RESEARCH ARTICLE

BON VIEW PUBLISHING

Tuning into Learning: Student Narratives from an Augmented Podcasts Intervention

Ajita Deshmukh[1,2,]* , **Dhairya Pandya[2]** and **Sameer Sahasrabudhe[2]**

[1] School of Education and Research, MIT Art Design Technology University, India

[2] Design, Indian Institute of Technology, India

Abstract: This study explores learner experiences with augmented podcasts (APs), an innovative multimodal learning format that uses two formats in this case: MP3 file + PDF file of the screenshots used in the lecture. Situated within a larger mixed method study, this qualitative analysis is an extension of the earlier quantitative findings that evaluated effectiveness of APs through test scores. The present study delves into the uncaptured nuances of learner engagement—examining how students interact with, experience, and make meaning from this novel format. Using purposive sampling from Tier 2 and Tier 3 Indian cities, the study makes use of interviews conducted with 20 students to identify perceived facilitators and barriers to AP use. Findings of this study reveal that this experience of learning using APs which encourage learner autonomy, fostered enhanced cognitive engagement, multisensory integration, and deeper involvement. It was noted that students valued the format's integrated cues and layered delivery, which encouraged focused learning and sustained attention. However, issues with pacing, language accessibility, and digital fatigue were mentioned by some students. This study adds to the growing discourse about inclusive, scalable, and learner-informed digital innovations in education by elevating the voices of students while confirming the pedagogical value of APs.

Keywords: educational podcasts, augmented podcasts, multimodal learning, digital learning, learner experience, educational technology, qualitative analysis, digital divide, cognitive engagement

1. Background and Context

The rapid expansion of digital education has led to ongoing efforts to develop instructional strategies that are both effective and accessible, catering to diverse learner needs and technological constraints. Online learning, particularly through Massive Open Online Courses (MOOCs) and other digital platforms, has established video-based content as a dominant format. However, internet connectivity disparities remain a significant barrier to equitable access. Addressing these challenges requires innovative, technology-driven solutions that enhance accessibility without compromising learner engagement.

At the same time, technology providers continue to develop user-friendly, multi-platform tools that allow users to consume educational content in diverse formats [1]. With exponential advancements in media integration and digital convergence, the education sector has seen new possibilities for incorporating alternative learning formats [2]. As learning environments become increasingly asynchronous and learner-centric, there is growing demand for instructional methods that offer greater flexibility and autonomy.

One such asynchronous medium gaining traction is the augmented podcast (AP), an evolution of traditional podcasts that integrates audio narration with synchronized slides to enhance comprehension and engagement. The widespread availability of podcasting platforms, including those supported by major technology companies, underscores their popularity. While the audio format has historical roots in radio broadcasting, its resurgence in the form of podcasts has provided educators with an accessible and adaptable learning tool [3]. Forbes and Khoo [4] have explored the educational applications of various podcast formats, yet challenges remain, such as lack of visual cues, potential distractions, and limited interactivity.

Despite these limitations, podcasts have been widely adopted as an educational tool, prompting further research into their effectiveness, learner engagement, and cognitive impact [5]. Recognizing the need for an alternative that mitigates bandwidth limitations while maintaining learning effectiveness, this study investigates the use of APs—a hybrid model that integrates audio narration with synchronized slides. This approach aims to bridge the digital divide by offering a low-bandwidth, high-engagement learning solution.

2. Literature Review

The section of Literature review explores three key areas that frame the study: the rise of digital learning trends, the evolution of the role of podcasts in education, and theoretical foundations that underpin the development of APs.

1) The Rise of Digital Learning & Asynchronous Education: There is a growing demand for flexible, self-paced learning, acceptance of MOOCs, particularly after the COVID-19 pandemic.

_____

*Corresponding author: Ajita Deshmukh, School of Education and Research, MIT Art Design Technology University, India and Design, Indian Institute of Technology, India. Email: ajita.deshmukh@mituniversity.edu.in

This shift highlights the importance of learner autonomy and responsive instructional design in the asynchronous educational process.

2) Evolution of the Podcasts and Audio-Based Learning in Education: This section addresses the evolution of podcasts as an educational tool. Their accessibility and ease of use have contributed to their widespread adoption. However, the exclusive reliance on audio alone may restrict comprehension and retention of information.

3) Implementing APs: An AP is a media format designed for this intervention, integrating audio content with visual slides. Users listen to the audio file and change the slide upon receiving cues from the audio. This multimedia format presents a solution to the audio alone nature of the podcast by utilizing principles of multimedia learning.

Figure 1 outlines three primary sections dealt with in the Literature review section. They are the rise of digital learning and asynchronous education, evolution of the podcasts and audio-based learning in education, and implementing APs: A rationale for enhanced learning.

2.1. The rise of digital learning and asynchronous education

This section discusses the emergence of MOOCs and online learning, emphasizing the significance of learner autonomy and self-paced learning within these educational models. It reflects on various studies that highlight key aspects of MOOCs and examines how researchers have explored the differences between asynchronous and synchronous learning, as well as students' preferences between these two learning methods. Additionally, it addresses the benefits associated with asynchronous learning, illustrating its advantages in fostering a more personalized educational experience.

2.1.1. The rise of MOOCs and online learning

Over the past few years following the COVID-19 pandemic, MOOCs and online teaching-learning have experienced substantial growth, coinciding with significant technological advancements. MOOCs are now recognized as a tool for enhancing traditional education models by improving students' learning efficiency through technology-driven teaching methods. In this context, learning efficiency is the ability to achieve higher learning outcomes while utilizing minimal resources such as time and effort [6]. Another significant advantage of online learning is the flexibility it offers, allowing students to study from any location without being restricted by geographical boundaries [7].
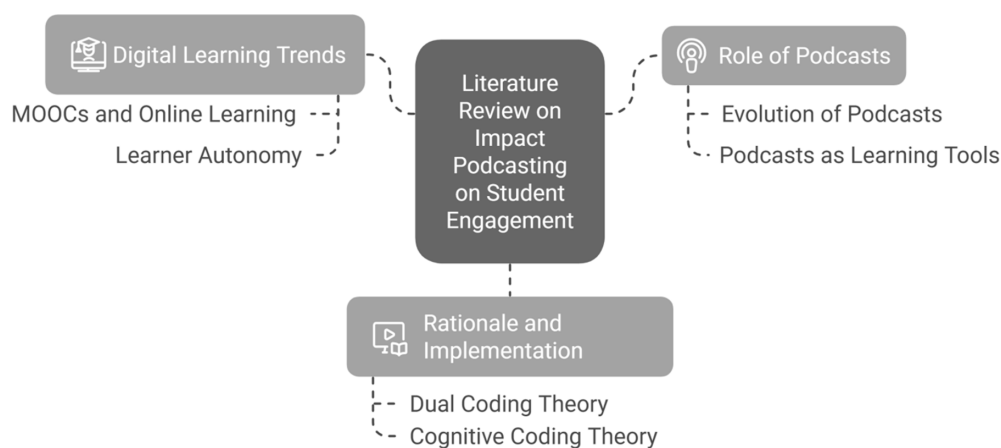
2.1.2. Enhancing student ownership through asynchronous learning

A systematic literature review on a large dataset indicated that while asynchronous learning was slightly more effective in enhancing student knowledge compared to synchronous online learning, the effect size was minimal. The self-paced nature of asynchronous learning not only enhances learner autonomy but also fosters a sense of ownership over the educational experience, allowing students to tailor their learning journeys to their unique needs and preferences [7]. According to Self-Determination Theory, learning at one's own speed promotes both competence and autonomy [8]. Kearney and Hartnett [9] have highlighted numerous studies demonstrating the positive outcomes associated with autonomy and ownership in learning, particularly through self-directed learning models. These benefits include increased motivation, the ability to connect formal and informal learning experiences as well as formal education and professional needs, and a sense of empowerment that extends into learners' professional lives. Additional advantages include improved metacognitive abilities, greater responsibility for learning, and an expanded knowledge base. Perera and Richardson [10] previously examined these aspects and emphasized the need to investigate which types of learning materials lead to better outcomes.

2.1.3. Advancements in educational media creation

Building on the previously discussed term of "learning efficacy," Billsberry and Alony [11] further expand on this term of learning efficiency from an educator's perspective, highlighting how technological advancements have simplified the creation of educational resources. With video-making becoming more user-friendly, cost-effective, and capable of producing high-quality content, individual academics can now develop instructional materials with minimal support. Given these advancements, video has emerged as a dominant medium in MOOCs. Deng [12] notes that video lectures serve as an effective instructional strategy for enhancing social engagement. However, for MOOCs to be truly effective, they must incorporate multiple channels for knowledge dissemination. The integration of animations, simulations, and other interactive elements can enrich MOOCs, thereby positively influencing learners'

Figure 1
Thematic structure of the literature review

and educators' behavioral intentions and actual usage, particularly in emerging fields such as health informatics [13].

