Computers in Human Behavior Reports 19 (2025) 100735

Contents lists available at ScienceDirect

Computers in Human Behavior Reports

journal homepage: www.sciencedirect.com/journal/computers-in-human-behavior-reports

Design and evaluation of a GenAI-based personalized educational content system tailored to personality traits and emotional responses for adaptive learning

Wentao Hu [a,*], Zichen Shao [b]

[a] Zhejiang Police College, Hangzhou, 310053, Zhejiang, China
[b] Hangzhou Caihe Middle School, Hangzhou, 310020, Zhejiang, China

ARTICLE INFO

Keywords:
Adaptive learning
GenAI
Personality traits
Emotional response
Design-based research

ABSTRACT

This research integrates personality traits and emotional responses with GenAI to create personalized educational content. Using a design-based approach, the Psychologically-Aware Generative Education (PAGE) system was developed to adapt learning materials based on learners' Big Five personality profiles and real-time emotional feedback. Quasi-experimental testing with 200 university students demonstrated that PAGE significantly enhanced emotional satisfaction (4.4/5 vs 3.6/5, Cohen's d = 1.05) and learning engagement compared to traditional adaptive systems, with 22 % higher task completion rates (87.6 % vs 72.3 %) and 34 % increased study duration. The system successfully tailored content style, difficulty, and support mechanisms according to individual psychological characteristics. Content personalization was particularly effective for students with high neuroticism, reducing dropout rates by 48 % and negative emotions. This study provides empirical evidence that psychological adaptation in educational technology produces more engaging learning experiences than solely cognitive-based approaches, contributing design principles for developing psychologically-aware AI systems in education. These findings offer practical implications for educational institutions seeking to implement more humanized and culturally responsive technological solutions.

1. Introduction

As personalized learning becomes central to digital education, the rise of GenAI—especially large language models (LLMs)—offers new opportunities for generating learner-specific content (Guettala et al., 2024). While systems like Knewton and ALEKS adapt content based on cognitive performance, they often neglect psychological variables such as personality traits and emotional states, which play a crucial role in shaping learning behavior.

Research indicates that learners with higher extroversion tendencies may prefer collaborative learning activities, while individuals with high neuroticism are more prone to cognitive overload when facing setbacks (Jansz et al., 2022). Additionally, emotional states—such as curiosity and anxiety—not only regulate learning motivation but also directly influence knowledge internalization efficiency through emotion-cognition interaction mechanisms (Wu et al., 2023).

This "psychological blind spot" hinders existing systems from achieving true personalized adaptation. This highlights the urgent need to explore educational content generation frameworks that integrate psychological characteristics.

This study introduces the Psychologically Aware Generative Education (PAGE) paradigm, which combines Big Five Personality Traits with real-time emotional feedback to guide GenAI-based personalized content generation. Theoretically, this model extends traditional adaptive learning approaches by embedding psychological and affective dimensions into system design and broadens GenAI's role in education from content delivery to comprehensive learner adaptation.

Practically, PAGE offers a framework for generating content that aligns with learners' personality profiles (e.g., open learners receiving exploratory narratives) and responds to their emotional states (e.g., supportive prompts for neurotic learners). This dual adaptation enhances learner engagement and supports more effective personalized education.

Specifically, this study addresses three questions:

* Corresponding author.
E-mail address: wthu@zju.edu.cn (W. Hu).

https://doi.org/10.1016/j.chbr.2025.100735
Received 27 May 2025; Received in revised form 15 June 2025; Accepted 25 June 2025
Available online 3 July 2025

- How can GenAI be designed to generate psychologically adaptive educational content based on personality and emotional cues?
- Does such content improve learner engagement and emotional satisfaction compared to traditional systems?
- What design principles ensure both educational effectiveness and responsible use?

Drawing on existing theories, we hypothesize:

(1) GenAI content tailored to Big Five traits will be perceived as more personally relevant than standard adaptive materials; (2) Integration of real-time emotional feedback will increase emotional satisfaction and reduce negative experiences; (3) Learners high in neuroticism or openness will show stronger effects; (4) Combined personality-emotion adaptation will improve engagement by at least 20 % over cognitive-only systems.

The remainder of this paper is structured as follows. Section 2 reviews the literature on adaptive learning, personality-learning associations, and GenAI applications in education. Section 3 details the Design-Based Research (DBR) methodology, encompassing psychological trait dataset construction, GenAI strategies, and evaluation processes.

Subsequently, Section 4 presents the research findings, and Section 5 discusses their theoretical implications and limitations. The paper concludes by proposing design principles for educational technology developers, offering a roadmap for building "Warm AI-Ed Agents"—intelligent educational systems that integrate psychological awareness with adaptive learning. Through this structure, we aim to provide both theoretical foundations and practical guidelines for developing Psychologically Aware adaptive learning systems.

The PAGE system was implemented in an introductory statistics course required for all social science majors across the three participating universities. This course covered fundamental statistical concepts including descriptive statistics, probability theory, hypothesis testing, and basic inferential statistics.

We selected this subject area for several reasons. First, statistics courses typically exhibit high variance in student engagement and emotional responses, making them ideal for testing psychological adaptation. Second, the conceptual nature of statistics allows for multiple explanation approaches that can be tailored to different personality types. Third, the subject matter can be presented with varying degrees of abstraction, exploration, structure, and real-world applications, providing flexibility for personality-based adaptations. Fourth, the cross-disciplinary nature of statistics enabled recruitment of participants from diverse academic backgrounds.

2. Literature review

This Section reviews core theories and current research in four areas: (1) adaptive learning and personalized education, (2) Big Five Personality Traits in educational contexts, (3) emotional responses in learning processes, and (4) GenAI opportunities and challenges in education. The Section establishes the theoretical foundation for our study's innovative contributions.

2.1. Adaptive learning and personalized education: from cognitive adaptation to a whole-person perspective

Adaptive Learning Systems (ALS) are a central research focus in educational technology, aiming to dynamically adjust learning pathways, content presentation, and interaction strategies based on individual differences to optimize learning experiences and outcomes (Brusilovsky & Mill'an, 2007). Early ALS, such as Intelligent Tutoring Systems (ITS), primarily relied on fine-grained cognitive models (e.g., Knowledge Space Theory, Bayesian Knowledge Tracing) to assess learners' knowledge levels and recommend appropriate resources or exercises accordingly (Corbett & Anderson, 1994). Modern commercial ALS platforms (e.g., Knewton, ALEKS, Smart Sparrow) have further advanced these foundations by leveraging large-scale learning data, enabling more flexible content delivery and adaptive pacing (VanLehn, 2011).

However, as noted in the introduction, mainstream ALS primarily personalize learning along cognitive dimensions, such as adjusting content difficulty, complexity, and sequencing or recommending subsequent learning modules based on prior performance (Shute & Zapata-Rivera, 2012). While this cognition-centric paradigm has achieved significant success, it often overlooks non-cognitive factors that are equally crucial to the learning process. Learning is not merely a cold, knowledge-transfer process but a complex socio-psychological activity involving motivation, emotions, attitudes, and personality traits (Shemshack et al., 2021). A key limitation of existing research is the lack of integration of psychological characteristics, which may lead to misalignment between personalized recommendations and learners' intrinsic preferences, learning styles, or emotional needs, thereby constraining the learner-centred potential of these systems (Ouf et al., 2017). Recent studies have demonstrated that cultural diversity and individual differences significantly impact students' learning behavioural patterns in online environments (Tlili et al., 2021), further highlighting the need for more holistic adaptation approaches.

Thus, expanding the scope of personalized education beyond cognitive adaptation to incorporate psychological dimensions such as personality and emotions—a whole-person perspective—represents a critical direction for the next generation of adaptive learning systems.

