data continues to encounter significant obstacles. One major challenge is the presence of data barriers. On one hand, existing mechanisms for data collection within educational systems are often incomplete, characterized by issues such as missing information and uneven sampling practices. These deficiencies result in analytical outcomes that fail to accurately represent the complexities of real-world scenarios. On the other hand, the absence of standardized definitions and methodologies for utilizing data indicators leads to inconsistent interpretations and measurements of the same concept, such as "learning engagement," across different institutions or departments. This lack of uniformity undermines the comparability and practical utility of the data. Furthermore, advanced analytical techniques, including deep learning models, often lack transparency and interpretability, making them inaccessible to non-technical stakeholders such as teachers and school administrators. This gap in understanding further complicates the translation of data insights into actionable policies. Another critical issue is the growing trust crisis surrounding data-driven approaches. Educators express concerns about the opaque nature of algorithmic models, fearing that non-transparent processes may result in biased decision-making or inequitable evaluation criteria. Additionally, heightened awareness of privacy protection has amplified apprehensions regarding the potential infringement of individual rights during data collection and processing. This is particularly evident in cases involving sensitive data types, such as physiological signals, facial expressions, and voice analysis, which often elicit strong objections from teachers and parents. These ethical and technical concerns directly affect the acceptance and feasibility of data-driven policies, necessitating solutions that prioritize technological transparency, stakeholder engagement, and the establishment of clear guidelines to address these challenges [5].

3.3. Designing Decision Support Systems Based on MMLA

Multimodal Learning Analytics offers robust technical support for evidence-based decision-making in education. Decision support systems leveraging this approach should prioritize the seamless integration and intuitive visualization of data. These systems can harness diverse data sources, including learning behaviors such as student interaction frequency and task completion rates, classroom environmental factors like lighting, noise levels, and seating arrangements, as well as teaching content characteristics such as textbook structure and difficulty levels. By synthesizing this information, "smart dashboards" can be developed to assist educators and administrators. These dashboards should feature capabilities such as real-time monitoring, dynamic alerts, and personalized recommendations. For instance, they can automatically detect patterns in students' attention distribution, identify comprehension challenges, and analyze participation trends during lessons. This enables teachers to receive immediate feedback, facilitating precise instructional adjustments tailored to students' needs. For school administrators, the system can aggregate and analyze data trends across various subjects, classes, or grades, providing valuable insights for resource allocation and evaluating teaching quality on a broader scale. To address potential trust issues among educators, the design of these systems should emphasize interpretability and ease of use. Complex analytical results should be translated into clear, actionable strategies through tools such as natural language generation or visual representations. Furthermore, stringent privacy protection measures, including data anonymization, access control, and transparent usage policies, must be implemented to ensure ethical operation. Ultimately, decision support systems grounded in multimodal analytics not only elevate the scientific rigor of educational management but also establish sustainable mechanisms for data-driven teaching practices, paving the way for intelligent education governance [7].

4. Design and Validation Pathways for Mixed Methods Research

4.1. The Necessity of Mixed Methods in Educational Data Research

In educational data research, relying on a single research paradigm often proves insufficient for capturing the multifaceted nature of teaching and learning processes. Quantitative research, which employs large-sample analysis and statistical modeling, is effective in identifying general patterns and trends, such as relationships between student performance and learning behaviors or the overall impact of specific teaching strategies [8]. However, this approach frequently overlooks individual variations, contextual influences, and dynamic changes within educational environments, making it challenging to fully understand the underlying reasons for certain behaviors or outcomes. By focusing primarily on broad trends, quantitative methods may fail to address the nuanced complexities that characterize real-world educational settings.

On the other hand, qualitative research offers a valuable lens for delving deeply into personal experiences, emotional perspectives, and interactive dynamics within educational contexts [9]. This approach is particularly adept at uncovering the mechanisms and processes that drive observed phenomena. However, due to its reliance on smaller sample sizes and its context-specific nature, qualitative findings often lack the generalizability needed to inform broader educational practices or policies. While it provides rich, detailed insights, its limited scope can restrict its applicability to diverse educational scenarios.

