Jiang, B. et al. (Eds.) (2025). Proceedings of the 33rd International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Extending the Textbook: Co-designing AI-Augmented Interfaces for Mixed Reality Learning in Indian Classrooms

Ganesh BENIWAL[a], Sunny Prakash PRAJAPATI[a], Syaamantak DAS[a*] & Rwitajit MAJUMDAR[b]
[a]*Centre for Educational Technology, Indian Institute of Technology Bombay, India*
[b]*Graduate School of Instructional Systems, Kumamoto University, Japan*
*syaamantak.das@iitb.ac.in

Abstract: Printed textbooks continue to serve as foundational learning resources in many classrooms across the Global South, yet they offer limited support for exploratory, multimodal, or adaptive learning. This paper presents the co-design and prototyping of the INTERACT Suite, a system of AI-augmented interfaces that extends the traditional textbook into a seamless, phygital learning environment. Through a combination of proposed mobile and Mixed Reality-based prototypes named *ext-book* and *interactouch;* the system enables learners to interact with textbook content through OCR-based AI explanations, gesture-driven media overlays, and structured annotation workflows. We describe a cooperative design study involving one subject matter expert and six university students, using generative AI-based storyboarding and low-fidelity prototyping to explore interface expectations and cognitive workflows. Key insights include the need for ergonomic design, modular scaffolding, fluid transitions between real and virtual spaces, and context-sensitive multimodal feedback. The study contributes empirically grounded design recommendations for building seamless learning environments that blend physical and metaverse contexts to support learner agency, reflection, and spatial interaction.

Keywords: Co-design, Mixed Reality Learning, Smart Textbooks, Educational Interface Design, Phygital Learning, Metaverse Technologies

1. Introduction

Textbooks remain central to classroom learning in many parts of the world, especially in resource-constrained settings such as Indian public schools. However, their linear structure and static content often limit learners' opportunities for exploration, contextualization, and interaction, capabilities that are increasingly important in 21st-century science education. With the emergence of intelligent systems and immersive technologies, there is growing potential to augment printed materials with digital overlays, AI-generated explanations, and embodied experiences that support both conceptual understanding and learner agency.

This paper explores how metaverse-based technologies, specifically mobile AI tools and Mixed Reality (MR) environments, can be integrated with existing school infrastructure to create phygital learning experiences. By embedding AI-driven feedback, multimodal content, and gesture-based interaction into the textbook reading process, we aim to create seamless transitions between physical and digital modes of engagement (Prajapati & Das, 2023; Saindane et al., 2023).

We present the early-stage design of the INTERACT Suite, a system of AI-augmented interfaces that reimagines school textbooks as active learning companions. The suite consists of:

1. *ext-book*: A mobile application that enables scanning of textbook content to access AI-powered explanations, media overlays, and note-taking features.
2. *Interactouch*: A MR prototype that anchors AI agents and educational media to physical textbooks using gesture-controlled interfaces on head-mounted displays.

To guide the development of these tools, we conducted a co-design study with one biology subject matter expert and six university students. Using generative AI-supported storyboarding and low-fidelity prototypes, participants engaged in reflective design activities to articulate their expectations, cognitive needs, and usability concerns. The co-design process surfaced several design imperatives for developing seamless, embodied, and adaptive learning interfaces grounded in the realities of Indian classrooms. The design of the INTERACT Suite was motivated by three gaps in current research: the absence of ergonomic strategies for managing transitions between mobile and immersive devices, the limited personalization available in AI-augmented textbooks, and the lack of participatory design studies situated in Global South classrooms. Addressing these gaps required co-designing with learners and subject experts to capture their expectations and constraints, thereby ensuring that the system remains pedagogically relevant and contextually feasible.

In this paper, we describe the co-design methodology, present selected user interface designs, and reflect on the broader implications for seamless metaverse learning. Our goal is to contribute design insights and interface concepts that bridge the gap between physical textbooks and emerging metaverse-based educational technologies.