2.2. The role of the Big Five Personality Traits in education: linking personality to learning behavior

To better understand individual differences among learners, personality psychology provides a valuable theoretical framework. The Big Five Personality Traits model—Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (or Emotional Stability)—is widely recognized for describing individual personality differences (McCrae & Costa, 1997). Extensive research has established strong correlations between these traits and various aspects of education:

- Openness: Learners high in openness tend to be curious, imaginative, and receptive to new ideas and experiences (Komarraju et al., 2011). They often prefer exploratory, unstructured learning tasks and are more inclined towards abstract theories and complex concepts.
- Conscientiousness: Highly conscientious learners are organized, self-disciplined, and achievement-oriented (Poropat, 2009). They demonstrate strong learning motivation, favor structured, goal-oriented materials, and excel at time and task management. Among the Big Five, conscientiousness is one of the strongest predictors of academic success.
- Extraversion: Extraverted learners are energetic, sociable, and stimulation-seeking (De et al., 2012). They often thrive in collaborative learning environments and class discussions but may be easily distracted, requiring dynamic and engaging content to maintain focus.
- Agreeableness: Learners with high agreeableness exhibit cooperative, trusting, and helpful tendencies (Graziano et al., 2007). They tend to perform well in team-based activities but may require additional support in competitive or critical-thinking-driven learning environments.
- Neuroticism (Low Emotional Stability): Learners high in neuroticism are more prone to anxiety, stress, and negative emotions (Chamorro-et al., 2003). They may struggle with academic challenges and setbacks, requiring greater emotional support and positive reinforcement. The stress levels of the learning environment significantly impact their learning outcomes.

Despite extensive research on these associations, several contradictions and knowledge gaps remain unresolved in the literature. First, while studies consistently link conscientiousness to academic achievement, there are notable inconsistencies regarding extraversion, with Western studies often finding negative correlations with academic performance (Poropat, 2009) while some East Asian studies report positive associations (Kim, 2020; won Kim, 2019). These cross-cultural differences suggest that personality-based adaptations must account for cultural context rather than applying universal design rules.

Second, most personality-learning research employs correlation designs with self-reported measures, raising questions about causality and potential confounding variables. Few studies have experimentally tested educational interventions specifically tailored to personality profiles. This methodological limitation creates uncertainty about which personality-based adaptations would be most effective in practice. In our design, we address this by empirically testing multiple adaptation strategies for each trait rather than assuming effects based solely on correlational findings.

Third, the existing literature tends to examine personality traits in isolation, neglecting potential interaction effects. For instance, high openness combined with high neuroticism might require different adaptations than high openness with low neuroticism. Our research advances beyond this limitation by considering trait combinations and their interactive effects on learning preferences.

Finally, systematically applying these insights to adaptive learning content design remains a challenge. Most existing studies focus on correlation analysis or macro-level pedagogical recommendations, lacking empirical exploration of how personality traits can be translated into concrete, dynamic, and personalized content generation strategies.

This study aims to bridge these gaps by leveraging GenAI to tailor educational content based on learners' Big Five personality profiles, creating truly personalized learning experiences that account for trait interactions, cultural considerations, and the need for empirical validation of adaptation strategies.

2.3. Emotional responses and learning: the dynamic process of emotion-cognition interaction

The role of emotion in learning has gained increasing attention in educational research. Learning is not a purely rational activity; it is intertwined with a range of emotional experiences, including curiosity, interest, enjoyment, confusion, anxiety, frustration, and even boredom (Chamorro-et al., 2003).

The Control-Value Theory of Achievement Emotions, proposed by Pekrun (Pekrun, 2006; Pekrun et al., 2002), provides a key framework for understanding emotions in academic settings. This theory suggests that a learner's perceived control over a task and its subjective value jointly shape their emotional experiences. Positive emotions tend to enhance motivation, deepen information processing, and foster creativity, while negative emotions have more complex effects that can either stimulate short-term effort or impair cognitive processes (Pekrun et al., 2011).

Emotion and cognition are deeply intertwined (Immordino et al., 2007). Emotional states not only regulate motivation and engagement but also directly influence cognitive resource allocation and information processing. Adaptive learning systems capable of recognizing and responding to learners' emotional states can enhance engagement and foster more empathetic learning experiences (Baker et al., 2010).

However, existing research on affective adaptation in learning systems mainly focuses on surface-level feedback adjustments or strategy prompts, rarely exploring how core learning content itself can be dynamically reshaped based on emotional states. This gap highlights the need for emotion-aware content generation frameworks, which this study seeks to address.

2.4. GenAI in education: opportunities, challenges, and research gaps

GenAI, particularly Large Language Models (LLMs) such as the GPT series, BERT, and their variants, has made groundbreaking progress in recent years, demonstrating powerful capabilities in natural language understanding, generation, and multimodal content creation (Brown et al., 2020; Vaswani et al., 2017). This brings unprecedented opportunities to the field of education. GenAI can rapidly produce vast and diverse educational resources, such as explanatory texts, exercises, case studies, dialogue scripts, and even teaching stories, theoretically enabling the creation of unique learning materials for each student (Binhammad et al., 2024; Kasneci et al., 2023). GenAI can power intelligent conversational agents (e.g., virtual tutors) that provide instant feedback, answer student questions, and guide inquiry-based learning, creating more interactive and immersive learning environments (Zawacki et al., 2019). GenAI can help teachers generate lesson plans, assessment criteria, and even preliminary grading of assignments, reducing teachers' workload and allowing them to focus more on higher-order teaching activities (UNESCO, 2023).

However, the application of GenAI in education also faces numerous challenges. The "hallucination" issue, where models may generate content that appears plausible but is factually incorrect or logically flawed, poses a significant risk in educational applications (Ji et al., 2023). Ensuring that generated content is not only informationally accurate but also aligns with effective teaching principles (e.g., constructivism, cognitive load theory) is essential to truly promote deep learning, rather than merely stacking information (Miao et al., 2023). GenAI models may learn and amplify social biases present in training data. This can lead to generated content that is unfair or stereotypical towards certain groups (Abid et al., 2021).

Crucially, although existing research has begun to explore the use of GenAI for personalized education (e.g., generating questions of varying difficulty based on knowledge levels), few studies have systematically integrated learners' deep psychological characteristics—particularly stable personality traits and dynamic emotional responses—into the core mechanisms of GenAI content generation. Current GenAI educational applications often remain "psychologically blind," failing to fully leverage their powerful generative capabilities to create content that truly addresses learners' inner needs and preferences.

2.5. Psychological adaptability in GenAI through prompt engineering and fine-tuning

Recent advances in GenAI techniques offer promising approaches for achieving psychological dimensions of content adaptability. Prompt engineering—the strategic design of instructions given to language models—enables these systems to generate content aligned with specific psychological variables without requiring extensive model architecture changes (Liu et al., 2023; Wang et al., 2022). This approach utilizes carefully crafted prompts that incorporate psychological principles and learner profiles to guide model outputs towards personalized content generation.

For instance, well-designed prompts can instruct models to generate explanations with varying levels of detail and reassurance based on a learner's neuroticism level, or to adjust the narrative style to match extraversion preferences. Reynolds & Zhang (Reynolds & Zhang, 2022) demonstrated that even simple prompt modifiers (e.g., "Write an explanation for a highly conscientious learner who prefers structured information") can yield outputs that align with psychological research on learning preferences. This approach offers flexibility and immediate implementation without requiring specialized model training.

For more sophisticated applications, fine-tuning techniques allow for systematic adaptation of base models to specific psychological dimensions. Kim et al. (Kim et al., 2023) showed that models fine-tuned on datasets containing paired examples of content variations matched to personality profiles achieved significantly higher alignment with psychological adaptation principles than prompt-only approaches. Such fine-tuning creates models that internalize the connection between psychological characteristics and appropriate content adjustments,

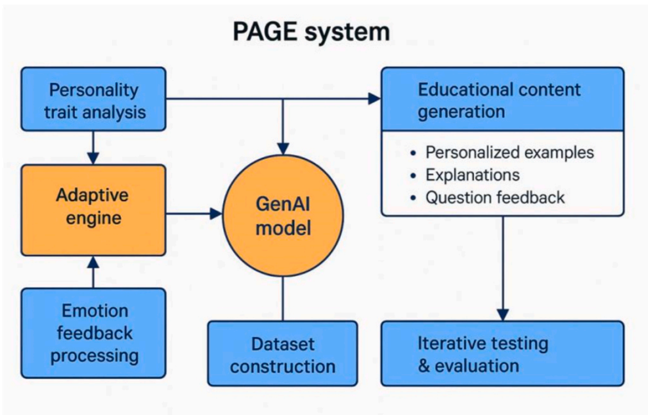

Fig. 1. The PAGE System Architecture Diagram illustrates the integrated workflow of personality trait analysis, emotion feedback processing, and GenAI content generation.

potentially improving generalization to novel situations.

