



Research on the Model Construction of AI-Empowered Multimodal Teaching in High School English from the TPACK Perspective

Wenxuan Guo, Ziliang Zheng*, Mingyu Zhang

School of Foreign Languages, Hubei University, Wuhan, 430062, China

*202331103011034@stu.hubu.edu.cn

Abstract. With the rapid development of educational informationization and artificial intelligence (AI) technology, the field of education is undergoing profound transformations. In particular, the cultivation of core competencies in senior high school English education is facing new challenges. Although the teaching competency development of senior high school English teachers has significantly improved and their ability to use technological tools is increasingly enhanced, current teaching practices still exhibit issues such as superficial application of technology and monotonous use of modality. Based on the TPACK theory, this study integrates AI technology and the concept of multimodal teaching to explore the organic combination of the three. It aims to construct a new teaching model termed "TPACK-AI-MM" and apply it to concrete high school English classroom instruction. The study also attempts to design a "Four-Layer Three-Stage Method" teaching procedure framework to help high school English teachers seek practical teaching solutions, provide robust support for cultivating core competencies in high school English education, and realize educational modernization.

Keywords: TPACK; AI; Multimodal Teaching; English Education; Core Competencies

1 Introduction

The rapid advancement of educational digitization and artificial intelligence is driving profound reconstruction in high school education. As a crucial component of high school education, English language education plays a vital role in cultivating students' abilities and enhancing their core competencies[1]. It also poses serious challenges to its traditional teaching models: Firstly, there is a "superficial" application of technology, where teachers' understanding of information technology often remains at a theoretical level and fails to be deeply integrated into all aspects of English teaching practice[2]. Secondly, the use of modalities tends to be "monolithic", with teachers overly relying on linguistic functions in teaching and neglecting the use of multiple senses and semiotic resources to support student learning[3]. Thirdly, there is an "in-

sufficient integration" of teaching elements. Influenced by various internal and external factors, teachers often lack the ability to deeply integrate AI technology with multimodal teaching, resulting in a tendency toward simply superimposing different elements in instruction [4].

The TPACK theory provides a classical perspective for understanding and addressing the aforementioned issues. However, in the era of artificial intelligence, the traditional TPACK framework exhibits certain limitations in terms of technology integration, contextual adaptation, and restoration of Educational Humanism[5]. Meanwhile, multimodal teaching approach in practice often falls into the pitfalls such as "excessive technological intervention" and "monolithic or cluttered modality use," lacking systematic and dynamic guidance[6]. Therefore, this study advocates, with the perspective of TPACK, to construct a new teaching modality driven by AI technology as the endogenous force and grounded in multimodal teaching as the integrative substrate. This modality responds to the inherent demands of TPACK theory development and provides a reality-facing and practical pathway for teachers.

2 Literature Review and Theoretical Foundation

The basic definition of the TPACK theory is a comprehensive framework of teaching competency formed when teachers effectively integrate Technological Knowledge (TK), Content Knowledge (CK), and Pedagogical Knowledge (PK) in instructional design and practice. This framework also provides a new perspective for enhancing the informationization teaching competency of high school English teachers[7]. However, as the educational ecology shifts toward "Human-Machine Collaboration", AI tools are increasingly becoming "partners" to humans, completely transcending the traditional positioning of technology as merely an auxiliary tool[5]. Therefore, achieving the integrated optimization of technology and leveraging AI to reshape components of TPACK, especially TK, TPK and TCK, has become an important practical pathway.

Multimodal Teaching is a method that emphasizes promoting student learning through multi-sensory experiences such as visual, auditory, and tactile stimuli. It particularly stresses the use of multiple semiotic resources, communication media technologies, and diversified expression of information[3]. With the widespread application of AI in education, utilizing diverse AI technology tools to implement multimodal teaching has become a trend[8]. Nonetheless, in the field of high school English teaching, the application of AI technology for multimodal teaching often falls into the misconception of "multimodality for multimodality's sake". Issues such as the simplistic piling up of technologies, non-refined modal selection, and a lack of core instructional objectives all impose unnecessary cognitive load on students' learning.

In summary, against the background of an intelligent and informatized era, although significant achievements have been made in the respective developmental fields of TPACK theory, the educational application of AI, and multimodal teaching, there remains a conspicuous gap in current research regarding how to systematically integrate these three elements based on classroom teaching practice to construct a novel, structured and operable teaching model. Based on this identified gap, this study

aims to construct an AI-empowered, novel multimodal teaching model for high school English from the TPACK perspective, thereby contributing to filling the research void while providing a new solution for the educational and teaching reform in high school English classrooms.

3 Research Design and Methods

This study is a theoretical construction type research. Its core objective is to construct a novel "TPACK-AI-MM" instructional model through systematic theoretical deduction and design. And three main research methods are employed.