2.2. Evolution of the podcasts and audio-based learning in education

This section explores the evolution of podcasts over time and their utilization as an educational tool. It studies the journey of the podcast and has investigated the effectiveness of podcasts in education, highlighting the reasons why learners prefer this medium. The discussion emphasizes the various advantages that podcasts offer, which contribute to their appeal as a valuable resource for learning.

2.2.1. The evolution and popularity of podcasts

The evolution of podcasting as an educational tool has been substantial. An article by Minooka [14] outlines this journey. The emergence of podcasting as a significant media format in the digital age can be traced back to the early 2000s with the convergence of the iPod and RSS (Really Simple Syndication) technology, leading to the coinage of the term "podcast" in 2004. Initially, an informal medium driven by independent creators and niche interests, podcasting's growth significantly accelerated with the proliferation of smartphones and user-friendly mobile applications, enabling broader access for a mainstream audience. This evolution marked a departure from traditional radio by offering a flexible, on-demand listening experience that caters to modern lifestyles. This journey continues as podcasts have now evolved into valuable tools for data collection in academic research, with a recent article outlining a seven-step methodology for analyzing their content to gain diverse insights [15]. Several factors fueled this expansion, including technological advancements that enhanced accessibility, the flexibility and convenience of consuming content anytime and anywhere, and the capacity of podcasts to foster intimate connections and active engagement between hosts and listeners, often building communities around specific topics [14].

2.2.2. Podcasts: A powerful learning tool

The use of audio as a medium for delivering information is not a recent phenomenon; it has long been practiced through radio broadcasts and, more recently, through podcasts available on digital platforms. Audio-based information dissemination is not a new concept; it has been utilized for years through radio broadcasts and, more recently, through podcasts on digital platforms. Podcasts have become a popular medium for both entertainment and learning, as they provide a unique blend of convenience, engagement, and accessibility [16]. A key reason for their appeal is their flexibility—listeners can consume content at their own pace, whether for leisure or educational purposes. Unlike traditional learning materials, podcasts allow multitasking, making them an ideal medium for individuals with busy lifestyles. Evidence from various studies supports the growing role of podcasts in information consumption. Tobin and Guadagno [17] highlight how individuals frequently integrate podcasts into their daily routines. A qualitative study by McCarthy et al. [18] found that people often listen to podcasts while commuting, performing household chores, or engaging in other activities. This ability to multitask while learning enhances their sense of productivity and knowledge acquisition, indicating both their familiarity with and preference for this format.

As podcasts have gained traction, their role in education has also been explored. Köhler et al. [19] reference multiple studies that demonstrate the positive impact of podcasts on learner perceptions. Their ease of use and flexibility make them particularly effective

as learning tools. For instance, a study conducted with 200 first-level students who were given a series of revision podcasts after completing a course in Information and Communications Technology found that students considered podcasts to be more effective revision tools than textbooks. They also found podcasts to be more efficient than their own notes in reinforcing learning [20]. Similarly, a qualitative investigation into the implementation of study casts in a middle school social study, found that students showed heightened motivation and engagement when using podcasts for test preparation [21].

These insights further reinforce the idea that podcasts can be impactful educational resources, promoting a more interactive and effective learning experience.

2.3. Implementing augmented podcasts: A rationale for enhanced learning

This section explores the principles of the AP. It examines various research studies that provide the rationale for integrating slides with podcasts. Additionally, it discusses Dual Coding Theory and Cognitive Load Theory, highlighting how these frameworks inform the functionality of APs. The incorporation of slide changes while listening to the podcast is developed with these theories in mind, enhancing the overall learning experience.

2.3.1. Integrating audio and slides for enhanced learning

Video is a well-established educational medium, and a systematic review by Mutlu-Bayraktar et al. [22] found that pairing graphics with audio narration is more effective than pairing graphics with text, as it leverages both visual and auditory channels to reduce cognitive load. In contrast, traditional podcasts lack this dual-channel benefit. To bridge this gap, APs, which combine audio narration with synchronized visual slides—have emerged as a promising alternative. A recent study by Sahasrabudhe et al. [23] found that APs are as effective as video for learning.

This innovation prompts a closer look at the pedagogical theories underpinning such media. By integrating audio and visuals, APs align with dual cding theory (DCT) and cognitive load theory (CLT), offering a theoretical and empirical foundation for their effectiveness in digital learning environments.

2.3.2. Dual coding theory in educational contexts

The DCT, proposed by Clark and Paivio [24], posits that the human mind processes information through two separate systems: verbal and visual. These systems interact to integrate linguistic and visual inputs, enhancing learning and recall through imagery and associative processing. DCT highlights the importance of structuring knowledge with both text and visuals to deepen understanding and retention. It also aligns with cognitive load principles—combining visuals and text helps reduce extraneous load and support meaningful learning. A meta-analysis by Mutlu-Bayraktar et al. [22] confirms that using signaling techniques lowers cognitive load and boosts comprehension.

A study testing DCT in statistical learning compared three conditions: text-only, text with diagrams, and diagrams with voice-over. The group exposed to diagrams with voice-over demonstrated the highest recall, affirming that audio-visual combinations enhance learning better than text alone or static visuals [25].

Overall, these studies underscore the importance of engaging both the visual and verbal systems in the learning process, suggesting that adding slides to audio presentations can enhance comprehension and retention.

2.3.3. Cognitive load theory in multimedia learning

Effective instructional design should optimize cognitive load, facilitate information transfer, and promote deep understanding to support meaningful learning outcomes [26]. Further researchers have pointed out several pitfalls to avoid extraneous cognitive load on working memory, one of them is split attention—like any audio with visuals that are difficult to read and understand or contain unnecessary information can hinder learning [27].

Transient effect iterated in CLT where temporary information (e.g., spoken words or animations that disappear) increases cognitive load and impairs learning compared to permanent formats (e.g., written text or static visuals) was studied and results showed that the transient condition led to poorer performance on retention and transfer tasks [28]. A common thread across studies is that keeping instructional content simple and focused helps students grasp it better. This aligns with the idea that working memory can only handle so much at once, so clear, concise audio or visuals help learners to retain information [26, 29, 30]. Students often struggle to remember information they only hear once, like a lecture or podcast, because it vanishes fast from their minds. Using the right mix of audio and visuals to present the information helps students learn without getting lost.

However, as mentioned by Skulmowski and Xu [31], interactive learning media, immersion, realism, disfluency, and emotional design all have the same design feature in common that learners are required to invest a small amount of extraneous load in order to allow certain forms of germane processing to occur. This cost–benefit approach also requires a strong focus on choosing appropriate assessment methods. If these conditions are met, digital learning can be used to foster learning despite minor increases in extraneous load. The same study has also revealed that moderate degrees of (inter-)activity tend to be the most beneficial for learning. Building on the findings of Skulmowski and Xu [31], this study takes the next step by examining the task of changing slides while listening to audio may introduce an additional cognitive demand, it can also enhance engagement and reinforce learning by encouraging active participation.

3. Research Gap/ Need for Study

While the effectiveness of educational podcasts has been well documented, most existing studies primarily rely on quantitative measures, such as pre- and post-test scores, quiz performance, or final grades [32]. Although these findings indicate measurable learning gains, they often fail to capture learners' subjective experiences, contextual engagement, and emotional or cognitive responses to the medium [33].

A previous controlled experimental study by Sahasrabudhe et al. [23] compared the efficacy of educational podcasts with traditional video content. While the study yielded valuable insights, it did not explore learner preferences, motivations, or sociocultural dynamics that may influence the reception of podcasts.

Recognizing this limitation, recent research has highlighted the need for qualitative investigations. For example, Conroy and Kidd [34] examined how podcasts enhance teacher–learner rapport, suggesting that the format may foster a sense of approachability and credibility. Similarly, Brew-Girard et al. [35] explored the experiences of medical students creating peer-to-peer educational podcasts, uncovering both pedagogical benefits and implementation challenges.

In light of these developments, the current study adopts a qualitative approach to examine learners' experiences with educational podcasts. This perspective offers deeper insight into how podcasts function as pedagogical tools across diverse educational contexts.

3.1. Research questions

This study seeks to answer the following research questions:

1) How do learners describe their experiences and patterns of engagement with APs as a learning medium?
2) What perceived facilitators and barriers influence learners' use of APs?

3.2. The study

This section addresses the research questions, outlines the experimental design, and details the development of the AP.

3.2.1. Developing the augmented podcasts

The development of the APs involved a structured process of audio-visual repurposing of existing instructional content. Audio tracks were first extracted from selected video lectures of the SWAYAM course Fundamentals of Banking and Insurance and converted into MP3 format. These audio files were then enhanced using Filmora software to embed audio cues—typically bell sounds—at key transition points. These cues were strategically placed at moments where visual changes occurred in the original video, helping listeners synchronize audio content with corresponding visuals.