The technical innovation of the PAGE framework lies in combining both approaches—utilizing carefully designed prompt templates that incorporate psychological variables while selectively applying fine-tuning to enhance the model's responsiveness to emotional signals and personality characteristics. This hybrid approach builds upon previous research while extending it through systematic integration of multiple psychological dimensions into a unified adaptability framework.

In summary, this study is grounded in the need for deeper personalization in adaptive learning, the profound insights from personality psychology and affective science on learner differences, and the robust technical support provided by GenAI. By integrating the Big Five personality traits and real-time emotional feedback data, we construct a GenAI-based personalized educational content generation model (the PAGE paradigm). This research aims to fill the existing research gap and explore a new path for psychologically and cognitively driven adaptive learning, thereby enhancing learning engagement, emotional satisfaction, and ultimate outcomes.

3. Methodology

This study aims to design, develop, and evaluate a GenAI-based personalized Adaptive Generation of Educational content system (named PAGE, Psychologically Aware Generative Education System). This system, as shown in Fig. 1, is capable of dynamically generating adaptive learning content based on learners' Big Five personality traits and real-time emotional responses. To achieve this goal and ensure a systematic and practice-oriented research process, this study employed the Design-Based Research (DBR) methodology (McKenney & Reeves, 2018; Reeves, 2006). DBR is a research paradigm that involves the systematic design, development, implementation, and iterative refinement of innovative educational interventions (in this case, the PAGE system) within real educational contexts. It aims to develop actionable design principles and generalizable theoretical knowledge, emphasizing the close integration of theory and practice, as well as the necessity of conducting research in complex, dynamic environments.

The DBR process in this study followed four main phases: Analysis & Exploration, Design & Construction, Evaluation & Reflection, and Implementation & Spread, focused on principles. Fig. 2 illustrates this iterative DBR cycle and how it was specifically implemented in our PAGE system development process, showing the key activities, feedback loops, and progression through each phase of the research.

3.1. Problem analysis and theoretical foundation construction

This phase aimed to deepen the understanding of the research problem, clarify design goals, and construct the foundational dataset required for the study.

Building on the literature review in Section 2, the specific manifestations of the "psychological blind spot" and its impact on learning experiences were validated and refined through a detailed review of the latest research in related fields (educational psychology, human-computer interaction, affective computing) and potential small-scale preliminary interviews or surveys (targeting potential learners or educators). This precisely defined the design requirements and expected goals of the PAGE system.

To train and evaluate the PAGE system, a dataset containing learners' psychological characteristics, interaction behaviors, and feedback was needed. Undergraduate students from universities were recruited, with a sample size of n = 200. Purposive sampling was used for recruitment. All participants signed an informed consent form before participating in the study, ensuring they understood the research purpose, process, data usage, and privacy protection measures. This study strictly adhered to ethical standards. Internationally recognized and reliable standardized scales were used for measurement, such as the Big Five Inventory-44 (BFI-44) (John & Srivastava, 1999) or NEO-FFI (Costa

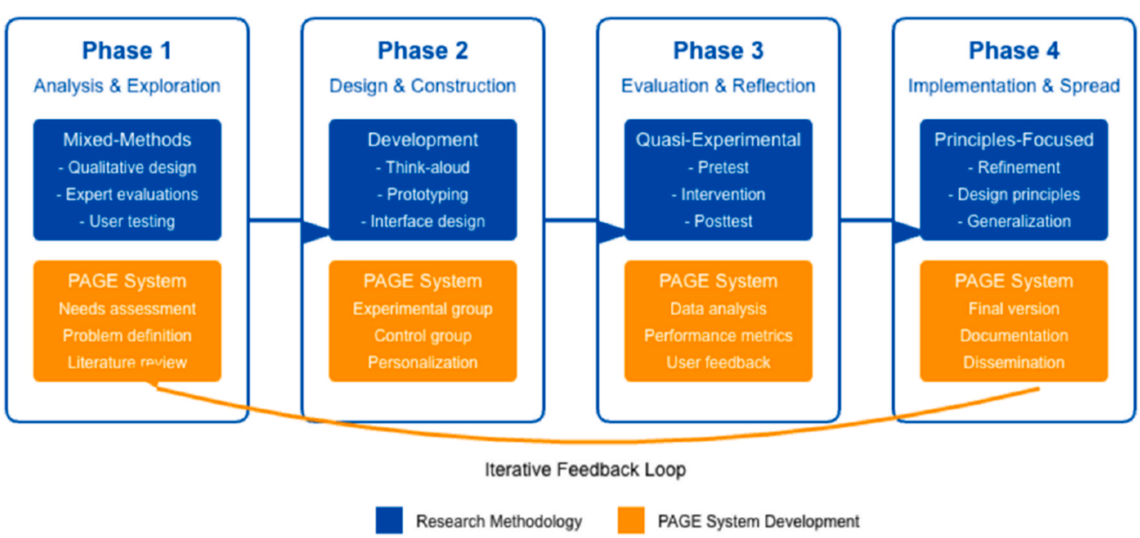

Fig. 2. The Design-Based Research (DBR) cycle implemented in the PAGE system development process, illustrating the four main phases and their interconnections with iterative refinement loops.

& McCrae, 1992). Participants completed the questionnaire before the study began to obtain their scores on the five dimensions: openness, conscientiousness, extraversion, agreeableness, and neuroticism. After completing learning modules or at key interaction points, learners were asked to report their current emotional state and satisfaction with the content by selecting predefined emotion labels (e.g., interested, confused, frustrated, satisfied) or rating on a Likert scale (e.g., 1–5 points from very dissatisfied to very satisfied). The Self-Assessment Manikin (SAM) (Bradley & Lang, 1994) was used to measure valence and arousal. Learners were encouraged to provide brief text feedback at specific points (e.g., when encountering difficulties or after completing tasks), describing their feelings, thoughts, or issues encountered. These texts were used for subsequent sentiment analysis and qualitative research.

The study received full approval from the university's Institutional Review Board and adheres to international ethical research standards. All data collection procedures were designed to protect participant privacy while gathering necessary information. Specifically, personal identifiers were separated from psychological and performance data using a coded system, with the linking keys stored in encrypted, password-protected databases accessible only to the principal investigators. Participants were informed that their data would be anonymized in reporting and were given the right to withdraw from the study at any time without penalty, with an option to have their data removed from analysis. Additionally, given the psychological nature of data collection, a protocol was established for referral to university counselling services should any participants experience distress during the study, though no such incidents occurred.

To ensure system reliability and validity, comprehensive evaluation criteria were established before deployment. Reliability assessments included: (1) technical stability testing (system uptime $>$ 99.5 %, response time $<$ 3 s); (2) content generation consistency ($>$90 % alignment between repeated prompts with identical parameters); and (3) inter-rater reliability for content evaluation (Cohen's Kappa $>$ 0.75). Validity measures encompassed: (1) content validity (expert assessment of alignment between generated materials and learning objectives); (2) construct validity (correlation between system-generated personality-based adaptations and theoretical predictions); and (3) criterion-related validity (relationship between system recommendations and actual learning outcomes). These metrics were continuously monitored throughout the study, with documentation of all system modifications and their impacts on reliability indicators.

The system automatically recorded learners' interaction behavior data to compute a comprehensive engagement index. This engagement index is calculated as a weighted composite score (ranging from 0 to 1) of the following metrics: completion rates of learning modules (30 %), average learning duration per session (20 %), page view depth (15 %), number of attempts and correctness rates of exercises (20 %), and frequency of active interactions with the system features (15 %). Higher values indicate greater engagement with the learning materials.

3.2. Solution design and prototype development (PAGE system)

Based on problem analysis, this study designed and constructed a GenAI model capable of integrating personality traits and emotional information to achieve personalized content generation. Additionally, a prototype of an adaptive learning system based on this model was developed.

In terms of the design and implementation of the GenAI model, the study first selected an appropriate foundational model, such as GPT-4 (via API calls) or open-source large models like Llama 2 and Mistral, and fine-tuned them according to actual needs. The criteria for model selection included the quality of generated content, contextual understanding, controllability, API availability, and cost. To ensure the model effectively adapted to the psychological characteristics of different learners, the study further designed a personalized generation mechanism, including prompt engineering and fine-tuning strategies. On the input side, the model received structured prompt information, which included the learner's personality profile (e.g., "high openness, moderate conscientiousness, low neuroticism"), current emotional state (e.g., "feeling confused" or "satisfaction score 4/5"), learning goals, and core knowledge points to ensure that the generated learning materials matched the learner's personality traits and immediate needs.