2. Background and Related Work

2.1 Smart Textbooks and AI-Augmented Learning

Smart textbooks bridge the gap between traditional print and digital interactivity by embedding features such as multimedia overlays, adaptive feedback, and learner-driven annotations. These systems are grounded in constructivist learning theory (Piaget, 1971; Vygotsky, 1978) and support personalized inquiry by allowing students to explore content in nonlinear, multimodal ways. Recent innovations in generative AI further extend these capabilities, enabling on-demand explanations, simplified vocabulary, and visual scaffolds that adapt to learners' contexts and queries. Empirical studies show that AR-enhanced and AI-enriched textbooks improve comprehension, engagement, and motivation in science subjects like biology (Chang et al., 2019; Koć-Januchta et al., 2020; Prajapati & Das, 2025). However, challenges persist regarding usability, cognitive overload, and the seamless integration of AI feedback within the textbook interaction flow (Yang et al., 2020; Ali et al., 2024).

2.2 Mixed Reality and Seamless Learning in the Metaverse

MR learning environments allow for spatial interaction with both digital and physical content, often leveraging gesture recognition, gaze tracking, and immersive visualization. These systems are particularly effective in conceptual domains that benefit from spatial reasoning and embodied cognition, such as cellular biology, physics, or geometry (Parong & Mayer, 2018; Li et al., 2019). The Cognitive Affective Model of Immersive Learning (CAMIL) (Makransky & Petersen, 2021) provides a theoretical basis for designing MR experiences that support presence, agency, and emotional engagement while managing cognitive load. Complementing this, the XR-Ed framework (Yang et al., 2020) offers design guidelines across six instructional dimensions - learner agency, feedback, modality, interactivity, sociality, and accessibility, tailored for MR in educational settings.

2.3 Co-Design and Participatory Development in TEL

Co-design methods in TEL foreground user participation (learners, educators, or domain experts), in shaping educational tools that are usable, inclusive, and pedagogically relevant. In our project, co-design played a central role in surfacing workflow expectations, ergonomic constraints, and multimodal needs specific to hybrid mobile–MR textbook use. Participants engaged in reflective storyboarding, low-fidelity prototype interactions, and scenario walkthroughs to iteratively refine features such as AI feedback panels, annotation tools, and gesture-driven overlays. Design recommendations emerging from these sessions include support for short modular interactions, progressive information disclosure, and clear transitions between physical and immersive tasks.

3. System Overview

3.1 INTERACT Suite and ext-book: AI-Augmented Mobile Textbook Interface

The INTERACT Suite is conceived as a unified phygital ecosystem that extends textbook use across mobile and MR platforms. Although informed by our earlier work on phygital textbooks (Prajapati & Das, 2023, 2025; Saindane et al., 2023) the unique value of the present design lies in its three interrelated features: (i) AI-powered explanations with multimodal scaffolds that adapt to different content types, (ii) gesture-driven spatial interactions that anchor learning to the physical textbook while enabling immersive overlays, and (iii) cross-device continuity through synchronized annotations, AI conversations, and shared learner models. These features distinguish INTERACT from prior smart textbook and XR systems by emphasizing modular engagement, ergonomic design, and seamless transitions between physical and virtual learning spaces.

The *ext-book* application allows students to interact with textbook content using their mobile devices by scanning printed text, diagrams, or exam-style questions. The application offers three key interaction modes: (1) Text Mode: Recognizes scanned paragraphs and generates contextual explanations using a large language model. It also provides options to simplify text, retrieve definitions, or summarize concepts. (2) Image Mode: On recognizing textbook diagrams or illustrations, the app overlays voice-narrated descriptions, linked videos, and 3D animations to support visual understanding. (3) General Mode: Designed for unstructured content such as exam questions, this mode offers scaffolded generative responses through "Explain," "Answer," and "More" buttons. These afford learners different levels of support and elaboration. Each interaction result can be saved into a unified Notebook View, enabling learners to revisit, organize, and reflect on their AI-generated content. This design supports self-regulated learning and is informed by Universal Design for Learning (UDL) principles (CAST, 2023), which emphasize flexible representation and learner autonomy. Figure 1 shows the application architecture.