Simultaneously, relevant screenshots were captured from the videos to represent key concepts, slide transitions, or visual explanations. These were then compiled into a PDF document, arranged in the same sequence as the original lecture. Each PDF served as a visual counterpart to its corresponding audio track.

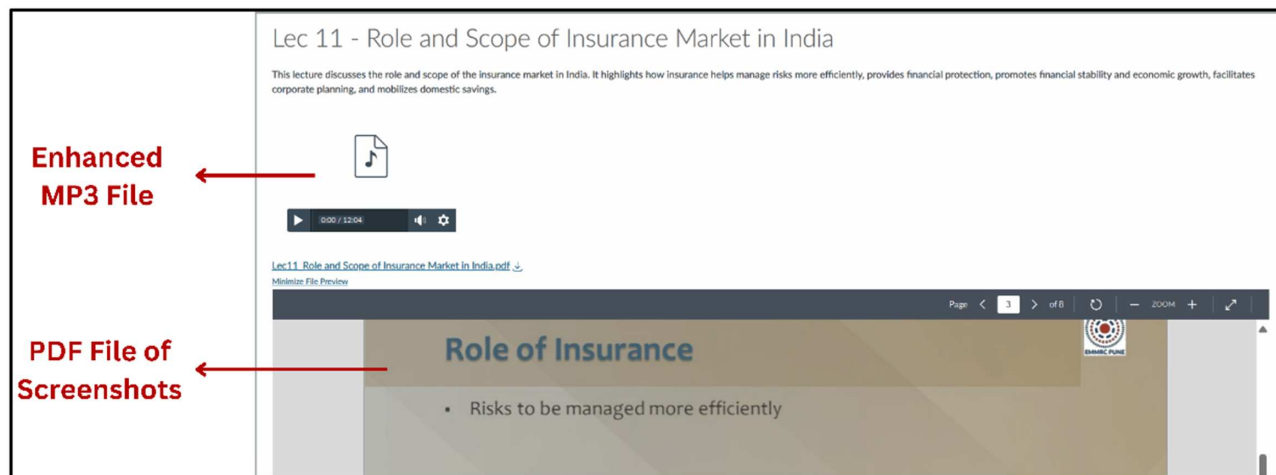The final AP consisted of two separate but coordinated components:

Figure 2 shows the structure of an AP in the CANVAS Learning Management System (LMS). Both files were uploaded to the CANVAS LMS on the same content page, with the audio placed above the PDF. The page also included the title of the lecture along with a concise description of its content. This layout allowed learners to listen to the audio while independently navigating the visuals, thus offering a multimodal, flexible learning experience.

3.2.2. The experiment

The experiment was piloted in one selected college of Mumbai. The choice of the city being deliberate in order to eliminate extraneous factors such as internet connectivity or infrastructural constraints. This ensured that the focus of the study remained on evaluating the effectiveness of the AP as a learning medium [23]. Building upon this foundation, the study was expanded to encompass a variety of contexts viz: two tier and three tier cities as well as field based and controlled institutions. This was aimed to assess the applicability and reception of AP across different learning environments.

The study centered around the delivery of 11 augmented lectures from the Fundamentals of Banking and Insurance course, originally hosted on SWAYAM. These 11 lectures were carefully selected in consultation with the course instructor and external

Figure 2
Layout and functional components of the augmented podcast on the CANVAS LMS

Note: A modified MP3 file containing the instructional audio and navigational cues

A PDF file featuring lecture visuals/screenshots

subject experts based on their pedagogical coherence, duration, and suitability for chunking. Each of the selected lectures was then converted into APs following the process described above.

The implementation of the study took place across two institutional settings:

1) Field Based Setting: In this setting, the students accessed the content over CANVAS LMS in their respective institutions. Prior to the intervention, an orientation session was held with each student group. This was to familiarize them with the LMS, guide them through the login process, demonstrate navigation of the augmented content, and address any technical issues. Students then completed the intervention from their respective campuses.

2) Controlled Institutional Setting: In this setting, the same content was delivered via the same LMS at the researcher's institution. Students accessed content individually, over a designated computer with a stable internet connection. Orientation session was conducted prior to the intervention, and technical support was given real time. This arrangement enabled real-time observation that led to invaluable insights through researchers' diary.

At the conclusion of the intervention, students of both the settings were given a structured feedback form to complete. The final item in this form invited participants to volunteer for a follow-up interview. Interested students were later contacted for interviews.

4. Methodology

To gain a deeper understanding of learners' experiences with APs as a learning format, a qualitative study was conducted utilizing semi-structured interviews. The participants included students who had engaged with APs derived from selected lectures of an online course. The interviews aimed to capture learners' levels of engagement, perceived usefulness, and any challenges they encountered. Notably, data were collected from students in two distinct learning environments: one offering remote, self-paced access and the other conducted in a structured, in-person institutional setting. This diversity in context provided a more nuanced understanding of how different learning environments influence learner engagement with digital content.

4.1. Methodology overview

This qualitative study was part of the larger sequential explanatory mixed-method research design. Following the quantitative study that tested the effectiveness of AP using test scores, the qualitative phase aimed to explore finer nuances of learner experience that could not be captured via quantitative analysis. Semi-structured interviews of 20 students who volunteered from contrasting learning environments: own institution and the research institution were conducted. This dual context format allowed for deeper insights on how learning environments shape engagement with different digital learning formats.

Thematic analysis was conducted through open coding, identifying patterns in engagement, comprehension, and format-related challenges. The research follows a constructivist approach, recognizing that learning is a subjective process shaped by prior knowledge, cognitive preferences, and environmental factors. Participants were selected based on their exposure to multiple digital learning formats, ensuring a diverse range of perspectives. Thematic coding allowed for the identification of key themes, learner behaviors, and instructional design considerations. Open coding was carried out independently by two researchers, systematically documenting manual coding resulting in a spreadsheet-based codebook. Codes that evolved were added in the codebook, and further analysis led to emergent themes. Cross validation ensured analytic rigor. Coding was iterative first then collaborative. Data saturation was deemed to have reached when no further themes evolved in final iteration. To mitigate bias, reflective memos were maintained, and regular debriefings were conducted within the research team to discuss emerging assumptions and ensure reflexivity. These steps were necessary in the absence of use of any software like NVivo for coding.

Through rigorous qualitative analysis, this study expands the discourse on multimedia learning, cognitive load theory, and digital instructional design. By capturing learner narratives, cognitive strategies, and engagement patterns, this research provides insights that complement the quantitative findings of the broader study, ensuring a holistic understanding of AP effectiveness. The findings aim to inform educators, content creators, and instructional designers on how to optimize digital learning environments for improved engagement and retention.

4.2. Participant profile/ sampling

A combined sampling approach was employed to balance contextual relevance with practical feasibility. Purposive sampling was used to select institutions from Tier 2 and Tier 3 cities, aligning with the study's focus on learners from non-metropolitan areas. Within this framework, convenience sampling guided the final selection based on institutional willingness, infrastructure availability, and scheduling compatibility. This ensured contextual diversity while addressing logistical constraints common in field-based educational research. Participating institutions were located in cities such as Mumbai, Pune, Satara, Bagalkot, Rajkot, Ahmedabad, and Gandhinagar. From these cities, students were invited to participate in the interviews on a voluntary basis. While students from Mumbai, Pune, Ahmedabad, and Gandhinagar could be interviewed, students from Rajkot, Satara, and Bagalkot were unable to participate in interview process due to examination schedules and other events at their respective institutions.

For the qualitative phase, 20 undergraduate students who had engaged with the podcast-based learning intervention and voluntarily opted for interviews were selected. The group included 8 males and 12 females, representing various academic years (from first to final year). Among the 20 students, 14 came from commerce background (6 males – 8 females), whereas 6 were from science background (2 males – 4 females). The goal was not statistical generalization but to gather rich, diverse insights into the perceived value, usability, and learning experience with podcasts.

Table 1 outlines the gender distribution across two research settings. The field-based setting included 4 students (2 males, 2 females), while the controlled setting comprised 16 students (6 males, 10 females).

Participant identifiers were anonymized using the format PFx.y or PMx.y, where "P" stands for participant, "M" or "F" denotes gender, "x" refers to the phase of the intervention (2 or 3), and "y" is the participant number within that phase. This system allowed easy referencing during analysis without compromising confidentiality.

This approach allowed the researchers to access a range of learner voices from different socio-academic contexts, providing depth and nuance to the exploration of podcasts as a pedagogical medium.