For the personalized design of output content, the model generated educational resources in various forms, including personalized narratives or cases (e.g., adjusting the exploratory and complexity of stories based on openness, or adjusting cooperation and conflict elements in narratives based on agreeableness), customized explanations (e.g., providing more structured, step-by-step explanations for highly conscientious learners, while using more vivid, conversational language for extraverted learners), and adaptive questions and feedback (e.g., adjusting the challenge level of questions based on neuroticism and dynamically adjusting question difficulty or providing additional hints based on the learner's emotional feedback). If initial tests indicated that prompt engineering alone could not reliably achieve the desired personalization and pedagogical alignment, the study collected or constructed small-scale "(psychological traits + emotions + learning goals) - > personalized educational content" sample pairs to fine-tune the selected LLM, making it more aligned with the specific needs of this study. Furthermore, to ensure content quality, the study established corresponding control mechanisms to ensure the factual accuracy of generated content (e.g., through knowledge base retrieval or expert review), pedagogical effectiveness (in line with basic instructional design principles), and ethical appropriateness (avoiding bias and inappropriate content). Therefore, the system needed to design post-processing rules or introduce manual review processes to ensure high-quality content output.

In the design of the adaptive learning system prototype, the study developed a web-based prototype system, whose architecture included a user interface, a learner model module (for storing and updating demographic information, personality traits, emotional states, and learning progress data), an adaptive engine (which calls GenAI to generate personalized learning content based on the learner model and preset rules or algorithms), a content presentation module, and a data recording module. The system's design philosophy incorporated constructivist learning theory, emphasizing learners' active exploration and meaning construction, aiming to support different learning preferences (related to personality traits) and regulate learning emotions through personalized content to promote deeper learning. For example, the system provided different levels of exploration space based on the learner's openness; offered structured learning path options based on conscientiousness levels; generated simulated discussion or collaboration task prompts based on extroversion characteristics to meet individualized learning needs.

In terms of adaptive logic, the study defined a series of explicit rules and algorithms to dynamically adjust learning content. For instance, when the system detected that "the learner's neuroticism score is above a certain threshold and recent emotional feedback shows 'frustration'," it automatically called GenAI to generate more soothing, detailed explanations and provide encouraging feedback; for learners exhibiting strong curiosity, the system prioritized providing exploratory tasks or case analyses to enhance their learning engagement. Through this mechanism, the system could adapt in real-time to changes in individual psychological states during the learning process, thereby enhancing learner engagement and satisfaction.

3.3. GenAI model implementation and prompt engineering

This study selected GPT-4 as the foundational model and implemented personalized content generation through carefully designed prompt engineering strategies, as shown in Fig. 3. Below are examples of key prompt templates used for different personality traits:

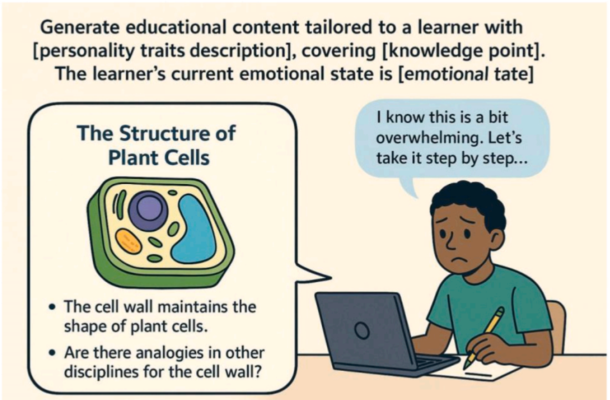

Fig. 3. Emotion- and Personality-Driven personalized Learning Experience.

- For high openness learners: "Generate an exploratory case study on [topic] with multiple perspectives and interdisciplinary connections. Include open-ended questions that encourage creative thinking. Current emotional state: [emotion]."
- For high conscientiousness learners: "Create a structured learning module on [topic] with clear objectives, step-by-step explanations, and progress checkpoints. Provide organized summaries after each section. Current emotional state: [emotion]."
- For high extraversion learners: "Develop an interactive scenario about [topic] with opportunities for discussion and collaboration. Use conversational language and incorporate social elements. Current emotional state: [emotion]."
- For high neuroticism learners experiencing anxiety: "Explain [topic] using gradual, scaffolded steps with encouraging feedback. Anticipate potential points of confusion and provide reassurance. Emphasize that mistakes are part of learning. Current emotional state: anxiety."

These prompts were dynamically combined with real-time emotional feedback to adjust tone, complexity, and support level. For instance, when a highly neurotic learner reported frustration, the system would generate content with more detailed explanations, supportive language, and confidence-building elements.

3.4. Multi-round iterative testing and evaluation

This study employed a Design-Based Research (DBR) methodology involving testing the PAGE system prototype in authentic contexts to gather feedback and iteratively refine the design. This process comprised three stages: expert review, user testing, and implementation validation.

The Expert Review stage involved 3–5 experts from educational technology, psychology, relevant subject domains, and AI ethics evaluating the system. Using methods such as Heuristic Evaluation and structured interviews, they assessed the pedagogical effectiveness, personalization appropriateness, usability, and potential risks. Based on this feedback, we revised the system's design logic and quality standards. Each stage builds upon findings from the previous phase, creating an iterative refinement cycle that progressively enhances the system's design and functionality.

The User Testing stage involved 10–15 learners testing the PAGE system in a simulated environment. Using Think-Aloud Protocol and semi-structured interviews, we collected feedback on system usability, content personalization, and emotional experiences. This feedback helped refine the GenAI model's personalization strategies and adaptive engine rules.

Table 1 documents the specific problems identified during each iteration and our corresponding system modifications. For instance, expert reviewers in the first iteration identified that initial prompt templates produced content with insufficient emotional support for high-Neuroticism learners, leading us to redesign prompt structures with explicit validation phrases. In the second user testing phase, participants found that content for high conscientiousness learners was sometimes perceived as excessively rigid, prompting modifications to incorporate modest exploratory elements even in structured formats.

Table 1
Key problems and modifications during DBR iterations.

Iteration	Identified Issues	System Modifications
Expert Review 1	• Content for high-Neuroticism lacked emotional support • Factual errors in complex content • Inconsistent style for trait combinations	• Redesigned prompt templates • Added knowledge verification system • Created hierarchical prompt structure
User Testing 1	• Emotional interface too intrusive • Exploration content lacked structure • personalization delays disrupted engagement	• Redesigned emotion reporting • Balanced exploration with organization • Added content pre-generation
Expert Review 2	• Emotional adaptation overrode learning goals • Privacy concerns for emotional data • Adaptation lacked transparency	• Added pedagogical constraints • Implemented data anonymization • Added adaptation notifications
User Testing 2	• Content for high conscientiousness too rigid • Failed to address context-specific frustrations • Misinterpreted engagement metrics	• Added exploratory elements to structured content • Created content-specific support templates • Refined contextual engagement metrics

The Implementation and Validation stage employed a Quasi-Experimental Design comparing an Experimental Group (using the complete PAGE system) and a Control Group (using standardized content or cognitive-based adaptation). The study was conducted in an authentic learning environment over several weeks. We compared outcomes on key metrics including emotional satisfaction, learning engagement, and learning outcomes to assess whether the PAGE system's personalization strategies enhanced the learning experience.

As shown in Fig. 5, the PAGE system interface was designed to facilitate both content delivery and emotional feedback collection, with components specifically tailored to support personalized learning experiences. The user interface integrated seamlessly with the backend adaptation engine, allowing for real-time content adjustments based on learner interactions and reported emotional states.

3.5. Data collection and analysis

Throughout the various stages of iterative testing and evaluation, particularly during Phase 3 of implementation verification, data were systematically collected and analyzed to assess the performance of the system.