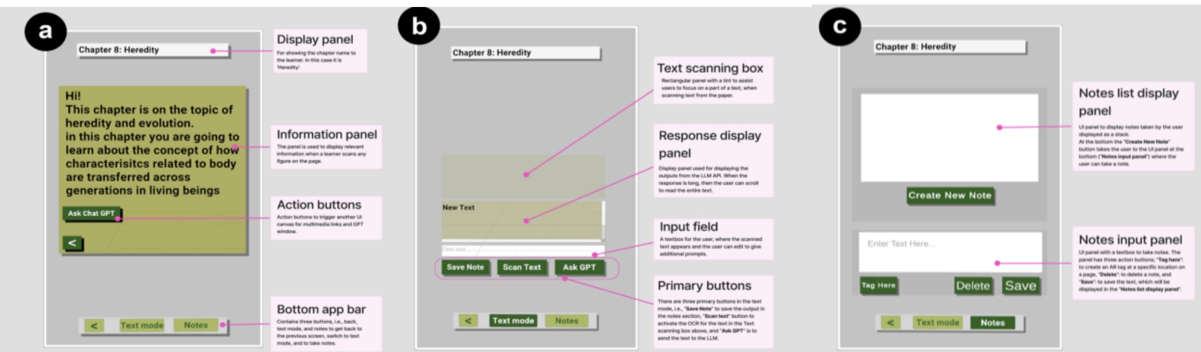

Figure 1: The application architecture integrates: (a) OCR and text/image detection using Google Cloud Vision API, (b) Generative AI responses through the OpenAI GPT API, and (c) Real-time analytics and logs via Google Firebase, which also supports future adaptive learning analytics.

3.2 Interactouch: Mixed Reality Textbook Overlay

The *Interactouch* prototype extends the *ext-book* experience into MR, enabling embodied interaction with spatially anchored educational media. Designed for head-mounted displays (e.g., Meta Quest 3), the system recognizes specific textbook pages and overlays corresponding 3D models, floating media panels, and interactive feedback windows. Key features include: (a) Gesture-Based Navigation: Learners use hand gestures to rotate, zoom, or activate multimedia layers (e.g., diagrams with narration, embedded AI chat), (b) Contextual

Anchoring: Media elements are linked to printed textbook locations, supporting continuity between physical and virtual spaces, and (c) Floating Interface Panels: The MR environment provides access to saved annotations, AI explanations, and multimedia visualizations in modular layers. Figure 2 shows a transition from physical to immersive digital experiences using *Interactouch* prototype.

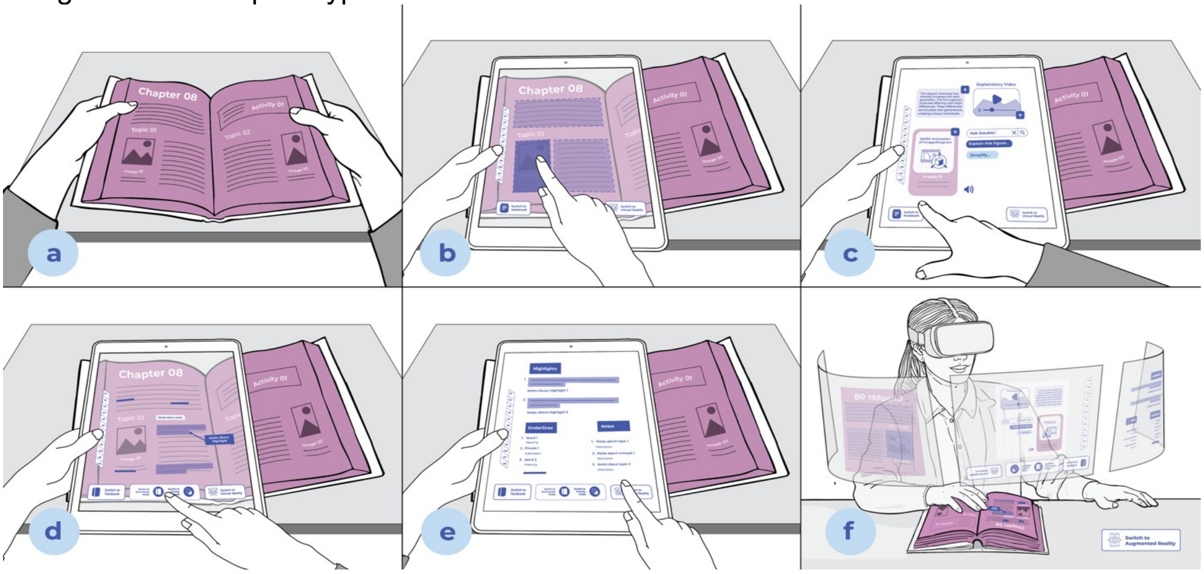

Figure 2. A typical navigation depicting a transition from physical to immersive digital experiences. The six frames represent key steps in the user journey, from textbook recognition, interaction with content, annotation features, to the switch to a mixed reality environment (Interactouch).

