Effetti di Lettura
Effects of Reading

Podagogy: turning podcast hype into real learning gains

Podagogia: trasformare l'entusiasmo per i podcast in reali vantaggi per l'apprendimento

Andrea Laudadio
Head of TIM Academy | TIM SpA | andrea@laudadio.it

ABSTRACT

Podcasts have become a popular medium for asynchronous learning, yet the efficacy of this approach remains a subject of debate. This integrative review synthesises 42 primary studies and 12 systematic reviews with the objective of evaluating learning outcomes and knowledge retention. The analysis yielded immediate knowledge gains in specific contexts (e.g., 6–20% in assessments), with Content Acquisition Podcasts (CAPs) and structured language learning demonstrating robust efficacy. However, a multitude of substantial reviews have documented effects that are modest, non-significant, or trivial (d = 0.19). Furthermore, several controlled trials have found no significant differences when compared with traditional instruction. The evidence for short-term retention (i.e. ″ 30 days) is moderate and generally non-inferior; however, confidence in long-term retention (> 30 days) and performance transfer remains low. Theory-driven design makes podcasts an effective supplement, yet they do not consistently surpass traditional methods. There is a clear need for more robust longitudinal studies.

Keywords: podcasts, educational technology, learning outcomes, knowledge retention

I podcast si sono affermati come un diffuso strumento per l'apprendimento asincrono, tuttavia la loro efficacia rimane un tema dibattuto. Questa revisione sistematica della letteratura analizza 42 studi primari e 12 revisioni sistematiche al fine di valutare gli esiti di apprendimento e la ritenzione delle conoscenze. L'analisi ha rivelato guadagni di conoscenza immediati in contesti specifici (es. 6–20% nelle valutazioni), con i Content Acquisition Podcasts (CAPs) e l'apprendimento linguistico strutturato che dimostrano una solida efficacia. Ciononostante, numerose revisioni sistematiche hanno documentato effetti modesti, non significativi o trascurabili (d = 0.19). Diversi studi controllati, inoltre, non hanno rilevato differenze significative rispetto ai metodi di istruzione tradizionali. Le evidenze sulla ritenzione a breve termine (″ 30 giorni) sono moderate e generalmente non inferiori; tuttavia, la fiducia nella ritenzione a lungo termine (> 30 giorni) e nel trasferimento delle competenze è bassa. L'uso di una progettazione basata su teorie dell'apprendimento rende i podcast un efficace strumento supplementare, ma non garantisce un superamento sistematico delle metodologie tradizionali. Emerge con chiarezza la necessità di studi longitudinali più robusti.

Parole chiave: podcast, tecnologia didattica, risultati dell'apprendimento, mantenimento delle conoscenze

OPEN ACCESS Double blind peer review

Volume 4 | n. 2 | dicembre 2025

Citation: Laudadio, A. (2025). Podagogy: turning podcast hype into real learning gains. *Effetti di Lettura / Effects of Reading, 4*(2), 90-104. https://doi.org/10.7347/EdL-02-2025-06.

Corresponding Author: Andrea Laudadio | andrea@laudadio.it

Journal Homepage: https://ojs.pensamultimedia.it/index.php/edl

Pensa MultiMedia ISSN 2785-7050 | DOI: 10.7347/EdL-02-2025-06

Authorship/Attribuzioni:

1. Introduction

The integration of digital media into educational practices has positioned podcasts as a prominent tool for asynchronous learning across various academic disciplines, particularly in medical and higher education settings. Broadly defined, a podcast is a series of digital media files, either audio or video, that a user can download or stream for learning purposes. The pedagogical application of this medium has led to the emergence of the term Podagogy. Coined to describe the intersection of podcasting and pedagogy, this term refers specifically to the theory and practice of using podcasts as an intentional educational tool. It moves beyond simple content delivery to encompass the evidence-based best practices and instructional design principles required to transform a podcast from a passive medium into an active learning instrument designed to achieve specific, measurable learning outcomes.

A critical distinction, often overlooked in general discourse but central to instructional design, lies between audio-only podcasts and video podcasts (vodcasts). This differentiation is paramount because the inclusion of visual elements fundamentally alters the cognitive processing demands placed upon the learner. While audio-only formats offer flexibility for learning during non-traditional activities (e.g., commuting), vodcasts, which combine auditory and visual information, must be carefully designed to avoid redundancy and cognitive overload.