The data collected can be categorized into three main types. Pre-test data includes demographic information, scores on the Big Five personality traits, and optionally, pre-test knowledge levels related to relevant academic disciplines. Process data consists of real-time or phase-specific emotional self-assessments, such as those measured through scales or labels, open-ended textual feedback from participants, and system-recorded indicators of learning engagement. These engagement indicators encompass metrics such as task completion time and rate, interaction frequency, and page dwell time. Post-test data consists of overall emotional satisfaction scores, a learning experience questionnaire evaluating subjective assessments of personalization, engagement, and learning outcomes, as well as optionally, post-test knowledge levels in relevant academic areas.

The analysis of the collected data involved both quantitative and qualitative methods. In terms of quantitative analysis, descriptive statistics were employed to calculate the means, standard deviations, and frequencies for all variables. Additionally, correlation and regression analyses were conducted to explore the relationships between personality traits, various types of emotional responses, and learning engagement and satisfaction. This analysis provided partial answers to Research Question 1, with further exploration in the discussion section. Pearson's correlation coefficient or Spearman's rank correlation, along with multiple linear regression models, were used to examine these relationships. To evaluate the effects of the intervention, independent samples t-tests or Mann-Whitney U tests were employed, depending on the data distribution, to compare post-test emotional satisfaction and learning engagement indicators between the experimental and control groups. When significant pre-test differences were observed or when controlling for variables such as pre-test knowledge levels or the direct influence of personality traits, Analysis of Covariance (ANCOVA) was applied to account for these factors, thus addressing Research Question 2.

For qualitative analysis, thematic analysis (Braun and Clarke, 2006) was used to analyze learners' open-ended text feedback and interview records. This process involved familiarizing oneself with the data, generating initial codes, searching for themes, reviewing and refining these themes, and ultimately defining and naming them. Thematic analysis also involved writing up the results. NVivo or other qualitative data analysis software was used to assist in managing and analyzing the data. The objective of this analysis was to gain a deeper understanding of learners' experiences with personalized content, including their perceptions of its strengths and weaknesses, the emotional changes they experienced, and their suggestions for improving system design. The key themes identified through this process helped explain the quantitative results, enhanced the understanding of the system's effectiveness, and provided valuable insights for refining the system's design principles.

Finally, the results from both the quantitative and qualitative analyses were integrated through a process known as triangulation. This approach provided a more comprehensive and nuanced interpretation of the findings, such as using qualitative interviews to explain satisfaction differences observed in the quantitative surveys. By combining both methods, the study was able to draw more robust and insightful conclusions.

3.6. Ethical considerations and data privacy

Given the psychological nature of this study and the collection of sensitive personal data, ethical considerations were prioritized throughout the research process. The study underwent rigorous ethical review and received full approval from the institutional review boards of all participating universities.

All data handling followed strict privacy protocols. We implemented end-to-end encryption for personality and emotional data during both storage and transmission. To address concerns regarding third-party API usage, we utilized a proxy server architecture that anonymized all learner data before API calls by removing personally identifiable information and replacing it with randomized identifiers. Additionally, we secured contractual agreements with our API provider, ensuring research data would not be retained or used for model training.

Informed consent procedures were comprehensive, with explicit disclosure of how psychological data would be used for personalization. Participants were provided with both opt-out options for specific data types and the ability to withdraw entirely from the study without penalty. During system usage, the interface-maintained transparency by briefly explaining personalization mechanisms (e.g., "This content has been adapted based on your personality profile to provide more exploratory material") and offering options to view alternative content formats.

We also addressed broader ethical implications of psychological profiling in educational contexts. To prevent potential psychological labeling effects, the system avoided explicit trait categorization in user-facing interfaces. To mitigate risks of creating filter bubbles based on personality traits, we designed the system to gradually introduce content outside a user's predicted preferences, expanding their exposure to diverse learning approaches. Regular ethical monitoring throughout the study included scheduled reviews of personalization effects across different demographic groups to identify and address any unintended biases or disparities in system effectiveness.

The Data Safety Monitoring Committee, comprising experts in psychology, educational technology, and data ethics, conducted fortnightly reviews of anonymized usage data to identify potential adverse effects. No significant adverse events were reported during the study, though three participants reported mild discomfort with emotion-tracking features and were immediately provided with alternative participation pathways as outlined in our ethical contingency protocol.

3.7. Summary of evaluation strategy

The evaluation strategy of this study is multi-dimensional and spans several stages.

Formative evaluation took place during the expert review and user testing phases. In these stages, both qualitative and quantitative feedback were used to continuously refine the system design and content generation strategies. Summative evaluation was conducted during the implementation verification phase, where a quasi-experimental design was employed to rigorously compare the learning outcomes under personalized (experimental group) versus non-personalized (control group) conditions. The focus of this evaluation was primarily on emotional satisfaction (as the primary metric), learning engagement (behavioural indicators), and learners' subjective experiences, which were captured through qualitative data. Additionally, the degree of personalization in the content generated by GenAI would be assessed, for instance, by having independent evaluators score the alignment between the content and user profiles. The educational value of the content would also be evaluated based on feedback from both experts and users. The goal was to answer the core research questions and provide empirical evidence for refining the design principles.

Through this systematic research approach, the study aimed not only to develop an effective prototype system but also to gain a deeper understanding of the mechanisms behind personalized content generation based on personality and emotions. Furthermore, the study sought to distil design principles that could serve as a reference for other researchers and developers in the field.

4. Results

This study adopted a Design-Based Research (DBR) approach to collect and analyze data regarding the relationship between learners' personality traits and emotional responses, as well as to evaluate the personalized learning content generated by a GenAI model and the overall performance of the PAGE (Psychologically Aware Generative Education) system. The results focused on three main aspects: (1) the associations between personality traits and emotional responses—and how these associations influence learning preferences; (2) the quality and educational value of the GenAI-generated content; and (3) the extent to which the PAGE system enhances learner engagement and emotional satisfaction.

4.1. Associations among personality traits, emotional responses, and learning preferences

A statistical analysis of baseline data was conducted, including learners' Big Five personality trait scores and emotional feedback, to verify whether significant associations exist between personality traits and emotional states or preferences during the learning process.

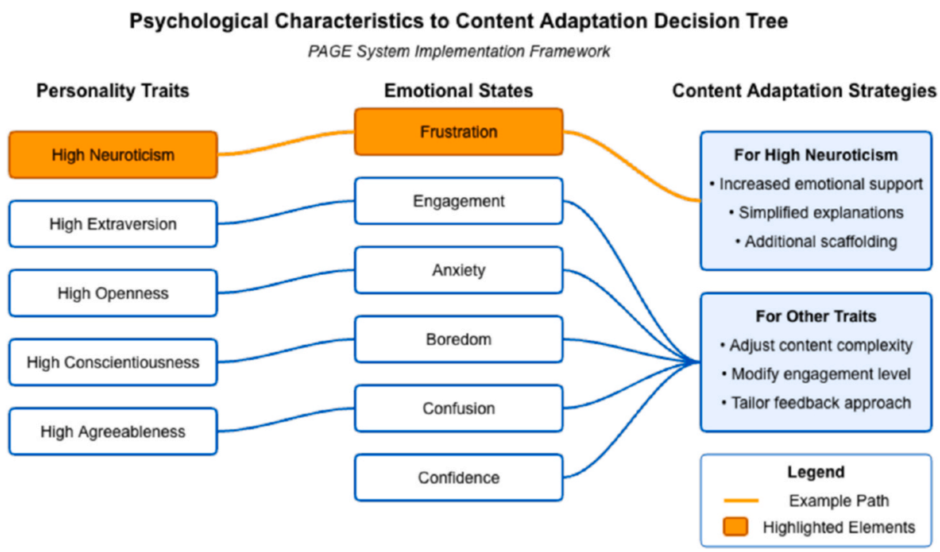

Fig. 4. Decision tree illustrating how different combinations of personality traits and emotional states trigger specific content adaptation strategies in the PAGE system. The diagram shows the primary adaptation pathways and their relationship to learning objectives.

