

Digital Textbook Presentation Mode Based on Cognitive Theory of Multimedia Learning

Tang Xin^{1,a}

¹Zhiyuan School of Liberal Arts, Beijing Institute of Petrochemical Technology, Beijing, China

^atangxin@bipt.edu.cn

Keywords: Cognitive Theory of Multimedia Learning; digital textbooks; presentation mode; cognitive load; instructional design

Abstract: Against the backdrop of the digital transformation of education, digital textbooks have replaced traditional paper textbooks and become the core teaching carrier of smart education. At present, some digital textbooks are plagued by such problems as the indiscriminate stacking of multimedia elements, disorderly information presentation, excessive cognitive load, and inefficient interactive design, failing to give full play to the educational advantages of multimedia teaching. Based on Richard E. Mayer's Cognitive Theory of Multimedia Learning, this paper sorts out the cognitive mechanism and core design principles of multimedia learning. From four dimensions including cognitive load regulation, multimodal information adaptation, interactive scenario optimization and hierarchical presentation design, it constructs scientific and feasible design strategies for the presentation modes of digital textbooks. The aim is to reduce learners' invalid cognitive consumption, optimize the path of knowledge reception and internalization, improve the teaching adaptability and learning effectiveness of digital textbooks, and provide theoretical reference and practical ideas for the optimal development and teaching application of digital textbooks.

1. Introduction

With the in-depth integration of artificial intelligence, big data and multimedia technology with education and teaching, the digital reform of education in China has been continuously deepened. As the core carrier of curriculum teaching, textbooks have officially entered a new stage of digital transformation and upgrading. Compared with the single presentation form of texts and static images of traditional paper textbooks, digital textbooks integrate diversified resources such as texts, audios, videos, animations, virtual simulations and interactive exercises. They possess advantages including dynamic content update, diverse presentation forms, multiple interactive scenarios, and unrestricted learning in time and space. Capable of adapting to various teaching modes such as autonomous learning, cooperative learning and inquiry-based learning, digital textbooks have become an important tool for empowering classroom teaching reform and implementing personalized education.

However, in the actual development and application process, most digital textbooks are confronted with the problem of "emphasizing form over connotation".[1] Developers excessively

pursue the richness of multimedia elements and blindly superimpose resources including images, texts, audios, videos and animations, while ignoring learners' cognitive laws and the limits of information processing. Furthermore, some digital textbooks suffer from chaotic graphic and text layout, redundant audio and video resources, ambiguous information hierarchy, and cumbersome interactive design.[2] These problems lead to learners' distracted attention and excessive cognitive load, making it difficult for them to accurately grasp core knowledge.

The Cognitive Theory of Multimedia Learning focuses on learners' cognitive processing procedures, reveals the internal laws of human beings in receiving, screening, organizing and integrating multimedia information, and clarifies the core design criteria for multimedia teaching. Based on this theory, this paper explores the scientific presentation and design path for digital textbooks, solves the cognitive adaptation problems existing in current textbook presentation, and facilitates the transformation of digital textbooks from "formal digitization" to "precision education".

2. Theoretical Basis

2.1 Core Connotation of the Cognitive Theory of Multimedia Learning

Proposed systematically by American educational psychologist Richard E. Mayer, the Cognitive Theory of Multimedia Learning constructs a complete cognitive model of multimedia learning based on the Dual Coding Theory, Cognitive Load Theory, and Constructivist Learning Theory. Its core viewpoint holds that learners process multimedia information through two channels: the visual channel and the auditory channel, and the cognitive capacity of both channels is limited. Essentially, human learning is an active process of knowledge construction. Rather than passively accepting information, learners establish connections between new and prior knowledge by screening, organizing and integrating multimedia information, and ultimately realize knowledge internalization.