4.3. Data collection

This paper presents findings from the qualitative component of a larger mixed-methods study investigating learner engagement with APs. While the larger study gathered structured feedback from 249 students via a Google Form, this paper focuses specifically on semi-structured interviews conducted with a purposively selected subset of those respondents. The feedback form assessed learners' perceptions of the AP experience regarding layout, accessibility, material usefulness, achievement of learning objectives, and suggestions for enhancement. It also invited open-ended reflections on the most valuable aspects of their learning experience. At the end of the form, students were asked if they would be willing to participate in a follow-up interview.

Out of the 249 students who completed the form, 26 expressed their willingness to participate and were subsequently interviewed—approximately 10% of the total sample. This two-pronged approach—voluntary expression of interest followed by purposive sampling—was crafted to ensure representation across institutions in both 2-Tier and 3-Tier cities. The interviews provided an opportunity for a deeper exploration of learners' engagement with the medium, the influence of format on learning support, and emerging themes that transcended the structured questionnaire. This included personal preferences, comparisons with previous digital learning experiences, and contextual challenges.

Appointments were scheduled in collaboration with institutional staff, and interviews were conducted on-site during pre-decided time slots. On the day of the interview, consenting participants were assembled in a designated classroom. To ensure confidentiality and minimize peer influence, each student was interviewed individually in a separate room with only the interviewer present.

Interviews were audio-recorded using a mobile phone recorder after obtaining verbal and written consent. To uphold the integrity of the process, students who had already been interviewed were instructed not to interact with those awaiting their turn. Each interview lasted approximately 15–20 minutes and followed an unstructured format, thus generating around 20 * 20= 400 minutes of data. Each audio file was anonymized and saved using a systematic coding scheme. For example, the identifier "PM2.2" denoted a male participant from Phase 2, listed second among the interviewees. The files were initially stored on the interviewer's secure external hard disk and subsequently transferred to the password-protected desktop of the Principal Investigator at the research institute, where all research materials were stored in compliance with ethical standards for data security and confidentiality.
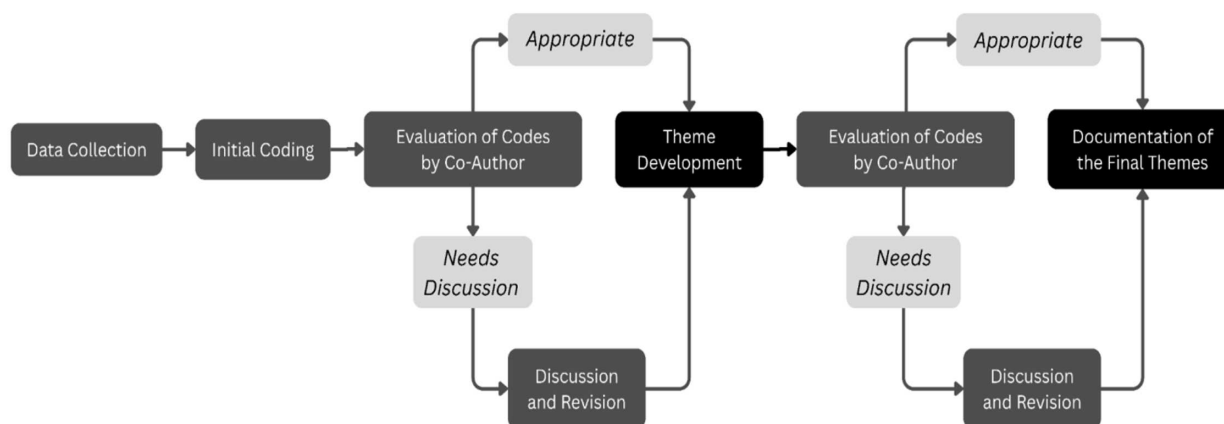
4.4. Qualitative data analysis: Coding and review process

To ensure methodological rigor and reduce the potential for researcher bias, the interview data underwent a systematic process of inductive thematic analysis. The data collection was carried out by one of the authors, who acted as the field investigator. The transcription and initial coding were undertaken by the primary researcher using an open coding approach, allowing codes

Table 1
Gender distribution of participants across field-based and controlled research settings

Gender distribution of students across research settings				
Sr. No.	Detail	Male	Female	Total Students
1	Students from field based setting	2	2	4
2	Students from controlled institutional setting	6	10	16
	Total students	8	12	20

Figure 3
Stepwise flowchart of the qualitative data analysis process

to emerge organically from the data without imposing a pre-existing framework.

To strengthen the credibility of the findings, the preliminary set of codes was independently reviewed by a co-author. This collaborative process involved multiple rounds of discussion to examine coding decisions, question underlying assumptions, and refine interpretations. A few codes were revised, merged, or reworded based on these peer debriefing sessions. This iterative dialogue helped establish inter-coder agreement and enhanced the analytical depth. Although NVivo or other qualitative software was not used, a detailed codebook was maintained using spreadsheets to document code definitions, frequencies, and relationships. This layered and collaborative approach ensured that the themes represented a balanced and reflective interpretation of the participants' voices. Figure 3 illustrates the process employed for the qualitative data analysis in this study.

5. Findings and Discussion

The analysis of learner experiences with APs revealed several key themes that shaped their engagement, comprehension, and preferences. By examining self-paced learning, cognitive load, multimedia integration, content length, and technical factors, this study uncovers how different aspects of digital learning impact student experiences.

The anonymized participant codes were used consistently during the coding process to trace responses back to intervention phases and demographic profiles, while maintaining participant confidentiality.

The thematic analysis identified six major themes, each addressing distinct aspects of the learning process. First, Learning Preferences & Engagement explores how interest, prior experience, and content structure influence engagement. Second, Cognitive Load & Processing Challenges highlight the mental effort required to balance audio and visual inputs. The third theme, Self-Paced & Flexible Learning, examines how learners navigate asynchronous learning formats, including multitasking and content chunking.

Moving beyond individual learning behaviors, the fourth theme, Role of Multimedia in Learning, focuses on how the combination of audio, slides, and visuals impacts retention and comprehension. Fifth, Challenges with Long-Form Content

addresses issues of fatigue and attention span in extended digital learning formats. Finally, Technical, Design & Instructor Presence explores how technical difficulties, instructional design, and the presence (or absence) of an instructor shape learner engagement.

These findings provide a nuanced understanding of learner interactions with APs, shedding light on both advantages and areas for improvement. The discussion that follows delves into each theme, supported by direct participant quotes, to offer deeper insights into their learning experiences. Figure 4 visually maps the thematic connections to RQ1 and RQ2, highlighting overlaps and distinctions. And Table 2 outlines which themes are associated with each research question (RQ1 and RQ2).

5.1. Enhanced cognitive engagement via audio-slide synchrony in augmented podcast

Learners indicated that the synchronized integration of audio and slides in the AP significantly enhanced their engagement and facilitated a deeper comprehension of the material. Instead of merely passively absorbing content, participants were required to concurrently process information across two distinct modalities. This dual engagement prompted increased cognitive effort, ultimately leading to improved retention of the information presented.