As a result, Mixed Methods Research (MMR) emerges as a crucial approach in educational data studies. By integrating the broad, overarching perspective of quantitative research with the detailed, context-sensitive insights of qualitative inquiry, MMR enables researchers to gain a more comprehensive understanding of educational phenomena. This methodology allows for the simultaneous exploration of large-scale patterns and individual-level intricacies, offering a balanced view that captures both the "forest" and the "leaves." For example, after identifying the effectiveness of a teaching strategy through statistical analysis, researchers can employ interviews or classroom observations to investigate the reasons behind its success, the specific groups of students who benefit most, and the conditions under which it is most effective. This dual approach not only enhances the explanatory depth and practical relevance of research but also provides more nuanced evidence to support decision-making in educational practices and instructional adjustments. By addressing the complex challenges inherent in educational systems, MMR serves as an indispensable tool for advancing both theoretical understanding and practical applications.

4.2. Case Study: A Mixed Methods Approach in a Vocational College

A case from a vocational college illustrates how mixed methods are applied in practice. In the redesign of a practice-oriented course, the institution utilized a Learning Management System (LMS) to gather data on students' online learning behaviors, including login frequency, assignment submissions, and participation rates. This was further complemented by physiological sensor data, which measured heart rate and galvanic skin response to assess students' emotional states. These combined data sources provided a comprehensive and objective understanding of student engagement and emotional alignment with the course content. Building on these quantitative findings, the research team conducted student group interviews and classroom observations to explore underlying factors. For instance, students who exhibited high engagement but low task completion were found to face challenges related to task comprehension or self-efficacy. Classroom observations revealed that the rapid pace of instruction caused difficulties for some students, leaving them unable to keep up [10]. By cross-validating the quantitative and qualitative data, the team implemented adjustments to the course structure and pace. Key concepts were refined, formative assessments and personalized support were introduced, and a more flexible feedback system was established. These modifications resulted in notable improvements in both student satisfaction and academic performance. Educators also expressed strong approval for how mixed methods research enhanced the adaptability and precision of teaching strategies. This case highlights the effectiveness of mixed methods in addressing the complexities of diverse learning environments,

particularly in vocational education, where practical and tailored approaches are essential for optimizing the learning process.

4.3. Mixed Methods as a Driver for Data Literacy and Teacher Development

Beyond instructional improvement, mixed methods research supports deeper teacher development and fosters a data-informed culture within educational organizations. By engaging in mixed methods research, teachers gain a more intuitive understanding of the data-driven logic that underpins student behavior and psychology. This process enables a shift from relying solely on experience-based teaching to adopting a more data-informed approach. Teachers begin to reflect critically on their instructional strategies and make adjustments based on empirical evidence, which enhances their ability to address diverse student needs effectively. This transformation can be described as a "data awakening," where educators develop heightened pedagogical sensitivity, improved analytical skills, and a commitment to continuous professional growth. Such a shift not only benefits individual teachers but also contributes to the overall quality of education within the institution.

Mixed methods research also promotes interdisciplinary and cross-role collaboration, fostering cooperative data processes among frontline teachers, instructional managers, and IT departments. This collaborative framework helps dismantle traditional barriers between technology and teaching, creating a more integrated and cohesive educational environment. For instance, teachers may acquire foundational skills in statistical analysis and data visualization, while technical staff become more actively involved in pedagogical discussions. This reciprocal exchange of knowledge and skills enhances the overall data literacy of the institution, enabling it to respond more effectively to emerging challenges. By bridging the gap between different roles, mixed methods research supports the development of a resilient and adaptive educational system that is better equipped to meet the demands of modern learning environments [11].

In essence, mixed methods research serves as more than just a methodological framework; it acts as a catalyst for transformative change in education. By integrating teaching, technology, and administrative processes through the lens of data, it provides a sustainable foundation for both teacher development and institutional innovation. This holistic approach ensures that educational organizations remain dynamic and responsive, fostering an environment where continuous improvement is not only encouraged but systematically supported [9].