Mayer divides the cognitive process of multimedia learning into three core stages. The first is the information selection stage, in which learners screen visual and auditory information relevant to learning goals from massive multimedia information and eliminate invalid interfering information. The second is the information organization stage, where learners sort out and integrate the fragmented information selected, so as to construct clear visual and auditory knowledge frameworks. The third is the information integration stage, in which learners associate the newly constructed knowledge structure with their existing knowledge reserve to complete the absorption and construction of new knowledge. These three stages are closely interconnected; unreasonable information presentation at any stage will hinder the progress of cognitive processing.[3]

2.2 Core Design Principles of Multimedia Learning

Based on the laws of cognitive processing, Mayer put forward a series of multimedia design principles adapted to teaching scenarios, which serve as the core basis for the presentation design of digital textbooks. The core principles are as follows:

First, the proximity principle. Relevant information such as graphics, texts, audios and videos corresponding to core knowledge points should be presented in spatial proximity and temporal synchronization. This avoids repeated line-of-sight switching and information recall of learners, thereby reducing unnecessary cognitive consumption. Second, the coherence principle. Redundant information irrelevant to knowledge points, including decorative images, background music and animation special effects, should be eliminated from textbook presentation to prevent irrelevant content from occupying limited cognitive resources. Third, the modality principle. For abstract

knowledge points, long paragraphs of static text should be preferably replaced with “audio explanation + dynamic images”. The auditory channel is adopted to assist the visual channel so as to improve the efficiency of information processing. Fourth, the redundancy principle. The repeated presentation of a single knowledge point through text, audio, video and other multiple forms should be avoided, so as to prevent cognitive overload caused by excessive redundant information. Fifth, the segmenting principle. Complex knowledge should be divided into independent modules and presented in segments to adapt to learners’ limited cognitive capacity and avoid excessive information input at a single time.

2.3 Adaptation Logic between Digital Textbook Presentation and Cognitive Theory

The core advantage of digital textbooks lies in multimodal information presentation. However, their essence is to serve learners’ knowledge construction rather than merely display technical effects. The core objective of the Cognitive Theory of Multimedia Learning is to optimize information presentation forms, adapt to human cognitive limits, reduce invalid cognitive load, and improve effective learning efficiency, which is highly consistent with the design goals of digital textbooks, namely “lightweight, precision and high efficiency”.

All presentation elements of digital textbooks, including texts, images, audios, videos and interactive functions, enter learners’ cognitive system through visual and auditory channels. Only by following the dual-channel processing rules, controlling cognitive load, and standardizing the logic of information presentation, can multimedia resources facilitate rather than hinder knowledge construction. Therefore, integrating the Cognitive Theory of Multimedia Learning into the presentation design of digital textbooks is an inevitable approach to solving the current confusion in textbook design and realizing the in-depth integration of technology and teaching.

3. Existing Problems in the Presentation Modes of Current Digital Textbooks

3.1 Excessive Stacking of Multimedia Elements and Severe Cognitive Overload

The development of most current digital textbooks falls into the misconception of “technology supremacy”. Developers one-sidedly believe that the more multimedia elements are embedded, the better the teaching effect of textbooks will be. According to the Cognitive Theory of Multimedia Learning, human cognitive capacity of dual channels is limited. Excessive invalid information consumes learners’ cognitive resources, making it impossible for them to concentrate on processing core knowledge and resulting in cognitive fatigue and distracted attention. Meanwhile, some textbooks repeatedly elaborate on the same knowledge points through texts, audios and videos, which violates the redundancy principle, further intensifies cognitive load, and greatly reduces learners’ learning efficiency.

3.2 Misaligned Information Presentation Violating Cognitive Processing Laws

Misaligned information presentation is a prevalent problem of digital textbooks, which is mainly reflected in spatial and temporal dimensions. In terms of space, most textbooks suffer from separated graphics and texts as well as isolated knowledge points and explanatory videos. Core textual knowledge points and auxiliary illustrations and analytical animations are scattered on different pages or areas. In terms of time, the audio and video explanations of some textbooks are not synchronized with textual content, and the progress of animation demonstrations is disjointed from the rhythm of knowledge explanation. Learners cannot match audio-visual information synchronously or quickly construct complete knowledge logic, which hinders the process of

information integration. In addition, some textbooks present dense and unsegmented long texts without clear hierarchy, and display all complex knowledge at one time. This exceeds learners' instantaneous cognitive capacity and fails to conform to the segmenting cognitive law.