4. Co-Design Study

To guide the development of the INTERACT Suite, we conducted a co-design study with the aim of capturing user expectations, usability concerns, and workflow insights for AI-augmented textbook systems that span both mobile and MR platforms. This early-stage study focused not only on interface preferences but also on the cognitive and ergonomic challenges of transitioning between physical and digital modes in the classroom.

4.1 Participants and Methodology

The co-design study involved one subject matter expert (SME) in biology and six university students (n = 6) with backgrounds in science education or HCI. The participants engaged in a three-part design workshop combining narrative exploration, prototype interaction, and reflective feedback.

1. AI-Assisted Storyboarding - Participants were introduced to a series of generative AI-created storyboards simulating academic reading scenarios. Each storyboard depicts a student using a printed textbook, scanning content with a mobile app, engaging in AI-supported conversations, and transitioning into a spatial MR environment. These visual narratives grounded the session in realistic learner workflows and sparked discussions on affordances, barriers, and expectations.
2. Low-Fidelity Prototype Walkthrough - Using a click-through prototype of the *ext-book* application built with Adobe Aero, participants explored proposed interface features including scanning modes, scaffolded AI buttons, media overlays, and the notebook view. Though not fully functional, the prototype mimicked the user's flow across different modalities and helped participants assess cognitive load and interaction coherence.
3. Reflective Discussion and Thematic Elicitation - Each session concluded with guided reflection on design affordances, cross-platform transitions, usability friction, and

learner agency. Notes were thematically clustered to identify common concerns and design recommendations.

All reflective discussions were audio-recorded and transcribed. The transcripts, together with researcher notes, formed the primary data sources. We conducted inductive thematic analysis in three stages: (i) open coding of participant utterances and notes, (ii) clustering of codes into provisional categories by two independent researchers, and (iii) refinement of categories through consensus-building discussions. This process ensured that the resulting design themes were empirically grounded and captured both convergent and divergent perspectives from participants.

4.2 Design Insights and Themes

Analysis of participant feedback revealed four core design themes relevant to metaverse-aligned educational systems:

1. Seamless Real–Digital Transitions: Participants emphasized the importance of minimizing friction when switching from reading physical books to engaging with digital overlays or immersive content. Context-sensitive scaffolds, QR codes, and synchronized progress tracking were suggested as aids.
2. Ergonomic Constraints and Cognitive Load: Sustained use of phones or head-mounted displays (HMDs) for reading-intensive tasks was viewed as potentially fatiguing. Participants preferred modular interactions with short bursts of immersive engagement rather than long, continuous sessions.
3. Multimodal Scaffolding and Visual Hierarchy: Learners valued layered information delivery, starting with essential explanations and gradually revealing detailed content (videos, 3D models). Overcrowded interfaces or overlapping overlays were flagged as overwhelming, especially in MR.
4. Cross-Device Continuity: Participants expressed interest in being able to start on mobile, explore deeper in MR, and return to a unified notebook or dashboard for revision. They recommended features like synced bookmarks, portable annotation layers, and persistent AI chat history.

5. Design Implications for Metaverse Learning

5.1 Design for Modularity and Transition

Metaverse learning tools should support modular, task-specific interactions that allow learners to enter and exit immersive or AI-augmented states without disrupting the learning flow. In our case, learners preferred short, focused interactions (e.g., scanning a paragraph, receiving an AI explanation, viewing a short animation) over continuous immersion. Seamless transitions between the mobile and MR interfaces, anchored to the same physical textbook context, emerged as a critical design requirement.

5.2 Continuity Across Physical and Virtual Devices

To fulfill the vision of phygital continuity, systems must be designed to track learning across platforms. Annotations, bookmarks, and AI conversations initiated on a mobile phone should persist into the MR environment and vice versa. This implies a need for shared learner models, synchronized content anchors, and cloud-based state management, all of which were anticipated in the architecture of the INTERACT Suite.