5. Pathways and Strategies for Systemic Educational Reform

5.1. From Local Optimization to Systemic Upgrade

In the ongoing progression of educational digitalization, numerous initiatives remain concentrated on isolated technological applications, such as the installation of advanced classroom monitoring systems, the deployment of platforms for analyzing learning data, or the trial implementation of data-driven teaching methods in select courses. While these efforts provide valuable insights and serve as exploratory steps, they are often confined to specific contexts and fail to catalyze a comprehensive transformation of the broader educational framework. To achieve a genuine shift toward digital-intelligent education, it is essential to move beyond a "tool-centric" approach and adopt a perspective rooted in "system thinking." This approach emphasizes the importance of coordination across all components of the educational ecosystem and necessitates structural adjustments to ensure cohesive progress [11].

A comprehensive upgrade of the educational system should be guided by a closed-loop mechanism that integrates data collection, strategic planning, feedback analysis, and iterative redesign. This framework ensures that data-driven principles are seamlessly embedded across various domains, including teaching methodologies, curriculum development, assessment practices, teacher training, and institutional governance. For instance, in instructional settings, the use of multimodal learning analytics can provide real-time insights into students' engagement and comprehension, enabling educators to

tailor their teaching strategies to individual needs. At the curriculum level, advanced data analysis can identify common learning challenges and optimize content delivery. In terms of assessment, combining process-oriented data with outcome-based evaluations can create a more balanced and comprehensive approach to measuring student progress. Furthermore, in governance, data-driven insights can facilitate more efficient resource allocation and the development of policies that address actual needs. By dismantling information silos and fostering vertically integrated and horizontally coordinated data systems, the digitalization of education can evolve from isolated innovations to a holistic and systemic transformation.

5.2. Establishing Multi-Level Data Governance Structures

Achieving systemic improvements in education requires the establishment of robust data governance mechanisms. Many schools and regions encounter challenges such as possessing data but lacking the expertise to utilize it effectively, dealing with fragmented datasets that are difficult to integrate, or experiencing a misalignment between policy directives and real-world data applications. These issues often arise from the absence of well-defined data management frameworks and collaborative processes that can bridge these gaps.

To address these challenges, educational systems should implement a multi-tiered data governance structure with clearly delineated responsibilities. At the classroom level, educators must be equipped with the skills to interpret data and make informed instructional decisions. At the institutional level, schools should establish specialized data teams or centralized platforms to manage data aggregation, enforce standards, and provide technical support. At the policy level, mechanisms for integrating and sharing data across agencies and platforms are essential to ensure that decision-making processes are both timely and evidence-based. Additionally, it is crucial to define clear data access rights and accountability boundaries to mitigate risks such as data silos, improper use, or inefficiencies caused by rigid digital systems.

The governance framework should also incorporate essential components such as rigorous data quality control measures, including validation protocols and strategies for handling missing values. Furthermore, workflows for regular data updates and maintenance should be established to ensure the reliability of the information [12]. Provisions for rapid response to technical or operational disruptions should also be included, serving as institutional safeguards to maintain the integrity and utility of data systems. These measures collectively contribute to the effective and sustainable use of high-quality data in educational contexts.

5.3. Enhancing Awareness of Educational Data Ethics and Trust in Technology

As technology and data become increasingly integrated into educational systems, fostering a robust ethical framework and trust-based environment is critical for ensuring sustainable advancements. The sensitive nature of learning data, which may include cognitive behaviors, emotional responses, and physiological signals, underscores the importance of addressing potential risks such as privacy violations, algorithmic biases, and improper data labeling. Mismanagement of such data could lead to significant ethical concerns and undermine confidence in technological applications within education.

In addition, the perception of being constantly "monitored" or "evaluated by algorithms" can create discomfort among teachers and students, potentially diminishing their willingness to engage with new initiatives. To mitigate these challenges, it is essential to establish and enforce comprehensive ethical standards and governance mechanisms for educational data. This includes promoting awareness and education on data ethics among educators, administrators, and technology developers, with a focus on legality, legitimacy, and transparency in data practices. Technical systems should be designed to ensure transparency and auditability, enabling traceability of algorithmic processes, clear boundaries for data usage, and mechanisms for reviewing decisions. Furthermore, a governance philosophy centered on human needs should be adopted, safeguarding students' rights to access, control, and delete their personal data. This approach positions

technology as a supportive tool for education rather than an authoritative force. By embedding technology within an ecosystem rooted in ethics, trust, and institutional backing, it becomes possible to achieve sustainable progress in enhancing educational equity, quality, and efficiency [12].