3.3 Chaotic Multimodal Collocation and Unrealized Value of Multimodal Presentation

Multimodal presentation is the core advantage of digital textbooks, yet the current collocation of presentation modes is highly arbitrary and lacks scientific adaptability. Some abstract theories and logically complex knowledge points are still presented through traditional long static texts, without adopting visual resources such as animations, simulation videos and dynamic schematic diagrams to decompose learning difficulties. In contrast, some simple and basic common-sense knowledge points are unnecessarily equipped with high-definition videos, 3D animations and other complex resources, leading to resource waste and cognitive redundancy. At the same time, most textbooks have unbalanced audio-visual collocation. Such designs fail to achieve the optimal effect of dual-channel collaborative processing, making it impossible to give full play to the educational value of multimodal resources.

3.4 Inefficient Interactive Design with Insufficient Adaptability and Personalization

The interactive design of digital textbooks aims to stimulate learners' initiative and assist knowledge internalization. However, most current textbook interactions are merely formalistic. In addition, the presentation modes are highly homogeneous, with unified display modes, resource quantities and presentation rhythms for all learners. There is no hierarchical design adapted to the characteristics of learners at different learning stages and with different cognitive levels. For lower-grade learners with weak cognitive ability, the textbooks contain excessive complex multimedia information; for upper-grade learners with strong cognitive ability, the presentation forms are overly simple and cannot meet the needs of in-depth learning. Consequently, the personalized adaptation capability of digital textbooks is severely insufficient.

4. Design Strategies for Digital Textbook Presentation Based on the Cognitive Theory of Multimedia Learning

4.1 Adhering to the Coherence Principle to Streamline Resources and Strictly Control Cognitive Load

Cognitive load regulation serves as the core prerequisite for the presentation design of digital textbooks, with the fundamental goal of eliminating invalid cognitive consumption and maximizing the retention of effective cognitive resources for knowledge construction. First, it is essential to strictly follow the coherence principle and comprehensively streamline redundant multimedia elements. Decorative special effects, irrelevant images, background music, redundant pop-up windows and other interfering information unrelated to knowledge points should be removed. All presented resources must serve the explanation of core knowledge points, eliminating the design misconception of "applying multimedia for the sake of multimedia". Second, information redundancy should be avoided. A single core modality supplemented by auxiliary modalities shall be adopted for the presentation of each knowledge point, so as to prevent repetitive elaboration through texts, audios and videos. For example, core concepts can be displayed with concise fixed texts and interpreted accurately through short audio explanations, without superimposing full-length synchronous videos.

Meanwhile, the segmenting principle should be fully implemented to modularize complex,

abstract and logically intensive knowledge. Long paragraphs are divided into short sentences of two to three lines with subheadings for hierarchical sorting. Complex principles, experimental procedures and knowledge systems are split into independent subsections and presented step by step. The next module will be unlocked only after learners complete the learning of the current one. This design adapts to learners' limited instantaneous cognitive capacity and effectively prevents information overload.

4.2 Following the Proximity and Modality Principles to Optimize Multimodal Information Adaptation Design

Based on the dual-channel cognitive processing rules of multimedia learning, the collocation and presentation positions of multimodal resources are optimized to realize audio-visual coordination and spatiotemporal adaptation. In terms of spatial proximity design, the principle of “integrated graphics and texts, synchronized audio and vision” is adopted. Relevant resources including knowledge texts, explanatory illustrations, principle animations and data charts are centrally presented in the same visual area, enabling learners to integrate information without frequent interface switching. For instance, when explaining scientific experimental principles, experimental procedure texts, dynamic experimental diagrams and key data annotations are displayed centrally to achieve spatial unification of textual interpretation and visual demonstration.