5.1.1. Audio-slide pairing promotes active cognitive processing
Learners consistently highlighted that the AP format required more active and sustained attention than traditional video-based materials. The integration of audio narration with visual slides created a dual-channel experience, which compelled them to remain mentally alert and cognitively engaged throughout the learning process. One participant remarked,
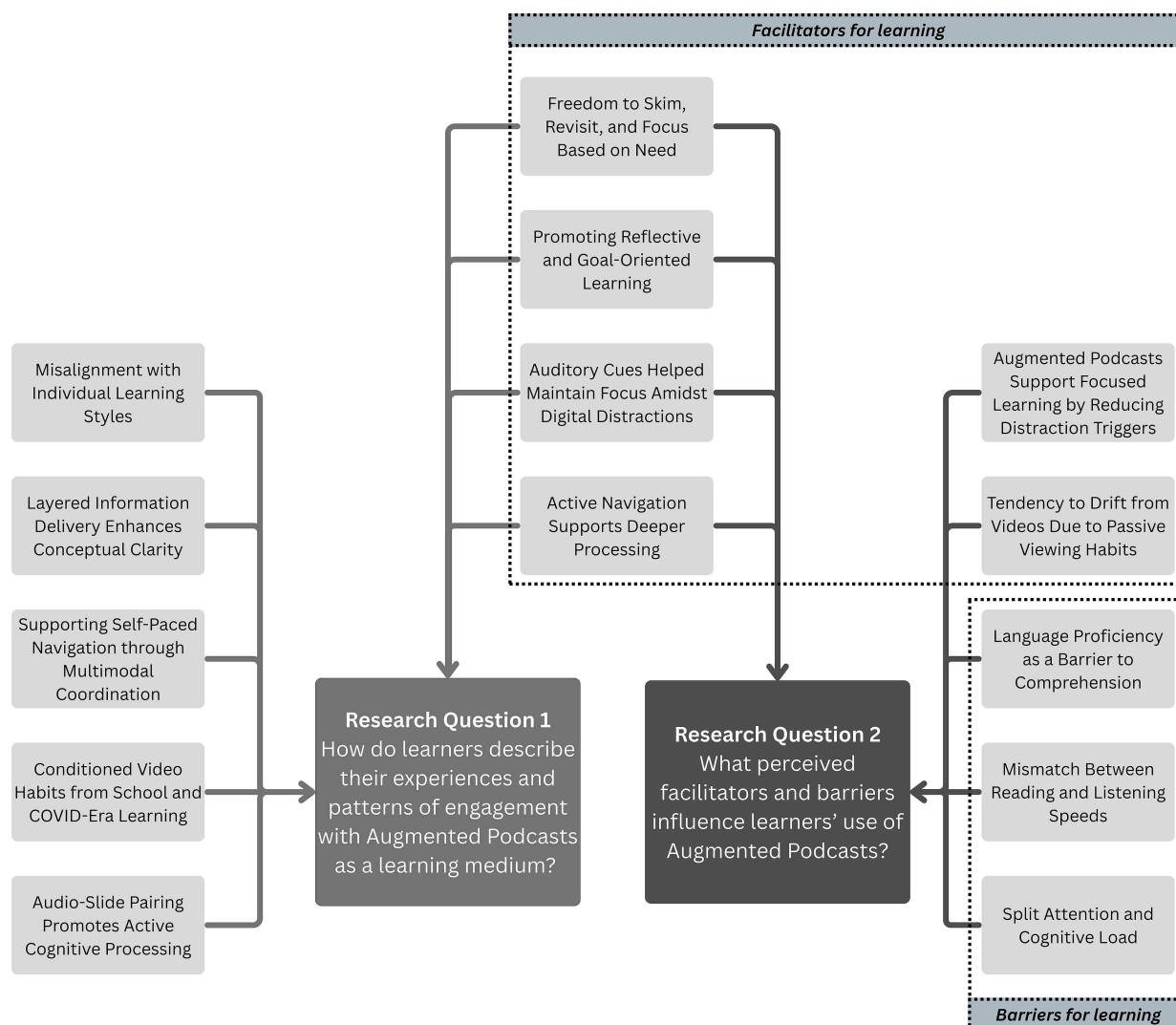
"When there was audio coupled with the PPT, I had to concentrate on both, which forced me to listen more attentively than I usually would in order to grasp the topic." (PF2.2)

"Because we can both read and listen with the slides, we can focus better on studying." (PF3.16)

"Because we can both read and listen with the slides, we can focus better on studying." (PF3.16)

"Because we can both read and listen with the slides, we can focus better on studying." (PF3.16)

Figure 4
Visual mapping of emergent themes to research questions

The necessity of coordinating what they heard with what they saw demanded a conscious effort, transforming the learning experience into a more intentional and immersive one. In contrast to passive video consumption, where learners often drift in and out of focus, the AP promoted a more concentrated mode of learning that cultivated a deeper understanding.

5.1.2. Layered information delivery enhances conceptual clarity

Participants highlighted how the APt format enabled a multifaceted understanding of complex topics by distributing content across complementary modalities. For instance, one learner remarked, "When studying the XYZ topic, X was explained through the slides, while YZ was further elaborated upon in the audio-visual format. This approach was quite beneficial." (PM2.1). Another remarked, "I think both audio and slide have their own benefits while learnt in this audio+slide combination" (PF2.2)

This strategic separation of visual and auditory content enhanced comprehension by allowing learners to absorb core concepts through one channel while receiving contextual elaboration through the other. The alignment between audio narration and slide content promoted integrative cognitive processing, moving beyond surface-level engagement. Notably, this cognitive involvement contrasts with learners' prior experiences with video content, which many found mentally disengaging—a theme further explored in the next section.

5.2. Video fatigue and digital drift

The transition from traditional video-based learning to the AP format seemed to address the digital fatigue often linked with lengthy videos. Participants noted that the passive habits developed during pandemic-era learning made it easy to lose focus on videos, while the AP required increased attention, thereby minimizing the likelihood of disengagement.

5.2.1. Tendency to drift from videos due to passive viewing habits

Participants frequently described a sense of detachment or boredom while watching traditional videos, often resulting in multitasking or abandoning the session midway. This contrasted with their experience of APs, where the need to toggle between slides and

Table 2
Alignment of emergent themes with research questions (RQ1 and RQ2)

Section	Subtheme	RQ1	RQ2
Enhanced Cognitive Engagement via Audio-Slide Synchrony in Augmented Podcast			
4.1.1	Audio-Slide Pairing Promotes Active Cognitive Processing	✓	
4.1.2	Layered Information Delivery Enhances Conceptual Clarity	✓	
Video Fatigue and Digital Drift			
4.2.1	Tendency to Drift from Videos Due to Passive Viewing Habits		✓
4.2.2	Augmented Podcasts Support Focused Learning by Reducing Distraction Triggers		✓
4.2.3	Conditioned Video Habits from School and COVID-Era Learning	✓	
Bell Sound Cue as a Prompt for Attention and Self-Regulation			
4.3.1	Auditory Cues Helped Maintain Focus Amidst Digital Distractions	✓	✓
4.3.2	Supporting Self-Paced Navigation through Multimodal Coordination	✓	
Learner Autonomy and Control in Navigating Content			
4.4.1	Active Navigation Supports Deeper Processing	✓	✓
4.4.2	Freedom to Skim, Revisit, and Focus Based on Need	✓	✓
4.4.3	Promoting Reflective and Goal-Oriented Learning	✓	✓
Barriers and Limitations to Augmented Podcast Engagement			
4.5.1	Misalignment with Individual Learning Styles	✓	
4.5.2	Split Attention and Cognitive Load		✓
4.5.3	Mismatch Between Reading and Listening Speeds		✓
4.5.4	Language Proficiency as a Barrier to Comprehension		✓

audio fostered sustained attention. One learner admitted, *"I would get bored watching the video, so I would put it aside and go to some other platform." (PF3.5)*

Some were candid to admit in different ways that conveyed the same thing *"There comes a moment while watching videos when you start to feel fatigued and bored, leading you to consider just leaving it." (PF3.5, PF3.14, PM3.15)*

These comments underline how video formats can inadvertently enable a "lean-back" posture that reduces cognitive engagement.

5.2.2. Augmented podcasts support focused learning by reducing distraction triggers

Unlike videos, which learners often associate with YouTube-style viewing that invites skimming, pausing, or background play, APs promote a more focused mode of engagement. As one participant observed, *"When the bell icon rang, I was reading the PPTs, so for me... audio plus PPT was excellent because you are listening, plus you are seeing, and you are reading." (PF3.5)*. Another stated about the required focus *"I had to focus on the audio as well as slide linkage with each other."* (PF2.2) while yet another remarked the focus shift, *"My main focus was on listening. Over there (in the slide), you are just reading the content, so you don't need to take that much load." (PF3.3)*

The absence of moving visuals may have helped reduce cognitive overload and visual fatigue, leading learners to remain more actively involved in meaning-making rather than drifting away.

5.2.3. Conditioned video habits from school and COVID-Era learning

Several learners acknowledged that their video preference stemmed not from its effectiveness, but from habitual exposure.

Different participants stated this effect in different ways. Some have a long exposure and conditioning since formative years *"I have been studying through video since 9th standard, so my preference is video," (PM3.4)*. While for some it is a conditioning due to compulsions of pandemic. *"After Corona, I've mostly studied for 10th and 12th grade by watching YouTube videos." (PM3.15)*. Another mentioned, *"These days after covid specially we have used videos and everything so this is different thing" (PF2.2)*

This indicates a form of conditioning, where the choice of modality is shaped by familiarity rather than perceived value—highlighting the potential of APs to disrupt passive habits and cultivate more deliberate learning practices.

5.3. Bell sound cue as a prompt for attention and self-regulation

The transition from traditional video-based learning to the AP format seemed to address the digital fatigue often linked with lengthy videos. Participants noted that the passive habits developed during pandemic-era learning made it easy to lose focus on videos, while the AP required increased attention, thereby minimizing the likelihood of disengagement.

5.3.1. Auditory cues helped maintain focus amidst digital distractions

Several participants pointed out the usefulness of the bell sound embedded in the AP as a prompt to shift attention back to the slides or to mark transitions. In a learning environment where distractions are just a click away, such cues serve as gentle nudges toward re-engagement. As one learner noted, *"I had to concentrate on both the audio and the slides." (PF3.1)*. Another noted how the bell sound helped to cut down distraction and proved as a cue to concentrate. *"The bell sound was actually good because whenever person gets distracted, so it was an alarm. For that you would need to concentrate." (PF2.6)*

This feedback illustrates how a small design element like a bell tone can act as a subtle yet effective anchor for learner focus.