Despite their rapid uptake, prior scoping and integrative reviews have consistently noted the limited availability of rigorous outcome evidence and highlighted mixed findings regarding effectiveness (Pierce, 2022). Podcast effectiveness is rooted in established theoretical frameworks. Multimedia learning theory (Mayer, 2014) posits that well-designed audio and visual combinations can enhance learning. Cognitive load theory (Sweller, 1988) offers a framework for understanding how design characteristics, such as duration, pacing, and information density, influence cognitive processing. Moreover, the Technology Acceptance Model (TAM) provides a framework for understanding the factors that influence adoption by students and staff, such as perceived usefulness and ease of use (Davis, 1989; Merhi, 2015).

Despite the podcast's widespread adoption, the extant evidence regarding its effectiveness is decidedly mixed. Previous, extensive, and exhaustive reviews yielded disconcerting conclusions. In a review of 30 studies, Hew (2009) and in a review of 53 studies, Kay (2012) both found predominantly small, inconsistent, or non-significant effects on learning outcomes. Heilesen (2010) similarly concluded that the empirical evidence for academic benefits remained weak. More recent reviews in specific domains, such as physician education, have expressed similar concerns, noting that many evaluations report only reaction-level outcomes (Kirkpatrick Level 1) and that aggregate methodological quality is often deficient (Cho et al., 2017; Wang et al., 2023).

This review addresses the ambiguity between enthusiastic podcast adoption and the weak, mixed evidence base for their effectiveness. We aim to synthesize the fragmented evidence by systematically examining four critical dimensions: (1) immediate knowledge acquisition and reported effect sizes; (2) delayed knowledge retention and durability; (3) comparative effectiveness versus traditional teaching methods; and (4) implementation moderators and contextual factors that may explain divergent findings.

1.1 Research Questions

This integrative review addresses four primary research questions:

(RQ1) What immediate learning outcomes and knowledge gains do educational podcasts produce across different educational contexts? (RQ2) How durable are podcast-based learning gains, particularly regarding knowledge retention over time? (RQ3) How does podcast effectiveness compare to traditional

teaching methods in controlled studies? (RQ4) What implementation factors and contextual moderators influence podcast effectiveness?

2. Methods

2.1 Study Design and Protocol

The present review follows the five-stage integrative review framework developed by Whittemore and Knafl (2005): problem identification, literature search, data evaluation, data analysis, and presentation. This methodology is particularly appropriate for this topic as it permits the synthesis of diverse study designs – including experimental trials, quasi-experimental studies, and qualitative analyses – whilst maintaining a structured approach to quality evaluation and data synthesis (Whittemore & Knafl, 2005). We completed the literature search in September 2025.

2.2. Search Strategy and Eligibility Criteria

We employed a purposive search strategy, encompassing a comprehensive array of databases to attain breadth. A comprehensive search strategy was employed, incorporating major bibliographic databases such as PubMed/MEDLINE, Scopus, Web of Science, and ERIC (Education Resources Information Center), along with Google Scholar. The search timeframe was set from January 2010 to September 2025 to capture the most relevant modern literature, although highly cited foundational papers from before 2010 (e.g., Abt & Barry, 2007; Evans, 2008; Hew, 2009) were included from the reference lists of synthesized reviews.

Keywords and search terms included variations of:

- Primary terms: podcast*, podcasting, educational podcast*, audio learning, content acquisition podcast*
- Outcome terms: learning outcome*, knowledge retention, knowledge acquisition, academic performance, effectiveness
- Context terms: education*, medical education, higher education, training, language learning

The screening process involved two stages (title/abstract review and full-text assessment) to identify relevant empirical studies. The initial database search yielded 853 records. After removing 111 duplicates, 742 titles and abstracts were screened for relevance. Of these, 97 full-text articles were assessed for eligibility against the inclusion and exclusion criteria. Following full-text review, 43 articles were excluded due to: (a) absence of objective learning outcome measures (n = 21), (b) non-educational contexts (n = 10), (c) conference abstracts without full papers (n = 6), and (d) other reasons including language restrictions (n = 4). This process resulted in the final inclusion of 54 studies: 42 primary empirical studies and 12 systematic or scoping reviews.