In terms of modality adaptation design, presentation forms are precisely matched according to the attributes of knowledge points to realize adaptive teaching based on knowledge characteristics. Static core knowledge such as basic concepts, definitions and formulas is mainly presented through concise texts and static charts to ensure rigor and standardization. For difficult knowledge including abstract principles, dynamic processes and spatial structures, the combined modality of “concise text+dynamic animation+voice explanation” is applied. Dynamic visual demonstrations decompose complex logic, while auditory voice interpretation assists understanding, reducing learning difficulty through dual-channel collaboration. For operational and inquiry-based knowledge, high-definition operational videos and virtual simulation interactive images are adopted to visualize abstract knowledge and dynamize static content. In accordance with the modality principle, voice explanation is prioritized to replace lengthy on-screen texts, so as to alleviate visual cognitive fatigue and improve information processing efficiency.

4.3 Optimizing Information Hierarchy to Construct Ordered Presentation Logic

Aiming at the problems of disordered content and ambiguous hierarchy in existing digital textbooks, a three-level information hierarchy system consisting of “core knowledge, auxiliary knowledge and expanded knowledge” is constructed based on the ordered rules of cognitive processing, so as to realize structured and gradient information presentation. The first level refers to core essential knowledge, which is highlighted through bold fonts, background highlighting and fixed display as the visual focus of the page, allowing learners to quickly grasp key content. The second level is auxiliary interpretative knowledge, including knowledge annotations, case analyses and error-prone prompts, which is presented in lightweight forms such as small fonts, side pop-up windows and floating prompts to avoid occupying core visual resources and support on-demand viewing. The third level is extended knowledge, including cutting-edge industry cases, interesting expansions and after-class inquiry tasks, which is presented in a hidden and click-to-unlock mode for autonomous learning of advanced students, balancing the needs of basic learning and in-depth learning.

Furthermore, the overall presentation logic of textbooks is unified in accordance with the

cognitive learning process of “introduction–intensive explanation–consolidation–expansion”. Each knowledge page strictly follows the sequence of situational introduction, core explanation, case support, exercise consolidation and knowledge expansion. This conforms to learners’ cognitive law of progressing from shallow to deep and from perception to internalization, helping learners construct a systematic knowledge framework.

4.4 Innovating Personalized Interactive Design to Adapt to Differentiated Cognitive Needs

Based on the cognitive differences among individual learners, the homogeneous presentation mode is broken to build a hierarchical, interactive and personalized textbook presentation system. First, a hierarchical presentation mode is designed. For lower-grade learners with weak cognitive ability, multimedia elements are simplified, with texts, images, animations and voice explanations as the main forms to reduce cognitive difficulty. For upper-grade learners with strong cognitive ability, logical maps, in-depth cases and simulation inquiry resources are added to meet the demands of deep learning. Meanwhile, two switchable modes, namely the basic mode and the advanced mode, are set for learners to independently adjust the presentation complexity according to their own learning level.

Second, precise interactive design is optimized. Invalid repetitive clicking interactions are eliminated, and immersive and inquiry-based interactions are designed targeting key and difficult knowledge points. For formula derivation, experimental operation and logical reasoning, interactive functions such as drag-and-drop operation, step-by-step deduction, virtual simulation experiments and error traceability are developed. This enables learners to actively process knowledge and deepen cognitive understanding through operational interaction. In addition, personalized functions including learning progress recording, key content marking, note synchronization and difficult knowledge collection are added to adapt to the cognitive habits of autonomous learning and help learners accurately break through knowledge weaknesses.

5. Key Practical Application Points for Digital Textbook Presentation Design

5.1 Prioritizing Teaching Nature to Balance the Relationship between Technology and Cognition

The core value of digital textbooks is to serve teaching and facilitate learning, and all presentation designs must adhere to the principle of “teaching priority and technology auxiliary”. In resource development and page design, technical showmanship should be avoided. The design of all multimedia elements, presentation forms and interactive functions shall take adapting to learners’ cognitive laws, achieving teaching objectives and breaking through key and difficult teaching points as the core criteria. Technology serves as a tool to optimize cognitive processing rather than the core of design. It is essential to constantly balance technical presentation and cognitive load, ensuring that every design can reduce learning difficulty and improve learning efficiency.