5.3.2. Supporting self-paced navigation through multimodal coordination

The bell also supported a smoother user experience in navigating the dual-modality format of slides and audio. It provided a clear signal for when to shift or expect new information on the screen, thereby supporting learners' sense of autonomy and pacing. Quite a few students noted, "When the bell rang, it was time to change the slide" (PF3.1). This contributed to a perception of increased control over the learning process and a more structured flow, even in the absence of visuals found in conventional videos. This sense of self-regulation extended beyond auditory prompts. Learners increasingly valued the autonomy they had in managing their own learning paths—an aspect further unpacked in the next theme.

5.4. Learner autonomy and control in navigating content

The transition from traditional video-based learning to the AP format seemed to address the digital fatigue often linked with lengthy videos. Participants noted that the passive habits developed during pandemic-era learning made it easy to lose focus on videos, while the AP required increased attention, thereby minimizing the likelihood of disengagement.

5.4.1. Active navigation supports deeper processing

Learners consistently highlighted how the AP format empowered them to take control of their learning by allowing them to pause, revisit, or skim through content at their own pace. This interactive autonomy stood in contrast to the often passive experience of watching a linear video. One learner described, *"When there was audio and PPT itself, I had to concentrate on both and I was forced to listen to it more carefully than I usually would to understand what the topic is, where is the next slide I should move on to." (PF2.2)* Another brought out the passivity of videos. *"You can read it if you want to, but with videos, you just feel like skipping them." (PF3.12)*

This suggests that the dual-modality design encouraged a more intentional and self-directed approach to learning.

5.4.2. Freedom to skim, revisit, and focus based on need

The format allowed learners to engage more selectively, focusing only on parts they needed and revisiting complex portions. As one participant shared, *"If I understand something, I just listen to the audio. But if I don't understand it, I look at the slides, and then the concept becomes clear." (PF3.12)*. Another compared it with PDF, *"When we read a PDF, we put in a bit of our effort, which helps it stay in our minds for the long term." (PM2.5)*. Yet another brought out the advantage of AP, *"There are advantages to this (audio + slides) too, because if something is missed, we can listen to it a second time." (PM3.17)*

This illustrates how content control supports individual learning preferences and promotes deeper engagement by allowing learners to adapt the pace and sequence based on their understanding.

5.4.3. Promoting reflective and goal-oriented learning

Participants described a heightened sense of purpose in how they approached content, using the flexibility to align their learning with personal goals. For instance, one student remarked *"The audio and slides made all the main points clear and easy to understand. I already knew the basics, but this format helped me grasp the whole topic—even the more complex parts—much better than a live lecture would have." (PF3.10)*. Many students shared how personal interests and goals affected learning. *"I am interested in banking and finance. I did it because that's my personal goal." (PM2.1)*. Another agreed, specifying the topic further, *"I was feeling very comfortable with the slides and the audio as well because I didn't know actually, I wanted some knowledge about cryptos and e money." (PF2.6)*. Yet another student confirmed the impact of interest on listening to the augmented podcasts, *"If there's interest, it happens (you pay attention). Basically, the whole game is about interest." (PM3.17)*

This points to learner-driven meaning-making, facilitated by the structure of the AP.

5.5. Barriers and limitations to augmented podcast engagement

Despite the pedagogical advantages of the AP format, several learners pointed out barriers that limited their ability to engage fully with the content. These challenges stem from individual learning preferences, processing limitations, and existing habits shaped by prior educational experiences.

5.5.1. Misalignment with individual learning styles

One of the students indicated nonpreference for audio mode of learning, suggesting that audio-based formats may not cater to all cognitive styles. *"I can never learn just by listening." (PF3.1)* Another observed the effect of a new topic. *"If we've studied this material before, then it's okay, but if we've studied a new topic, it's tough to adjust with audio+slide" (PM3.17)*

Learners who tend to be more visually or kinesthetically inclined often struggle to maintain focus or effectively comprehend content delivered through auditory channels. Although the AP aimed to offer a multisensory experience, its dependence on auditory processing may disadvantage those whose primary learning preferences do not align with this mode.

This observation is consistent with existing literature on learning styles, which suggests that individuals vary in how they best receive and process information [36, 37]. Although the validity and applicability of learning styles remain debated, learners' perceived preferences still play a crucial role in motivation and engagement. For example, students who mainly consider themselves visual learners might struggle to maintain focus when key concepts are presented through spoken explanations, even if accompanied by visual aids. One participant remarked that they tended to "zone out" while listening, whereas reading or watching videos helped them stay more alert. *"My mind was a bit scattered, thinking about what people were doing and how they were doing it. I was also wondering who was where in the course but when the bell rang, I had to come back to the course" (PF3.1)*

Therefore, it is essential to consider learner variability when creating educational media, especially formats that engage various sensory demands.

5.5.2. Split attention and cognitive load

The simultaneous requirement to listen to audio narration and read from the slides presented cognitive challenges for some learners. While the dual-channel delivery was engaging and immersive for many, others found the need to process two streams of

information—text and sound—mentally taxing. For such learners, the division of attention sometimes led to missing either the audio or the visual track, affecting comprehension and continuity.

This phenomenon aligns with Cognitive Load Theory [38], which underscores the limitations of working memory when processing multiple inputs. High extraneous load—such as managing unsynchronized audio and visual elements—can hinder integration of information. Mayer's [39] Cognitive Theory of Multimedia Learning similarly stresses that learning is most effective when modalities complement rather than compete. When this balance is disrupted, learners may disengage from one channel to manage cognitive overload.

This was reflected in participant statements such as, *"While reading the slides I was missing on the listening part of the audio." (PF2.6)* And *"You have to read in that (audio + slides). In video, it's like you only have to watch." (PM3.9)*

These highlight the need for better-paced integration or the possibility of giving learners more control over timing and navigation to ease the cognitive load.

5.5.3. Mismatch between reading and listening speeds

One of the concerns voiced by a few participants was the misalignment between the speed of the audio narration and their individual reading pace. While the AP was designed to facilitate dual-channel learning—combining auditory narration with visual slides—several students pointed out that the fixed pace of the audio track did not always correspond to their own cognitive rhythm.

For some learners, the narration felt too slow, leading to impatience or disengagement. Others found the audio progressed too quickly for them to simultaneously read and comprehend the content on the slides. As one student put it, *"The speed of the audio and our reading speed don't match for me." (PM3.9)* And another added, *"I had to quickly change the slide, which was a bit more challenging." (PM2.5)*

This perceptual gap resulted in interruptions in comprehension, requiring learners to pause or re-listen, which potentially disrupted the immersive flow of the session.

This issue resonates with principles from Cognitive Theory of Multimedia Learning [39], particularly the importance of temporal contiguity—presenting words and pictures simultaneously and in synchrony to foster understanding. When narration is not aligned with a learner's internal processing speed, it can generate extraneous cognitive load [38], leading to frustration or surface-level learning. Research in self-regulated learning emphasizes the importance of learner control in digital environments [40]. The inability to adjust playback speed or navigate content freely limited students' capacity to tailor the learning experience to their needs. Such mismatches highlight the value of adaptive features such as adjustable audio speed or the option to toggle between guided and self-paced modes.

In future design iterations, incorporating interactive controls for pacing could enhance usability and learner satisfaction, particularly in diverse classrooms where reading fluency and cognitive processing speeds vary significantly.

5.5.4. Language proficiency as a barrier to comprehension

While only two participants in the study reported language-related challenges, their experiences highlight important considerations regarding linguistic accessibility in multimedia learning environments. These learners expressed difficulty in following English audio narration, noting that spoken English—particularly in academic contexts—can be harder to process in real-time than written text. Few participants shared, *"If they only speak, it doesn't help*

me as much, and my English is a bit weak." (PF3.1). "We can't listen attentively and read at the same time.. . .not everyone is proficient in English" (PF3.16). "It was totally English...somewhere in between there is also Hindi, so our interest remains" (PM3.4)

Interestingly, when probed about preferences for Hindi or regional language options, one participant indicated that, even if the audio were in Hindi, they would still prefer the slides in English. *"I prefer English paragraphs" (PM3.4)*

This underscores a nuanced dynamic: while students from vernacular backgrounds may struggle with spoken English, they often prefer or expect academic content in English script due to its dominance in higher education. Such preferences reflect the complex interplay between language proficiency, educational norms, and learner comfort, a factor echoed in studies on bilingual learners and medium of instruction [41, 42].

Research supports the notion that limited proficiency in the language of instruction can impede comprehension in multimedia learning. Ramsin and Mayall [43] found that second language learners' performance in multimedia environments can be influenced by the modality of content delivery, which have implications for comprehension and retention. Additionally, studies have shown that the integration of captions or transcripts can aid learners with lower language proficiency by providing visual reinforcement of auditory information [44].

Taken together, the affordances and limitations discussed across these themes underscore the nuanced dynamics of learner engagement in digital environments—insights that inform the concluding reflections of this study.