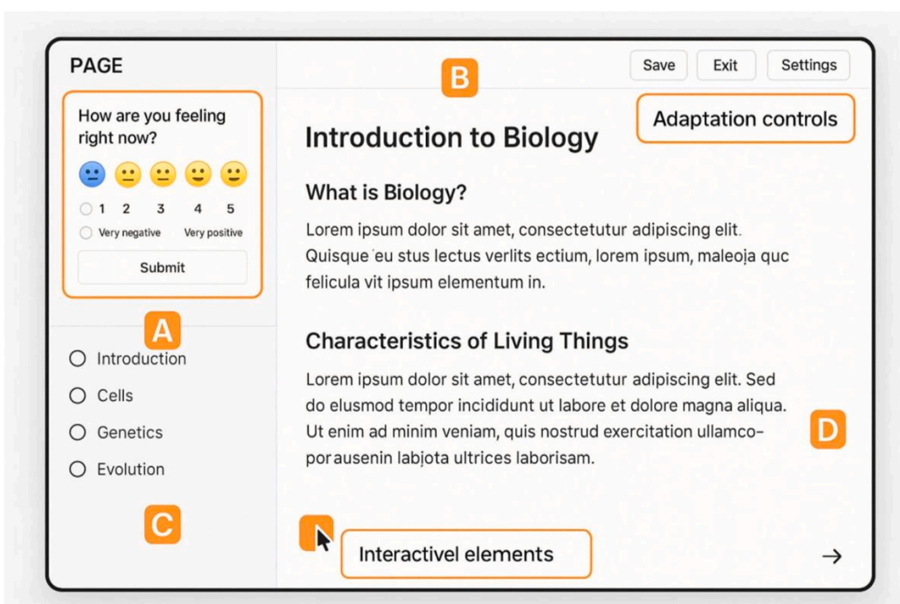

Fig. 5. The PAGE system user interface showing key components: (A) current learning module and progress indicators, (B) emotional feedback input mechanism allowing learners to report their affective state, (C) personalized content area dynamically generated based on personality profile and emotional state, (D) interactive elements tailored to personality traits, and (E) adaptive support panel with difficulty adjustment options.

Descriptive statistics showed that participants' scores on the five traits were generally consistent with established norms, though slight skewness was observed for extraversion and neuroticism.

Students were categorized as 'high' on specific personality dimensions using both normative and sample-specific approaches. Using the BFI-44 (John & Srivastava, 1999), raw scores for each dimension were first converted to standardized T-scores based on established norms for undergraduate populations. Participants scoring at or above the 70th percentile ($T > 55$) on a given trait were classified as 'high' on that dimension. This approach resulted in approximately 32 % of participants categorized as high in openness, 28 % high in conscientiousness, 35 % high in extraversion, 30 % high in neuroticism, and 33 % high in agreeableness.

During the learning sessions, learners frequently reported feeling "satisfied," "interested," or "confused," with an overall mean emotional satisfaction score of approximately 3.8 ($SD = 0.8$) on a 5-point scale.

Correlation analyses revealed meaningful relationships between personality traits and emotional responses or learning preferences. Extraversion correlated positively with positive emotions and enjoyment of interactive content. Conscientiousness was associated with higher post-task satisfaction and lower levels of disorganization. Neuroticism showed a positive correlation with anxiety and frustration, and a negative correlation with overall emotional satisfaction. Openness was linked to curiosity and engagement with exploratory or interdisciplinary content. Agreeableness was modestly related to preferences for collaborative learning. These findings supported the inclusion of psychological traits in content personalization and informed key design decisions in the PAGE system, as visualized in Fig. 4.

Contrary to the impression that might have been given in our results reporting, learners high in agreeableness were indeed included in the

Table 2
Correlation analysis of personality traits with learning experience.

Traits	Emotional Satisfaction	Learning Engagement	P - value
Extraversion	0.42	0.38	<0.01
Conscientiousness	0.56	0.61	<0.01
Openness	0.48	0.52	<0.01
Agreeableness	0.31	0.29	<0.05
Neuroticism	−0.45	−0.39	<0.01

Table 3
Content examples for different personality profiles.

Personality Type	Generated Content Example (Statistical Correlation)
High Openness	"Let's explore correlation across disciplines - from economics to psychology and ecology. Consider: How might correlation reveal patterns in climate data? What philosophical questions arise about causation vs. correlation? Try approaching from different theoretical frameworks … "
High Conscientiousness	"Lesson Objective: Master correlation concepts. 1. Definition: Correlation measures relationship strength/direction between variables. 2. Formula: Pearson's r = $\frac{\sum(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum(x_i-\overline{x})^2\sum(y_i-\overline{y})^2}}$ 3. Key Properties: Values from −1 to +1 4. Verification: Complete exercises 3.1–3.4."
High Neuroticism (frustrated)	"Many find correlation challenging at first - that's normal. Let's break it down: correlation simply measures how things change together. When one increases, does the other increase (positive) or decrease (negative)? Don't worry about formulas yet - focus on the basic idea. Confusion is part of learning, and we'll work at your pace."

full study and system design. The apparent underemphasis in our results discussion reflected our finding that agreeableness showed weaker correlations with adaptation preferences and learning outcomes ($r = 0.31$ and $r = 0.29$, respectively) compared to the other four dimensions. While still statistically significant ($p < 0.05$), these correlations were less pronounced than those observed for traits like conscientiousness ($r = 0.56$ and $r = 0.61$) and Openness ($r = 0.48$ and $r = 0.52$).

Design decisions included: (1) more emotional support for high-Neuroticism learners, such as real-time difficulty adjustments and encouraging feedback; (2) exploratory modules for high-Openness learners, dynamically adjusted by engagement data; (3) structured pathways with progression markers for high conscientiousness individuals; and (4) social or interactive elements for highly extraverted learners. These decisions were implemented via prompt engineering strategies to align content generation with learner profiles.

4.2. Quality assessment of GenAI-Generated content

GenAI serves as the core of the PAGE system by producing learning materials tailored to learners' psychological profiles. As shown in Table 2, the quality of the generated content was evaluated by examining personalization, educational value, and feedback from both experts and end-users, supplemented by independent evaluators' assessments of "learner profile–content" alignment. In terms of user perceptions, most learners in the experimental group felt that the system adequately adapted to their personality traits and emotional states. On a five-point Likert scale, the average rating for content personalization was around 4.2 (SD = 0.8), significantly above the neutral point of 3.0, suggesting that the system effectively matched presentation style, difficulty level, and emotional support with users' profiles. From the perspective of independent raters, the model generally performed well in modifying language or content style based on key traits such as Extraversion, Openness, or Neuroticism, achieving moderate to high inter-rater reliability (Cohen's Kappa).

Illustrative samples showed successful personalization: high-Openness learners received exploratory case studies; high-Neuroticism learners received step-by-step guidance with reassuring language; and high conscientiousness learners received structured content with clear objectives. Table 3 provides examples of how the same statistical concept was presented differently based on personality traits. Expert reviewers acknowledged the effectiveness of these approaches, though early content iterations showed factual inaccuracies due to GenAI limitations. In response, the research team refined prompts and implemented post-processing rules. These refinements reduced error rates and improved user satisfaction. Some participants found the content slightly verbose or occasionally misaligned, which highlights the need for ongoing iteration.