5.3 Learner-Centric Personalization and Agency

Metaverse learning environments must be adaptive to the learner's goals and rhythms. The co-design study revealed that learners value control over how they engage with content (e.g., starting with AI explanations, then exploring visuals, or generating notes). Designing interfaces that foreground learner agency, while still offering scaffolded support, is essential to balancing autonomy with structure in hybrid physical–digital environments.

6. Conclusion and Next Steps

This paper presented a co-design-led exploration of how printed school textbooks can be transformed into AI-augmented, multimodal interfaces that support seamless transitions between physical and virtual learning spaces. Through the development of the INTERACT Suite, comprising the *ext-book* mobile application and the *Interactouch* MR prototype, we demonstrated how design thinking, participatory methods, and theoretical grounding in UDL, CAMIL, and XR-Ed can inform the creation of intelligent phygital learning environments. In the next phase of this work, we aim to deploy the ext-book application in classroom settings to evaluate its usability and learning impact with a larger student population. Usability testing of the *Interactouch* prototype will focus on gesture interaction and spatial layout refinement. We also plan to implement a shared learner model to enable cross-device continuity and expand our co-design activities to include school teachers and students, ensuring curricular alignment and contextual relevance.

References

Ali, D., Fatemi, Y., Boskabadi, E., Nikfar, M., Ugwuoke, J., & Ali, H. (2024). ChatGPT in teaching and learning: A systematic review. Education Sciences, 14(6), 643. https://doi.org/10.3390/educsci14060643

CAST. (2023). Universal design for learning guidelines version 3.0. https://udlguidelines.cast.org/ Chang, K. E., Chang, C. T., Hou, H. T., Sung, Y. T., Chao, H. L., & Lee, C. M. (2019). Development and behavioral pattern analysis of a mobile guide system with augmented reality for painting appreciation instruction in an art museum. Computers & Education, 71, 185–197. https://doi.org/10.1016/j.compedu.2013.09.007

Koć-Januchta, M. M., Schönborn, K. J., Roehrig, C., Chaudhri, V. K., Tibell, L. A. E., & Heller, H. C. (2020). Connecting concepts helps put main ideas together: Cognitive load and usability in learning biology with an AI-enriched textbook. International Journal of Educational Technology in Higher Education, 17(1), Article 1. https://doi.org/10.1186/s41239-019-0174-1

Li, W., Liu, X., Zhang, Q., Zhou, B., & Wang, B. (2019). VR-enhanced cognitive learning: Method, framework, and application. Applied Sciences, 9(8), 1712. https://doi.org/10.3390/app9081712

Makransky, G., & Petersen, G. B. (2021). The Cognitive Affective Model of Immersive Learning (CAMIL): A theoretical framework for immersive learning design. Educational Psychology Review, 33(3), 937–965. https://doi.org/10.1007/s10648-020-09586-2

Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. Journal of Educational Psychology, 110(6), 785–797. https://doi.org/10.1037/edu0000241

Piaget, J. (1971). The theory of stages in cognitive development. McGraw-Hill.

Prajapati, S. P., & Das, S. (2023). Designing a Phygital Interface for Textbooks to Support Active Collaborative Learning. *28th International Conference on Intelligent User Interfaces*, 1–4. https://doi.org/10.1145/3581754.3584117

Prajapati, S. P., & Das, S. (2025). Understanding Students' Active Reading in Phygital Learning Environments: A Study of Smartphone-Based Textbook Companions in Indian Classrooms. *Frontiers in Education*, *10*. https://doi.org/10.3389/feduc.2025.1660133

Saindane, D., Prajapati, S. P., & Das, S. (2023). *Converting Physical Textbooks into Interactive and Immersive'Phygital'Textbooks: A Proposed System Architecture Design for Textbook Companion Apps.* 88–93.

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.

Yang, K., Zhou, X., & Radu, I. (2020). XR-Ed framework: Designing instruction-driven and learner-centered extended reality systems for education. Smart Learning Environments, 7(1), Article 13. https://doi.org/10.1186/s40561-020-00131-x