First, the "Literature Analysis Method", is used to systematically review and analyze relevant literature in the fields of "TPACK theory", "AI technology", and "multimodal English teaching". The focus is on clarifying the developmental trajectories of major theories to establish a solid theoretical foundation for this study, while simultaneously grasping academic frontiers and precisely identifying research gaps.

Second, the "Theoretical Deduction Method" is applied. Using the TPACK theory as the framework, this method organically integrates AI technology and multimodal teaching to systematically and logically construct the "TPACK-AI-MM" theoretical model. It elucidates the interrelationships and connecting mechanisms among the various internal elements.

Finally, the "Case Design Method" is utilized. Taking specific high school English teaching scenarios as the entry point, an instructional process model termed the "Four-Layer Three-Stage" method is designed. This serves to make it more intuitive and tangible within concrete teaching processes and application scenarios, thereby enhancing the comprehensibility, operability, and practical value of the "TPACK-AI-MM" model.

4 Construction of the TPACK-AI-MM Theoretical Model

This study attempts to construct an integrated "TPACK-AI-MM" model: using the discipline-innovated TPACK theory as the adaptive core, AI technological tools as the permeable empowering medium, and multimodal teaching as the concretized, operable teaching processes, so that the three form an organically nested and dynamically adaptive whole.

The core of this modal lies is the contextualized reshaping of the classical TPACK framework for the context of "high-school English teaching". In the high-school English context, content knowledge (CK) must be specified as the cross-cultural thinking underlying English discourse, grammar, etc.; pedagogical knowledge (PK) also needs to be anchored to language acquisition pathways such as task-driven learning and interactive communicative situations. Technological knowledge (TK), in the era of artificial intelligence, has already expanded into a new dimension of AI-TK, which is crucial in assisting teaching[9][10]. Particularly key is that these three do not exist in isolation but should be a dynamic, interactive, and continuously generative process. It is precisely this dynamic interaction that forms pedagogical content knowledge (PCK),

technological content knowledge (TCK), technological pedagogical knowledge (TPK), and ultimately integrates and optimizes into the TPACK theory oriented toward specific teaching contexts[7].

In the era of artificial intelligence, the essential impact of generative AI lies in its reshaping of TK and TPK. In high school English teaching practice, teachers' understanding and application of TK often appear static and one-directional. For example, traditional teachers may only use multimedia interactive whiteboards to assist teaching. With the development of the AI era, the use of teaching technological resources is shifting toward a more generative, interactive, and adaptive "intelligent era." This also catalyzes innovation in TPK. For instance, teaching evaluation can leverage AI to achieve a leap from "result judgment" to "process insight," and instructional design can shift from "experience-based estimation" in traditional models to data-informed and intelligent "precise adaptation."

To avoid falling into the misconception of "multimodality for multimodality's sake" in multimodal teaching practice, this study distills three key principles: "Purpose-Oriented, Clear Prioritization, and Functional Complementarity." In any instructional design, English language teaching and its core competencies are the ultimate goal. Visual, auditory and other multimodal elements must be employed systematically while serving instruction objectives, fulfilling auxiliary and reinforcing functions, and ensuring that all modalities work synergistically toward achieving teaching goals. Artificial intelligence and various technological tools need to be embedded in the complete continuum of "teaching" and "learning," forming a closed-loop system that supports teaching and learning, realizing full-chain participation from diagnosing learning situations and generating contexts to providing interaction and feedback evaluation[3][6].

In summary, the "TPACK-AI-MM" model attempts to depict a new type of teaching modality. The dynamic triangle of TPACK theory serves as its "most powerful brain," primarily responsible for teaching decision-making. And artificial intelligence acts as the "neural network" permeating throughout, playing the roles of perception, processing, and empowerment in specific teaching situations. Multimodal teaching functions as the "body and trunk" guiding its actions, ensuring coordinated activities. These three ultimately serve the concrete classroom practice of English teaching, constituting an interdependent, mutually reinforcing, and highly systematic organic whole.

5 The Model Framework for AI-Empowered Multimodal Teaching in High School English

Based on the aforementioned theoretical model of "TPACK-AI-MM" and the specific realities of high school English teaching, this study attempts to design the "Four-Layer Three-Stage Method" as a practical framework for AI technology-empowered multimodal English teaching from the TPACK perspective. This study elaborates the practical pathway of the "TPACK-AI-MM" modality in concrete teaching contexts from four major layers—the Objective Layer, Process Layer, Role Layer, and Evaluation Layer—and three major stages: Pre-class, In-class, and Post-class.