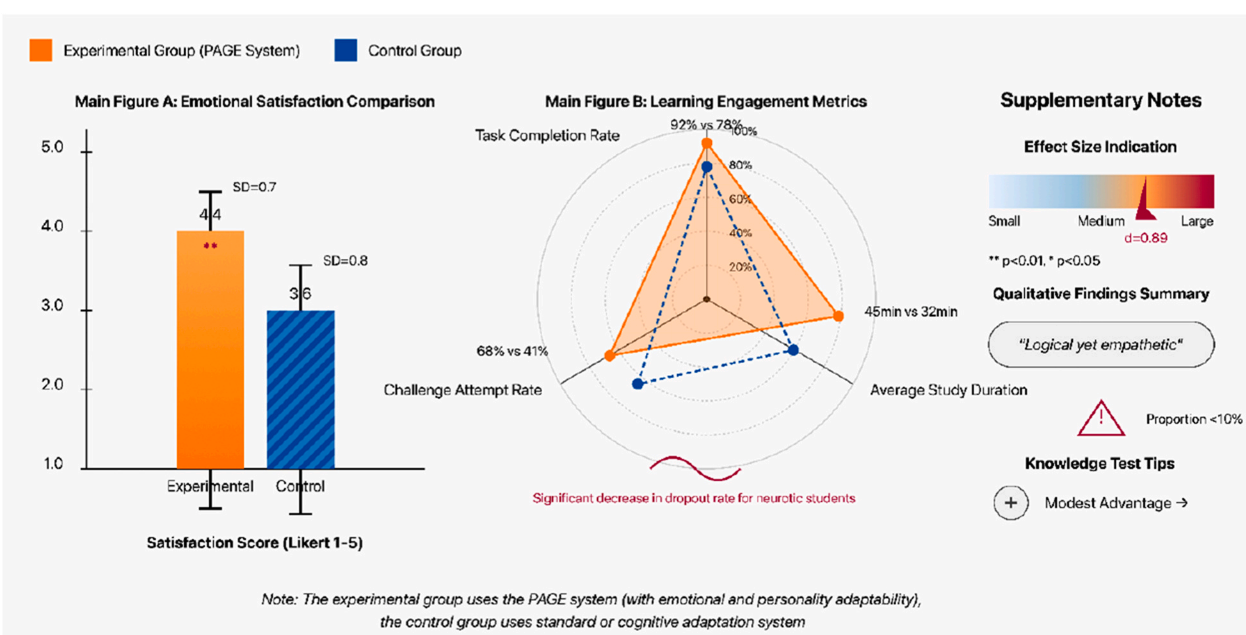

Fig. 6. Comparison of key evaluation metrics between the experimental group and control group.

Table 4
Comparison of key evaluation metrics between the experimental group and control group.

Evaluation Metric	Experimental Group (n = 100)	Control Group (n = 100)
Emotional Satisfaction (1–5 scale)	4.4 (SD = 0.7)	3.6 (SD = 0.8)
Overall Learning Engagement Index	0.78 (SD = 0.12)	0.63 (SD = 0.15)
Task Completion Rate Average Study Duration (minutes)	87.6 %	72.3 %
Voluntary Challenge Attempt Rate	42.3 (SD = 8.7) 64.8 %	31.5 (SD = 9.2) 38.2 %
Content personalization Perception (1–5 scale)	4.2 (SD = 0.8)	2.8 (SD = 1.0)

4.3. *Evaluation of the PAGE adaptive learning system*

As shown in Fig. 6, a quasi-experimental study was conducted to compare the PAGE system (experimental group) against a standard or cognitively adaptive control system (control group) in terms of emotional satisfaction and engagement. Random assignment or matching procedures ensured that the two groups did not differ significantly in demographic variables, baseline knowledge, or personality traits prior to the intervention. As shown in Table 4, over the course of several weeks, quantitative analyses indicated that the experimental group reported notably higher emotional satisfaction at the end of the study, with a mean score of approximately 4.4 (SD = 0.7), compared to 3.6 (SD = 0.8) in the control group. This difference was statistically significant ($t(198) = 7.43$, $p < 0.001$), and the effect size (Cohen's $d = 1.05$) suggested a large impact, indicating that psychologically aware content considerably enhanced learners' positive affect.

Learner engagement also differed considerably as measured by our composite engagement index (detailed in Section 3.4): the experimental group showed significantly higher overall engagement (0.78 vs. 0.63, $p < 0.01$), including higher task completion rates, longer average study times, and a greater willingness to attempt optional challenges. The engagement index calculation allowed us to quantitatively compare multidimensional engagement factors between the two groups.

Participants attributed improved engagement to the adaptive content and emotional support features. Learners high in Neuroticism had reduced dropout rates, attributed to adaptive emotional responses. For clarity, we define 'dropout' in this study as the discontinuation of a learning module before completion, specifically measured by: (1) not returning to the system within 72 h after leaving a module incomplete, and (2) completing less than 70 % of the assigned activities within that module. Using these criteria, the experimental group showed a 12 % overall dropout rate compared to 24 % in the control group ($\chi^2 = 8.76$, $p < 0.01$), with the most pronounced difference observed among participants scoring in the top quartile for Neuroticism (15 % vs. 38 %, $\chi^2 = 10.21$, $p < 0.001$). A supplementary knowledge test showed a modest advantage for the experimental group ($M = 78.4$, $SD = 8.2$ vs. $M = 74.7$, $SD = 9.1$), though this difference did not reach statistical significance ($p = 0.08$). Interestingly, we also found that the personalization effect was not uniform across all personality dimensions—the system appeared most effective for learners scoring high on Openness and Neuroticism, while differences for learners with varying levels of agreeableness were minimal and non-significant. Additionally, contrary to our expectations, content personalization based on conscientiousness showed inconsistent results, with some high conscientiousness learners preferring more exploratory rather than structured content. These unexpected findings warrant further study on a larger scale or over a longer duration to better understand the complex interplay between personality dimensions and adaptive content.

5. Discussion

This study evaluated the PAGE system—a GenAI-based adaptive learning framework integrating Big Five personality traits and emotional feedback. A DBR approach supported iterative design and testing. Results showed improvements in emotional satisfaction, learner engagement, and personalization.

5.1. *Advantages and contributions of a GenAI-Driven psychologically aware personalization*

This study demonstrated that integrating psychological variables into GenAI-based content personalization improves learner experience. PAGE exceeded traditional cognitive-only systems in emotional satisfaction and engagement. By dynamically adjusting content to personality and emotional state, PAGE provided real-time adaptation beyond static personalization.

The system generates content tailored by narrative style, complexity, and support mechanisms. When frustration was detected, PAGE simplified tasks and offered encouragement; when interest rose, it provided challenging material. This dual adaptation model allowed the system to maintain motivation while minimizing negative affect.

5.2. *Challenges and limitations*

Several limitations remained. GenAI models could produce inaccurate or misleading content, which may undermine learning. We reduced such issues with prompt engineering and post-processing, but complete elimination was not guaranteed.

Relying on self-reported data for emotions and cognition introduced bias. More accurate sensing would require costly or intrusive techniques. The sample size ($n = 200$) and demographic homogeneity also limited generalizability. Cultural, contextual, and developmental differences could influence system effectiveness.

5.3. *Design principles for psychologically aware GenAI educational systems*

Drawing on multiple iterations in our DBR programmed and synthesizing both empirical findings and broader discussions, we proposed a set of design principles that may guide future developments in similar systems, building on the adaptation strategies illustrated in Fig. 4:

- Develop adaptive learner models integrating cognitive, personality, and emotional dimensions, with mechanisms to monitor key psychological indicators dynamically to enable adaptive content generation.
- Tailor personalization for style and complexity to psychological traits (e.g., exploration for high openness, structured pathways for high conscientiousness).
- Incorporate real-time emotional support mechanisms that respond to emotional states with simplified tasks and encouraging feedback for negative emotions, or more challenging content for positive emotions
- Ensure content quality through knowledge verification and expert review at various stages of model output, using external knowledge bases or peer review for sensitive topics.
- Promote learner autonomy through transparent recommendations and user control options, allowing learners to understand why they received certain content and override or reject recommendations as needed

Additionally, integrating pedagogical theories—such as constructivism and cognitive load theory—can ensure that these adaptive approaches facilitate genuinely meaningful learning. Longitudinal monitoring and iterative improvements remain essential for preventing novelty effects, maintaining system effectiveness, and addressing unforeseen issues.

Table 5
Mapping of personality traits and emotional states to content adaptation features.

Personality	Emotion	Content Adaptation Features
High openness	Interest	• Interdisciplinary examples • Theoretical explorations • Creative application questions • Multiple perspectives
High conscientiousness	Confusion	• Numbered sequential steps • Clear learning objectives • Detailed explanations • Verification checkpoints
High neuroticism	Frustration	• Reassuring language • Simplified explanations • Smaller content chunks • Confidence-building feedback
High extraversion	Satisfaction	• Conversational style • Social learning activities • Interactive components • Group exercises

5.4. Implications for educational stakeholders

PAGE offered actionable insights for educational stakeholders. Platform developers could enhance personalization by integrating psychological models. Educators could use GenAI tools to tailor materials for diverse learners. Policymakers should define standards to safeguard equity, transparency, and data protection. With responsible use, psychologically aware GenAI systems could improve both learner experience and learning outcomes.

6. Conclusion

This study successfully developed and validated the Psychologically Aware Generative Education (PAGE) system, a novel personalized learning approach integrating GenAI with Big Five personality traits and real-time emotional feedback, employing a Design-Based Research methodology. Our findings demonstrated GenAI's capability to produce psychologically aware educational content. The PAGE system led to significantly higher emotional satisfaction (4.4/5 vs. 3.6/5) and enhanced learning engagement (e.g., 22 % higher task completion rates and 34 % increased study duration) in experimental group participants compared to traditional methods, underscoring the importance of addressing learners' psychological needs.