6. Conclusion

The findings of this study provide valuable insights into how learners engage with APs and the factors that influence their effectiveness as a digital learning tool. While APs offer flexibility, accessibility, and a multimodal approach, their success depends on learner preferences, cognitive processing, and instructional design choices. The findings reveal a nuanced picture shaping learners' engagement with this format.

Thematically, the AP was largely seen as an effective tool for fostering enhanced cognitive engagement, deepened understanding, and multisensory learning. Learners appreciated the demand for active listening and the mental effort involved in toggling between audio and visual content. This led to more intentional learning and improved retention for many. Additionally, the format enabled layered content delivery, where different aspects of a concept were explained through distinct but complementary channels- audio + slides, reinforcing comprehension. Built-in cues such as the periodic bell sound were noted as subtle but effective mechanisms to sustain attention and reduce cognitive drift.

The study revealed that mismatches in pacing between reading and listening disrupted learners' engagement. Language proficiency also emerged as a key concern—students from vernacular backgrounds found the academic English audio challenging, though they were more comfortable reading English-scripted slides. While only a few students voiced this issue, it underscores the broader challenge of language inclusivity in a linguistically diverse country. The findings suggest an ingrained familiarity with English text, even among students more fluent in regional languages, raising important considerations for designing accessible digital content.

To enhance inclusivity, future iterations of APs should incorporate features such as adjustable playback speeds, multilingual

subtitles, and transcripts. These affordances would offer learners greater control and support diverse linguistic needs.

In light of these findings, the AP presents a promising format for deep learning but necessitates adjustments, such as customizable playback speeds, multilingual support, optional transcripts, and flexible content modes, to enhance its effectiveness and inclusivity. Ultimately, this study contributes to a deeper understanding of multimodal learning in post-secondary education and provides a learner-informed roadmap for developing equitable and engaging educational technologies.

6.1. Limitations of the study

The study though offers valuable insights into learner engagements with media, especially APs that have certain limitations. The purposive and convenience sampling, though from the spread of tier-2 and tier-3 cities, limits the generalizability of the study. The qualitative data are based on self-reported interviews, which maybe influenced by recall or social desirability. Additionally, the study focused on a limited domain that of commerce, which may not reflect the performance of APs in other domains.

6.2. Implications for digital learning design

Based on these findings, educators and instructional designers should consider the following:

1) Optimize content structure—Use concise, well-organized slides with bullet points to facilitate learning.
2) Leverage multimodal strategies— Combine audio, visuals, and text effectively to enhance retention.
3) Prioritize engagement and interaction—Explore ways to integrate Q&A, storytelling, and real-world examples.
4) Address digital fatigue—Break content into short, structured modules to sustain learner attention.
5) Enhance technical accessibility— Improve slide clarity, navigation, and user experience to minimize distractions.

6.3. Final thoughts and future directions

This study contributes to the growing discourse on digital learning innovations, particularly in bridging accessibility and engagement through APs. While APs present a low-bandwidth, flexible learning alternative, their design must account for cognitive load, learner preferences, and instructional presence. This study was focused on the finance and commerce domain; however, the effectiveness of APs could be different for other domains such as Science, Mathematics, Linguistics, and Humanities . These specific experiments can be conducted to assess the effectiveness of APs in these domains for focused research establishing the connect between the content and the medium (where a specific medium like podcasts is a preferred medium for a certain subject domain). Longitudinal studies for evaluating retention and conceptual understanding across disciplines, regions, and languages could further establish and refine its applicability. Future research could also explore interactive APs, AI-driven personalization, and hybrid models that blend the benefits of both video and audio-based learning.

Given the increasing emphasis on digital equity and inclusion in national policies like NEP2020, APs offer a viable and scalable solution for bandwidth constrained environments. Policymakers and stakeholders should consider including Aps in repositories such as DIKSHA and include these in faculty development programs to promote equitable access and pedagogical innovation in diverse contexts.

Ethical Statement

All ethical considerations have been taken due care of as per guidelines and standards and were ensured by The Institutional Ethical Committee, Indian Institute of Technology, Gandhinagar. All participants of the study provided their written consent before participating in the study.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support this work are available upon reasonable request to the corresponding author.

Author Contribution Statement

Ajita Deshmukh: Conceptualization, Methodology, Validation, Formal analysis, Data curation, Writing – original draft, and Writing – review & editing. **Dhairya Pandya:** Methodology, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization, and Project administration. **Sameer Sahasrabudhe:** Conceptualization, Writing – review & editing, Supervision, Project administration, Funding acquisition.

References

[1] Evans, C. (2008). The effectiveness of m-learning in the form of podcast revision lectures in higher education. *Computers and Education/Computers & Education*, *50*(2), 491–498. https://doi.org/10.1016/j.compedu.2007.09.016

[2] Carson, L., Hontvedt, M., & Lund, A. (2021). Student teacher podcasting: Agency and change. *Learning, Culture and Social Interaction*, *29*, 100514. https://doi.org/10.1016/j.lcsi.2021.100514

[3] Besser, E. D., Blackwell, L. E., & Saenz, M. (2022). Engaging students through educational podcasting: Three stories of implementation. *Technology, knowledge and Learning*, *27*(3), 749–764. https://doi.org/10.1007/s10758-021-09503-8

[4] Forbes, D., & Khoo, E. (2015). Voice over distance: A case of podcasting for learning in online teacher education. *Distance Education*, *36*(3), 335–350. https://doi.org/10.1080/01587919.2015.1084074

[5] Gunderson, J. L., & Cumming, T. M. (2023). Podcasting in higher education as a component of universal design for learning: A systematic review of the literature. *Innovations in Education and Teaching International*, *60*(4), 591–601. https://doi.org/10.1080/14703297.2022.2075430

[6] Abhishek, N., Kulal, A., Divyashree, M. S., & Dinesh, S. (2025). Effectiveness of MOOCs on learning efficiency of students: A perception study. *Journal of Research in Innovative Teaching & Learning*, *18*(1), 145–164. https://doi.org/10.1108/jrit-12-2022-0091

[7] Baba, K., Elfaddouli, N.-E., & Cheimanoff, N. (2021). A comparative study of synchronous and asynchronous learning during COVID-19 Crisis. In *4th International Academic Conference on Education, Teaching and Learning*, 38–49. https://www.dpublication.com/wp-content/uploads/2021/10/10-8080.pdf

[8] Congmeng, J., & Mazlan, A. N. (2025). Exploring educational technology, online communities, and self-paced learning in enhancing graduate students health promotion skills in physical education. *Journal of Pedagogical Research*, *9*(1), 56–72. https://doi.org/10.33902/jpr.202527502

[9] Kearney, A., & Hartnett, M. (2016). Students taking charge of their learning (2016): What are the issues? *The International Journal of Pedagogy and Curriculum*, *23*(3), 35. https://doi.org/10.18848/2327-7963/cgp/v23i03/35-51

[10] Perera, L., & Richardson, P. (2010). Students' use of online academic resources within a course web site and its relationship with their course performance: An exploratory study. *Accounting Education*, *19*(6), 587–600. https://doi.org/10.1080/09639284.2010.529639

[11] Billsberry, J., & Alony, I. (2023). The MOOC post-mortem: Bibliometric and systematic analyses of research on massive open online courses (MOOCs), 2009 to 2022. *Organizational Behavior Teaching Review*, *48*(4), 634–670. https://doi.org/10.1177/10525629231190840

[12] Deng, R. (2023). Effect of video styles on learner engagement in MOOCs. *Technology Pedagogy and Education*, *33*(1), 1–21. https://doi.org/10.1080/1475939x.2023.2246981

[13] Alharbi, A. H. (2023). Investigating the acceptance and use of massive open online courses (MOOCs) for health informatics education. *BMC Medical Education*, *23*(1), 1–10. https://doi.org/10.1186/s12909-023-04648-9

[14] Minooka, L. (2024). The rise of podcasting: Evolution, impact, and future directions. *Global Media Journal*, *22*(72), 473–473. https://doi.org/10.36648/1550-7521.22.72.473

[15] Kulkov, I., Kulkova, J., Rohrbeck, R., & Menvielle, L. (2024). Leveraging podcasts as academic resources: A seven-step methodological guide. *International Journal of Qualitative Methods*, *23*. https://doi.org/10.1177/16094069241266197

[16] Chan-Olmsted, S., & Wang, R. (2020). Understanding podcast users: Consumption motives and behaviors. *New Media & Society*, *24*(3), 684–704. https://doi.org/10.1177/1461444820963776

[17] Tobin, S. J., & Guadagno, R. E. (2022). Why people listen: Motivations and outcomes of podcast listening. *PLoS One*, *17*(4), e0265806. https://doi.org/10.1371/journal.pone.0265806

[18] McCarthy, J., Porada, K., & Treat, R. (2023). Educational podcast impact on student study habits and exam performance. *Family Medicine*, *55*(1), 34–37. https://doi.org/10.22454/fammed.55.183124