Inclusion criteria: Empirical studies (quantitative, qualitative, or mixed-methods) reporting objective learning outcomes; educational or training contexts; interventions involving audio or video podcasts; and published in English. We also included systematic and scoping reviews to synthesise a “review of reviews”.

Exclusion criteria: Studies reporting only learner satisfaction (Kirkpatrick Level 1) without learning outcomes; non-educational contexts; and conference abstracts without full papers.

2.3 Data Extraction, Quality Assessment, and Synthesis

Quality was assessed by synthesising indicators from the included literature rather than conducting an independent risk-of-bias assessment, consistent with integrative review methodology.

Methodological Quality Indicators: We systematically extracted MERSQI scores from included systematic reviews where reported (e.g., Cho et al., 2017; Wang et al., 2023). Outcomes were categorized using the Kirkpatrick framework: Level 1 (Reaction: learner satisfaction), Level 2 (Learning: knowledge acquisition), Level 3 (Behaviour: application of knowledge), and Level 4 (Results: organizational or patient outcomes). Studies were classified by methodology (e.g., RCT, quasi-experimental) to assess study design rigor.

We centred data extraction on the following: the characteristics of the study, the intervention's details, the outcome measures (including effect sizes and statistical significance), and the implementation moderators. A formal meta-analysis was not feasible due to marked heterogeneity in study designs, populations, and outcome measures. Instead, we employed a narrative synthesis to identify, tabulate, and synthesize patterns across the included studies.

3. Results

3.1 Study Overview

This integrative review included 42 primary empirical studies and 12 systematic/scoping reviews published between 2007 and 2025.

Distribution by Study Design (Primary Studies):
- Randomised Controlled Trials (RCTs): 26% (11 studies)
- Quasi-experimental studies: 33% (14 studies)
- Observational/cohort studies: 31% (13 studies)
- Qualitative studies: 10% (4 studies)

Distribution by Educational Context (Primary Studies):
- Medical/health education: 45% (19 studies)
- Teacher education: 19% (8 studies)
- Higher education (non-medical): 17% (7 studies)
- Language learning: 12% (5 studies)
- K-12 education: 7% (3 studies)

Methodological Quality: Mean MERSQI scores extracted from included medical education reviews were 10.9 (range: 7.5–14.0), indicating moderate methodological quality in that specific field (Cho et al., 2017; Wang et al., 2023). We observed a predominant focus on Kirkpatrick Level 1 (Satisfaction) and Level 2 (Learning) outcomes across the entire body of literature.

3.2. Immediate Knowledge Gains and Learning Outcomes (RQ1)

The evidence for immediate knowledge gains is decidedly mixed, revealing a clear divergence between the modest effects seen in large-scale reviews and the strong effects found in studies of specific, theory-driven interventions.

3.2.1. Evidence from Comprehensive Reviews

Large systematic reviews provide a sobering context for podcast effectiveness claims.

Hew (2009): A comprehensive review of 30 studies revealed a pronounced emphasis (93.3%) on higher education and a general absence of rigorous comparative designs. Within this review, a single comparative study by Abt & Barry (2007) found a negligible effect size (d = 0.19) for podcasts compared to written materials.

Kay (2012): A comprehensive analysis of 53 empirical studies on video podcasts revealed that while students reported high satisfaction, only 7 of the 53 studies reported statistically significant higher test scores with podcasts. This finding suggests that the evidence for superior learning outcomes is limited and inconsistent.

Heilesen (2010): This systematic examination of podcast efficacy research from 2004–2009 concluded that, despite examples of modest gains, the overall empirical evidence for academic benefits remained weak.

3.2.2. Evidence from Controlled Trials and Quasi-Experimental Studies

In contrast to the often inconsistent and modest findings of large-scale reviews, numerous individual controlled studies have reported significant short-term learning gains, particularly when podcasts were integrated into well-defined instructional designs.