6. Conclusion

The development of big data in education is accelerating a transition from technology-assisted teaching toward more comprehensive and systematic pedagogical transformation. Grounded in multimodal learning analytics and evidence-based decision-making, this paper proposes a mixed-methods research framework that integrates data collection, analysis, interpretation, and responsive action across the full educational process. By incorporating multimodal data, the framework broadens the scope of observable learning behaviors, improves the precision of instructional diagnosis, and supports more informed decisions in teaching, assessment, and educational management. At the same time, mixed-methods inquiry enhances both explanatory depth and practical applicability, enabling quantitative patterns to be interpreted alongside contextual evidence. Future work should further strengthen interdisciplinary collaboration, methodological refinement, and technical integration so as to support the intelligent optimization of educational systems and promote a new stage of data-informed decision-making and evidence-driven educational improvement.

References

1. S. M. Asadzadeh and A. Azadeh, "An integrated systemic model for optimization of condition-based maintenance with human error," *Reliability Engineering & System Safety*, vol. 124, pp. 117-131, 2014.
2. E. S. Parris, M. J. Carey, and H. Lloyd-Thomas, "Feature fusion for music detection," in *Eurospeech*, Sep. 1999, pp. 2191-2194.
3. G. E. Marcus and M. B. MacKuen, "Anxiety, enthusiasm, and the vote: The emotional underpinnings of learning and involvement during presidential campaigns," *American Political Science Review*, vol. 87, no. 3, pp. 672-685, 1993.
4. D. Sargiotis, "Establishing a Data Governance Culture: Change Management and Leadership," in *Data Governance: A Guide*, Cham: Springer Nature Switzerland, 2024, pp. 391-418.
5. C. Navarretta and P. Paggio, "Classification of feedback expressions in multimodal data," in *Proceedings of the ACL 2010 Conference Short Papers*, Jul. 2010, pp. 318-324.
6. E. E. Jang, D. E. McDougall, D. Pollon, M. Herbert, and P. Russell, "Integrative mixed methods data analytic strategies in research on school success in challenging circumstances," *Journal of Mixed Methods Research*, vol. 2, no. 3, pp. 221-247, 2008.
7. A. Walker, M. Recker, M. B. Robertshaw, J. Osen, H. Leary, L. Ye, and L. Sellers, "Integrating technology and problem-based learning: A mixed methods study of two teacher professional development designs," *Interdisciplinary Journal of Problem-Based Learning*, vol. 5, no. 2, p. 7, 2011.
8. E. Morvant, A. Habrard, and S. Ayache, "Majority vote of diverse classifiers for late fusion," in *Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR)*, Berlin, Heidelberg: Springer Berlin Heidelberg, Aug. 2014, pp. 153-162.
9. V. Venkatesh, S. A. Brown, and H. Bala, "Bridging the qualitative-quantitative divide: Guidelines for conducting mixed methods research in information systems," *MIS Quarterly*, vol. 37, no. 1, pp. 21-54, 2013.
10. K. L. M. Ang, F. L. Ge, and K. P. Seng, "Big educational data & analytics: Survey, architecture and challenges," *IEEE Access*, vol. 8, pp. 116392-116414, 2020.
11. S. Ray and M. Saeed, "Applications of educational data mining and learning analytics tools in handling big data in higher education," in *Applications of Big Data Analytics: Trends, Issues, and Challenges*, Cham: Springer International Publishing, 2018, pp. 135-160.
12. P. Long and G. Siemens, "Penetrating the fog: Analytics in learning and education," *EDUCAUSE Review (Online)*, 2011.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of Publisher and/or the editor(s). Publisher and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.