Theoretically, the PAGE paradigm extends adaptive learning by empirically supporting personalized instruction that addresses both cognitive and affective dimensions. Practically, this research offers a validated system prototype and a set of iteratively derived design principles, alongside actionable guidelines for the broader educational ecosystem, to guide the development and implementation of psychologically aware AIED systems.

Despite these promising results, limitations include the short-term nature of the intervention and the need for broader sample diversity to confirm cross-cultural validity. Future research should prioritize longitudinal studies to assess long-term impacts on learning outcomes and metacognitive skills, investigate diverse populations, and continue advancing technical solutions and ethical discussions surrounding data privacy, algorithmic bias, and learner autonomy. Addressing these challenges is imperative for progressing towards a more holistic, ethically sound vision of AI-driven education that effectively integrates both cognitive and psycho-emotional dimensions into the learning experience.

Qualitative feedback indicated that learners felt the system "understood" them and offered personalized care. As shown in Table 5, the system employed specific adaptation features tailored to personality-emotion combinations. This table provides side-by-side comparisons of actual content adaptations for different psychological profiles, with annotations highlighting the adaptation mechanisms. The mapping between psychological states and content features formed the foundation of the PAGE system's adaptation logic, guiding the prompt construction and content generation for each personality-emotion combination.

While positive overall, concerns included GenAI reliability and data privacy. These issues required further attention in future development.

CRediT authorship contribution statement

Wentao Hu: Writing – review & editing, Writing – original draft, Formal analysis, Data curation, Conceptualization, Funding acquisition, Methodology, Validation, Visualization. **Zichen Shao:** Project administration, Data curation, Formal analysis, Methodology, Validation, Visualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work is supported by the Provincial-Level Key Undergraduate Teaching Reform Projects (Second Batch) under the 14th Five-Year Plan (Grant No. JGZD2024080).

Data availability

Data will be made available on request.

References

Abid, A., Farooqi, M., & Zou, J. (2021). Persistent Anti-Muslim bias in large language models. *arXiv preprint arXiv:2101.05783*.
Baker, R. S., D'Mello, S. K., Rodrigo, M. M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies, 68*(4), 223–224.
Binhammad, M. H. Y., Othman, A., Abuljadayel, L., Al Mheiri, H., Alkaabi, M., & Almarri, M. (2024). Investigating how generative AI can create personalized learning materials tailored to individual student needs. *Creative Education, 15*(7), 1499–1523. https://doi.org/10.4236/ce.2024.157091
Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*(1), 49–59.
Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101.
Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems, 33*, 1877–1901.
Brusilovsky, P., & Millán, E. (2007). User models for adaptive hypermedia and adaptive educational systems. *The adaptive web*, 3–53.
Chamorro-Premuzic, T., & Furnham, A. (2003). Personality predicts academic performance: Evidence from two longitudinal university samples. *Journal of Research in Personality, 37*(4), 319–338.
Corbett, A. T., & Anderson, J. R. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge, user modeling and user-adapted interaction. *4*(4), 253–278.
Costa, P. T., & McCrae, R. R. (1992). *Revised NEO personality inventory (NEO- PI-R) and NEO five-factor inventory (NEO-FFI) professional manual*. Psychological Assessment Resources.
De Feyter, T., Caers, R., Vigna, C., & Berings, D. (2012). Unraveling the impact of the big five personality traits on academic performance: The moderating and mediating effects of self-efficacy and academic motivation. *Learning and Individual Differences, 22*(4), 439–448.
Graziano, W. G., Habashi, M. M., Sheese, B. E., & Tobin, R. M. (2007). Agreeableness, empathy, and helping: A person x situation perspective. *Journal of Personality and Social Psychology, 93*(4), 583–599.
Guettala, M., Bourekkache, S., Kazar, O., & Harous, S. (2024). Generative artificial intelligence in education: Advancing adaptive and personalized learning. *Acta Informatica Pragensia*, (3), 460–489. https://doi.org/10.18267/j.aip.235, 2024.
Immordino-Yang, M. H., & Damasio, A. (2007). We feel, therefore we learn: The relevance of affective and social neuroscience to education. *Mind, Brain, and Education, 1*(1), 3–10.
Jansz, S., Molenaar, I., Knoop-van Campen, C., Demmers, K., & van Gog, T. (2022). Personality traits, motivation, and performance in computer-based personalized

learning. *Computers & Education, 187*, Article 104446. https://doi.org/10.1016/j.compedu.2022.104446

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y., Madotto, A., & Raffel, C. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys, 55*(12), 1–38.

John, O. P., & Srivastava, S. (1999). The big five trait taxonomy: History, measurement, and theoretical perspectives. In *Handbook of personality: Theory and research* (Vol. 2, pp. 102–138). The Guilford Press.

Kasneci, E., Sessler, K., Kuchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Gunnemann, S., Hullermeier, E., et al. (2023). Chatgpt for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, Article 102274.

Kim, S. W. (2020). Meta-analysis of parental involvement and achievement in East Asian countries. *Education and Urban Society, 52*(2), 312–337.

Kim, J., Park, S., & Lee, H. (2023). Personality-aware language models: Adapting text generation for user engagement. In *Proceedings of the 11th international conference on learning analytics & knowledge* (pp. 112–121).

Komarraju, M., Karau, S. J., Schmeck, R. R., & Avdic, A. (2011). The big five personality traits, learning styles, and academic achievement. *Personality and Individual Differences, 51*(4), 472–477.

Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys, 55*(9), 1–35.

McCrae, R. R., & Costa, P. T., Jr. (1997). Personality trait structure as a human universal. *American Psychologist, 52*(5), 509–516.

McKenney, S., & Reeves, T. C. (2018). *Conducting educational design research*. Routledge.

Miao, Z., Pan, Y., & De la Torre, J. (2023). Are large language models ready for education? A case study on chatgpt for educational assessment. *arXiv preprint arXiv:2305.12404*.

Ouf, S., Abd Ellatif, M., Salama, S. E., & Helmy, Y. (2017). A proposed paradigm for smart learning environment based on semantic web. *Computers in Human Behavior, 72*, 796–818.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315–341.

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology, 36*(1), 36–48.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist, 37*(2), 91–105.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135*(2), 322–338.

Reeves, T. C. (2006). *Design research from a technology perspective*. Routledge.

Reynolds, L., & Zhang, D. (2022). Prompt engineering for educational language models: A comprehensive analysis of strategies and applications. *Journal of Educational Computing Research, 60*(4), 1182–1214.

Shemshack, A., Kinshuk, & Spector, J. M. (2021). A comprehensive analysis of personalized learning components. *Journal of Computers in Education, 8*(4), 485–503.

Shute, V. J., & Zapata-Rivera, D. (2012). Adaptive educational systems. *Adaptive Technologies for Training and Education, 7*(27), 1–35.

Tlili, A., Wang, H., Gao, B., Shi, Y., Zhiying, N., Looi, C.-K., Liu, D., & Huang, R. (2021). Impact of cultural diversity on students' learning behavioral patterns in open and online courses: A lag sequential analysis approach. *Interactive Learning Environments*, 1–20doi. https://doi.org/10.1080/10494820.2021.1946565

UNESCO. (2023). Guidance for generative ai in education and research. *Tech. rep.* United Nations Educational, Scientific and Cultural organization.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist, 46*(4), 197–221.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems, 30*.

Wang, Y., Kordi, Y., Mishra, S., Liu, A., Smith, N. A., Khashabi, D., & Ha- jishirzi, H. (2022). Self-instruct: Aligning language models with self-generated instructions. *arXiv preprint arXiv:2212.10560*.

won Kim, S. (2019). Is socioeconomic status less predictive of achievement in East Asian countries? A systematic and meta-analytic review. *International Journal of Educational Research, 97*, 29–42.

Wu, J., Zhang, Y., Chen, J., & Keller, J. M. (2023). Emotion recognition in educational contexts: A systematic review of methods, applications, and implications. *Computers & Education, 193*, Article 104645.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1), 1–27.