5.2 Balancing Universality and Adaptability to Suit Multi-scenario Teaching

Digital textbooks need to adapt to diverse teaching scenarios, including in-class collective teaching, after-class independent learning, and online remote learning. In the in-class teaching scenario, the presentation design should highlight core knowledge, simplify secondary content, adapt to the rhythm of collective teaching, and facilitate teachers’ efficient explanation and students’ accurate focus on key points. In the independent learning scenario, functions such as voice explanation, step-by-step analysis, interactive Q&A and expanded resources should be improved to

meet learners' needs for independent inquiry and knowledge supplementation. In the remote learning scenario, resource loading speed should be optimized and the interface layout simplified to avoid stuttering caused by complex special effects and ensure the continuity of cognitive processing.

5.3 Establishing a Dynamic Optimization Mechanism for Continuous Design Iteration

The presentation design of digital textbooks is not static and requires continuous optimization based on practical teaching feedback. Relying on big data technology, textbook development teams can collect and analyze learning data such as learners' learning duration, page staying time, interaction frequency and error distribution, so as to accurately locate knowledge points with cognitive difficulties and deficiencies in presentation design. For modules with excessive cognitive load, insufficient presentation adaptability and poor interactive experience, iterative optimization should be carried out dynamically in accordance with the core principles of multimedia learning cognition, realizing the dynamic adaptation between textbook presentation design and learners' cognitive needs.

6. Conclusion and Prospect

As the core carrier of educational digital transformation, digital textbooks have a direct impact on teaching quality and learning effect through the scientificity of their presentation modes. The Cognitive Theory of Multimedia Learning reveals the internal laws of human multimedia information processing, providing a solid theoretical foundation for the presentation design of digital textbooks.[4] Current digital textbooks generally suffer from problems such as excessive stacking of multimedia elements, severe cognitive overload, misaligned information presentation, chaotic multimodal collocation, and inefficient interactive design.

Based on the core principles of the Cognitive Theory of Multimedia Learning, this paper constructs a set of scientific and implementable presentation design strategies for digital textbooks from four dimensions: strict control of cognitive load, optimization of multimodal adaptation, clarification of information hierarchy, and innovation of personalized interaction. By streamlining redundant resources, conforming to cognitive laws, realizing hierarchical and orderly presentation, and empowering precise interaction, the strategies effectively solve the prominent problem of "rich forms but insufficient practical effects" of digital textbooks, assist learners in efficiently completing knowledge selection, organization and integration, and ultimately promote in-depth learning.

In the future, with the continuous development of digital and intelligent technologies, digital textbooks will be further upgraded toward intelligence, personalization and immersion.[5] Subsequent research can further integrate artificial intelligence, virtual reality and big data technologies to explore the presentation design path of intelligently adaptive and immersive digital textbooks. This can precisely match the cognitive characteristics and learning habits of different learners, truly realize the personalized "one textbook for one student" presentation mode, enable digital textbooks to better adapt to the development needs of smart education, and continuously empower the high-quality development of education and teaching.

References

- [1] Yin Qiaoling, Xia Yonggeng. *Connotation, Value and Construction Path of Interactive Digital Textbooks*[J]. *Primary and Secondary School Teaching & Textbook Research*, 2024(05), 12-16.
- [2] Li Feng, Sheng Jie, Huang Wei. *Breakthrough Points of Educational Digital Transformation: The Design and Implementation of Intelligent Textbooks*[J]. *Journal of East China Normal University (Educational Sciences)*, 2023(03), 89-98.

- [3] Richard E. Mayer. *Multimedia Learning*[M]. Translated by Sheng Qunli. Beijing: The Commercial Press, 2018.
- [4] Yuan Xiaoqun, Su Hua, Hu Lixin. *Research on Innovative Methods of Resource Organization, Content Generation and Presentation of Textbook Publishing in the Digital Intelligence Environment*[J]. *View on Publishing*, 2025(22), 45-49.
- [5] Zeng Tianshan, Wang Yanfang. *Empowering Basic Education Curriculum and Teaching Reform with Textbook Digitization*[J]. *Journal of The Chinese Society of Education*, 2024(01), 23-28.