[19] Köhler, D., Serth, S., & Meinel, C. (2024). Promoting content variety in MOOCs: Increasing learning outcomes with podcasts. *Frontiers in Education*, *9*, 1–17. https://doi.org/10.3389/feduc.2024.1339142

[20] Oslawski-Lopez, J., & Kordsmeier, G. (2021). "Being able to listen makes me feel more engaged": Best practices for using podcasts as readings. *Teaching Sociology*, *49*(4), 335–347. https://doi.org/10.1177/0092055x211017197

[21] Cain, J. P. (2020). A qualitative study on the effect of podcasting strategies (studycasts) to support 7th grade student motivation and learning outcomes. *Middle School Journal*, *51*(3), 19–25. https://doi.org/10.1080/00940771.2020.1735867

[22] Mutlu-Bayraktar, D., Cosgun, V., & Altan, T. (2019). Cognitive load in multimedia learning environments: A systematic review. *Computers & Education*, *141*, 103618. https://doi.org/10.1016/j.compedu.2019.103618

[23] Sahasrabudhe, S., Deshmukh, A., Pandya, D., & Majumdar, R. (2025). Augmented podcasts: A low bandwidth solution for effective learning. *Indian Journal of Educational Technology*, *7*(1), 289–308. Retrieved from https://journals.ncert.gov.in/IJET/article/view/849

[24] Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*, 149–210. https://doi.org/10.1007/BF01320076

[25] Alty, J. L. (2002). Dual coding theory and education: Some media experiments to examine the effects of different media on learning. In *The Proceedings of EDMEDIA 2002: World Conference on Educational Multimedia, Hypermedia & Telecommunications*, 42–47.

[26] Sweller, J. (2011). Cognitive load theory and E-Learning. In *International Conference on Artificial Intelligence in Education*, 5–6. https://doi.org/10.1007/978-3-642-21869-9_3

[27] Bledsoe, R., & Richardson, D. (2020). View of maximizing learning using cognitive load theory. *ETH Learning and Teaching Journal*, *2*(2), 148–152. https://doi.org/10.16906/lt-eth.v2i2.121

[28] Wong, A., Leahy, W., Marcus, N., & Sweller, J. (2012). Cognitive load theory, the transient information effect and e-learning. *Learning and Instruction*, *22*(6), 449–457. https://doi.org/10.1016/j.learninstruc.2012.05.004

[29] Sithole, S. T. M. (2019). Enhancing blended learning materials using cognitive load theory. *Journal of Modern Accounting and Auditing*, *15*(1), 40–43. https://doi.org/10.17265/1548-6583/2019.01.004

[30] Sujatha, U., & Rajasekaran, V. (2024). Optimising listening skills: Analysing the effectiveness of a blended model with a top-down approach through cognitive load theory. *MethodsX*, *12*, 102630. https://doi.org/10.1016/j.mex.2024.102630

[31] Skulmowski, A., & Xu, K. M. (2021). Understanding cognitive load in digital and online learning: A new perspective on extraneous cognitive load. *Educational Psychology Review*, *34*(1), 171–196. https://doi.org/10.1007/s10648-021-09624-7

[32] Kelly, J. M., Perseghin, A., Dow, A. W., Trivedi, S. P., Rodman, A., & Berk, J. (2022). Learning through listening: A scoping review of podcast use in medical education. *Academic Medicine*, *97*(7), 1079–1085. https://doi.org/10.1097/ACM.0000000000004565

[33] Okonski, R., Toy, S., & Wolpaw, J. (2022). Podcasting as a learning tool in medical education: Prior to and during the pandemic period. *Balkan Medical Journal*, *39*(5), 334–339. https://doi.org/10.4274/balkanmedj.galenos.2022.2022-7-81

[34] Conroy, D., & Kidd, W. (2023). Using podcasts to cultivate learner–teacher rapport in higher education settings. *Innovations in Education and Teaching International*, *60*(6), 861–871. https://doi.org/10.1080/14703297.2022.2102528

[35] Brew-Girard, E., Brown, R., Salter, E., Hattersley, C., Hodge, O., & Leonard, X., ..., & De Souza, S. (2023). Hunting for pearls: A qualitative analysis of the reflections of students

creating psychiatric podcasts. *Advances in Medical Education and Practice*, *14*, 1157–1166. https://doi.org/10.2147/AMEP.S411072

[36] Nurmalisa, Y., Sunyono, S., Yulianti, D., & Sinaga, R. M. (2023). An integrative review: Application of digital learning media to developing learning styles preference. *International Journal of Information and Education Technology*, *13*(1), 187–194. https://doi.org/10.18178/ijiet.2023.13.1.1795

[37] Coffield, F., Moseley, D., Hall, E., & Ecclestone, K. (2004). *Learning styles and pedagogy in post-16 learning*. Retrieved from: http://www.leerbeleving.nl/wp-content/uploads/2011/09/learning-styles.pdf

[38] Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

[39] Mayer, R. E. (2005). Cognitive theory of multimedia learning. In E.R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). Cambridge University Press. https://doi.org/10.1017/cbo9780511816819.004

[40] Bylieva, D., Hong, J., Lobatyuk, V., & Nam, T. (2021). Self-regulation in E-learning environment. *Education Sciences*, *11*(12), 785. https://doi.org/10.3390/educsci11120785

[41] Annamalai, E. (2005). Mother tongue education. In K. Brown (Ed.), *Encyclopedia of language and linguistics* (pp. 342–345). Elsevier.

[42] Mohanty, A. K. (2006). Multilingualism of the unequals and predicaments of education in India: Mother tongue or other tongue? In O. Garcia, T. Skutnabb-Kangas, & M. E. Torres-Guzmán (Eds.), *Imagining multilingual schools* (pp. 262–283). Multilingual Matters. https://doi.org/10.21832/9781853598968-014

[43] Ramsin, A., & Mayall, H. J. (2019). Assessing ESL learners' online learning self-efficacy in Thailand: Are they ready? *Journal of Information Technology Education Research*, *18*, 467–479. https://doi.org/10.28945/4452

[44] BavaHarji, M., Alavi, Z. K., & Letchumanan, K. (2014). Captioned instructional video: Effects on content comprehension, vocabulary acquisition and language proficiency. *English Language Teaching*, *7*(5), 1–16. https://doi.org/10.5539/elt.v7n5p1

Annexure 1: Interview Protocol

Tuning into Learning: Student Narratives from an Augmented Podcasts Intervention

Note: This interview protocol was developed as part of the qualitative phase of a mixed-methods study exploring the learner experiences with Augmented Podcasts (APs). This was used to conduct semi-structured interviews to capture students' perspectives and perceptions and their contextual barriers and enablers in their lived experience of APs. The questions and the probes were designed to encourage open-ended reflections while maintaining consistency across interviews. The interview protocol was reviewed and validated internally.

Informed Consent and Ethical Consideration:

Before the interview, participants were informed about the purpose of the study, their rights as participants, and the voluntary nature of participation. Consent was collected via a Google Form. Participants were assured that their responses would be anonymized and that this data would be solely used for research purposes. Ethical approval from the overseeing institute was ensured.

Interview Protocol

Main Question	Probes / Follow-Up Prompts
1. Which type of e-learning package have you just completed?	- Was this your first experience with an Augmented Podcast? - How did it compare with other formats like videos or plain slides? - Where and how did you access it (e.g., phone, laptop, and Wi-Fi/data)?
2. The e-learning package met the overall learning objective.	- In what ways did it meet the objective? - Did the format help you understand key concepts? - Did the format support your learning pace and style?
3. If the overall learning objective was not covered, please provide further details.	- Which parts of the content felt unclear or insufficient? - Did the mode of delivery (audio + PDF) hinder understanding? - Were there any distractions or any such issues?
4. How would you rate access to the e-learning package?	- Was it easy to download and access the materials? - Did you face any technical or connectivity issues? - Did the file size affect your ability to use it?
5. How would you rate the layout of the e-learning package?	- Were the slides well organised and easy to follow? - Were the visuals and text appropriate and readable? - Did the layout help or hinder your focus?
6. How would you rate the materials in helping you learn?	- Was the combination of audio and slides helpful? - Did you find the pace appropriate for your understanding? - Were you able to revisit or review materials easily?
7. What were the main benefits of the e-learning package which you accessed?	- What features did you find most helpful? - Did it support self-paced learning or multitasking? - How did it affect your engagement compared to other formats?
8. What were the most valuable things that you have learnt?	- Can you share specific takeaways from the content? - Did the format help you apply or retain the concepts? - Was your learning deeper or more reflective?
9. What would you have liked to see more of in the e-learning package?	- Were there any missing features or content? - Would interactive elements or visuals help? - Would multilingual support or transcripts improve your experience?
10. Do you have any other comments or anything else that you would like to share with us?	