The "Objective Layer" must be precisely aligned with the core competencies of high school English teaching, which is the ultimate goal of all educational and teaching activities. Multimodal teaching should be actively employed to cultivate and apply English language skills, thereby enhancing students' language proficiency[1]. Utilizing the multicultural materials provided by AI to cultivate students' cultural awareness is also important. During the teaching process, teachers should learn to use AI to assist in designing inquiry-based tasks, helping students hone their thinking quality in the process of problem-solving. Meanwhile, personalized learning path planning created by AI and human-AI collaborative exploration should be leveraged to develop students' learning abilities[5].

The "Process Layer", as the core operational chain of this framework, runs systematically through the entire cycle of Pre-class, In-class, and Post-class stages, aiming to achieve precise presupposition, immersive interaction, and expansive transfer. First, in the "Pre-class stage", teachers should employ AI to conduct multimodal analysis of students' learning status, interpret students' knowledge weaknesses and learning style preferences, etc. Then assist students in generating personalized learning resource packages, such as micro-lectures, pictures and other materials to help students prepare for pre-class preview. Next, in the "In-class stage", teachers should focus on creating authentic contexts to stimulate students' cognition, while conducting multimodal collaborative inquiry and encouraging group cooperation among students, thereby emphasizing students' agency in the application of AI technology. Finally, in the "Post-class stage", teachers should utilize AI technology to assist in customizing differentiated homework tasks and consolidation exercises, monitor and analyze students' learning situations, provide personalized diagnosis, feedback, and learning path suggestions for students, thereby achieving a closed-loop teaching process[10].

The "Role Layer" requires both teachers and students to jointly construct a new type of teacher-student relationship characterized by "Human-Machine Collaboration". The teacher's role should shift from the traditional knowledge transmitter to the helmsman of "AI-assisted multimodal teaching", designing teaching contexts and guiding the cultivation of core competencies. Meanwhile, the student's role should also transform from passive knowledge recipients into explorers in "Human-Machine Collaboration" and participants in "multimodal learning." This approach not only emphasizes the student-centered principle in education but also fosters the construction of a new teaching community[5][9].

The "Evaluation Layer" focuses on leveraging the comprehensive function of an embedded multimodal AI educational assessment system[9]. Teachers should establish a comprehensive evaluation system that combines process and outcome assessments. Using AI to automatically record and analyze data on students' participation, interaction frequency, resource usage, etc., thereby completing students' process evaluation. Simultaneously, relying on multi-dimensional evaluation rubrics and modern evaluation technology tools, teachers could assess students' multimodal learning outcomes to complete high-quality outcome evaluation, achieving comprehensiveness and developmental nature in evaluation [3].

6 Support System and Pathways for Model Implementation

Achieving the systematic enhancement of high school English teachers' AI-TPACK literacy is a crucial key to the practicalization of this modality. The focus lies in constructing a multi-level, multi-dimensional, and sustainable support system. This requires us to start from the cultivation of teachers' internal capabilities, promoting their self-breakthrough in the field of intelligent education. And it also asks us to optimize the external environment, creating favorable conditions for the enhancement of teachers' AI-TPACK literacy[4]. High school English teachers should follow the developmental path of "Cognition-Acceptance-Design-Innovation" to systematically improve their ability to integrate AI technology with TPACK. Meanwhile, schools should also provide corresponding cultivation strategies, such as offering specific training courses for teachers, organizing AI-TPACK themed workshops and hosting related seminars, etc.[9][10].

Regarding technological integration, this study advocates a "Lightweight and Ecological" practical strategy. It is recommended to rely on the National Smart Education Platform for primary and secondary schools and prudently introduce lightweight, highly operable, secure and controllable domestic AI tools such as Wenxin Yiyan and KIMI. This approach promotes the natural integration of tools into existing teaching processes, rather than indiscriminately pursuing the application of emerging technologies, balancing the multiple relationships among educational ethics, teachers' usage habits, and tool technical efficacy. For instance, in an actual instructional context, taking a high school English themed writing session on "Environmental Protection" as an example, the teacher can guide students in using KIMI to efficiently extract key information from relevant English reports as material support. Subsequently, students may utilize "Wenxin Yiyan" to generate a preliminary draft framework for their writing, upon which they can make personalized modifications and linguistic refinements. Finally, the teacher can monitor progress in real time via the platform and provide targeted feedback. This process not only integrates naturally into existing teaching workflows but also demonstrates the lightweight yet effective support offered by AI tools in stimulating learning interest and enhancing information processing efficiency.

Collaborative support from institutional and cultural environments is also indispensable. It is necessary to urge schools to provide systematic support in areas such as flexible curriculum scheduling, teaching evaluation and incentive systems, hardware facility updates, and network environment optimization. This will foster a cultural atmosphere that encourages teaching innovation and tolerates exploration and trial-and-error. For example, establishing "Teaching Innovation Experimental Courses" can offer flexible guarantees and trial space for teachers to innovate in teaching models.

Resource development for high school English education should shift toward a "Human-Machine Collaboration" model. Constructing a school-based AI multimodal teaching resource repository is an important pathway, encouraging the establishment of a resource development model characterized by "AI preliminary generation + teacher collective review + in-depth refined processing." By utilizing the efficiency advantages of AI large-scale models to generate foundational teaching resources, subject teacher

teams can then conduct professional review, in-depth optimization, and systematic archiving. This approach aims to create a high-quality, localized school-based resource repository, ensuring the precision and applicability of teaching resources.

7 Potential Challenges and Systemic Countermeasures for the Model

In the process of integrating the "TPACK-AI-MM" modality into high school English teaching, teachers face multiple realistic challenges and urgently require systematic and prudent coping strategies.

First, facing cognitive and competency challenges, teachers need to recognize the nature of technological anxiety and skill deficiency issues. Teachers can rely on school-enterprise collaboration mechanisms, integrating two-tier resources of disciplinary knowledge and enterprise AI technology tools to create "out-of-the-box" intelligent instructional design templates, thereby reducing the difficulty of using AI tools.

Secondly, there are challenges related to teaching effectiveness and curriculum adaptability. Inappropriate use of multimodal teaching and AI tools may cause cognitive overload in students. This requires strengthening the principle of "Purpose-Driven Modality Design", clarifying the teaching functions and optimal application timing of each modality, and establishing an efficient and meaningful multimodal learning environment[1].

Furthermore, AI-generated content may contain biases or deviate from the high school English curriculum standards. Teachers must consistently adhere to the principles of "Dominant Multi-layered Review" and "Creative in-depth processing" to ensure the direction and scientific rigor of teaching.

Finally, ethical and security challenges cannot be ignored, especially issues concerning data privacy, algorithmic bias, and information ethics. In the teaching process, teachers should prioritize educational technology products that are registered with the Ministry of Education and comply with data security regulations. Additionally, they should appropriately integrate AI ethics education into teaching to foster students' critical awareness of technology and sense of responsibility[5].

8 Conclusion

Based on the practical needs for the transformation of high school English teaching in the era of educational intelligence and informationization, this study attempts to construct the "TPACK-AI-MM" model and its "Four-Layer Three-Stage Method" instructional process framework. It aims to provide a path for the transformation that is both theoretical support and practical feasibility. This model not only promotes the development and innovation of TPACK theory in the era of artificial intelligence but also organically combines it with localized multimodal English subject teaching, realizing a temporal expansion of localization and discipline-specific multimodal English teaching. Furthermore, from a practical perspective, it offers teachers a systematic

reference framework covering conceptual renewal, process reconfiguration, and instructional model reform.

However, this study is currently still in the theoretical construction stage, and its actual effectiveness and applicable boundaries still require verification and refinement through large-scale empirical research. Future research can be conducted based on this foundation, particularly focusing on its specific effects on cultivating students' language ability and core competencies, and continuously promoting the deep integration of artificial intelligence with multimodal teaching practice in English education from the TPACK perspective, achieving genuine integrated innovation.

References

1. Ji, R. R. (2024). Implementation strategies for multimodal teaching in senior high school English. *New Curriculum Teaching*, (23), 52–54.
2. He, T. (2025). Research on the improvement paths of digital teaching capabilities of college English teachers under the TPACK framework. *English Square*, (26), 120–123.
3. Huang, G. L. (2025). Exploration of the application strategies of multimodal teaching method in senior high school English teaching. *Teacher*, (16), 77–79.
4. Liu, Q. H. (2024). Research on the improvement paths of intelligent literacy for teachers in private undergraduate colleges under the AI-TPACK theoretical framework integrating artificial intelligence technology. In *Proceedings of the 2024 Cultural Information Development Forum* (Volume II).
5. Wu, H. Q., & Wu, C. Z. (2025). Evolution and innovation of teachers' knowledge frameworks in the GenAI era: From TPACK to GHEX-IPACK. *E-Education Research*, (10), 23–31.
6. Ma, Y. Q., & Zhao, L. Z. (2023). Integration of the TPACK framework and multimodal English teaching models in the new media era. *Overseas English*, (13), 78–80.
7. Zhang, Q. (2025). Paths for improving the information-based teaching capabilities of college English major teachers from the perspective of TPACK theory. *Education and Teaching*, (12), 148–151.
8. Deng, X. M. (2025). Research on the current situation of multimodal teaching in senior high school English classes supported by modern educational technology. [Master's thesis, Guizhou Normal University].
9. Wang, P. J., Wang, Y. E., & Wang, S. (2023). Research on the artificial intelligence literacy framework for normal school students based on the AI-TPACK theory. *China Information Technology Education*, (20), 67–71.
10. Zhao, Y. (2023). Research on the development of TPACK for college foreign language teachers in the context of the artificial intelligence era. *China Journal of Multimedia & Network Teaching*, (S1), 61–64.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