Research on Content Acquisition Podcasts (CAPs). CAPs are characterized by a structured, theory-driven format, incorporating multimedia support and explicit instruction, resulting in brief podcasts. A substantial body of research has demonstrated the efficacy of CAPs in facilitating learning, with findings consistently indicating the presence of robust learning effects across diverse studies. In the field of teacher education, for instance, a study encompassing 164 participants revealed that students utilizing CAPs exhibited marked enhancements in performance when compared to those receiving text-based instruction. The findings showed a 20% improvement and moderate to large effect sizes (² = 0.11–0.26) (Kennedy & Thomas, 2012). Conversely, in a separate study involving K–12 students with learning disabilities (n = 279), an enhanced CAP design that integrated explicit instruction with a keyword mnemonic strategy resulted in elevated vocabulary post-test scores compared to all other groups (Kennedy, 2020). CAPs have also been demonstrated to exert a direct influence on teaching practices. Pre-service teachers who utilize this format have been observed to augment their use of behavior-specific praise, exhibiting substantial effect sizes (Tau-U = 0.58–1.0) (Miller & Uphold, 2021).

The domain of language learning offers further evidence of these effects. A study involving Iranian EFL learners (n = 70) revealed that podcast-integrated speaking instruction resulted in a 16.9% improvement in speaking skills (² = 0.23) and an 18.3% increase in academic engagement (² = 0.18). These results indicate that podcast-integrated instruction outperforms traditional instruction (Hosseini & Ghaemi, 2025). The findings from prior reviews have indicated that structured podcast activities can significantly enhance listening comprehension and learner motivation (Hasan & Hoon, 2013; Indahsari, 2020).

Positive outcomes have also been observed in the fields of medicine and higher education. In the context

of surgical education, a significant positive correlation was observed between podcast-based learning and increased knowledge assessment scores, with a mean increase of 20.2 percentage points ($p < 0.001$). O'Neill et al. (2023) reported this finding. A more extensive study ($n = 856$) on video podcasts in a pre-calculus course identified small to moderate effect sizes ($d = 0.27–0.47$), thereby substantiating the efficacy of this medium in facilitating content acquisition (Kay, 2014).

However, these robust outcomes are not universally applicable. A substantial body of well-designed studies has reported null findings, thereby reinforcing the conclusions of large reviews. For instance, a randomized controlled trial in teacher education (O'Bannon et al., 2011) found no significant difference in student achievement between podcast and lecture groups. In the field of exercise physiology, a study was conducted that compared the effects of podcasts and printed transcripts. The study revealed only trivial, non-significant effects ($d = 0.19$) (Abt & Barry, 2007). Other research has also found that podcasts have a positive effect on student satisfaction (see Baker et al., 2007; Lakhal et al., 2007). However, these findings have not been replicated in terms of improved grades or course performance.

Table 1. summarizes key studies examining immediate learning outcomes, highlighting the variability in effect sizes and contexts

Author(s) & Year	**Context**	**Intervention Type**	**Main Outcome**	**Effect Size**
Abt & Barry (2007)	Higher Education (Exercise Physiology)	Podcast vs. Written Transcripts	Negligible effect	$d = 0.19$
Kennedy & Thomas (2012)	Teacher Education	Content Acquisition Podcasts (CAPs)	20% improvement in performance	$\eta^2 = 0.11–0.26$
Hosseini & Ghaemi (2025)	Language Learning (EFL)	Integrated podcast instruction	16.9% improvement in speaking skills	$\eta^2 = 0.23$
O'Neill et al. (2023)	Medical Education (Surgical)	Podcast-based learning	Mean increase of 20.2 percentage points	$p < 0.001$
Kay (2014)	Higher Education (Precalculus)	Video Podcasts	Content acquisition	$d = 0.27–0.47$
O'Bannon et al. (2011)	Teacher Education	Lecture replacement with podcasts	No significant difference	Not reported

Collectively, the evidence for immediate learning gains is moderate. Numerous controlled studies have yielded equivocal results. Large-scale reviews (Hew, 2009; Kay, 2012; Heilesen, 2010) have documented modest or null effects in aggregate, while strong positive effects appear to be concentrated in specific, theory-driven instructional contexts – particularly CAPs and structured language learning interventions. Concurrently, numerous rigorous randomized controlled trials have demonstrated an absence of a quantifiable benefit over conventional instructional formats (O'Bannon et al., 2011), emphasizing the pivotal role of context, design, and implementation fidelity in determining podcast efficacy.

3.3 Knowledge Retention and Long-Term Learning (RQ2)

Evidence regarding retention is predominantly concentrated in the short term (″ 30 days), while data concerning long-term knowledge durability remains limited and inconsistent.

In the short term, the results are generally positive. In the context of teacher education, the CAPs intervention demonstrated that learners in the podcast group exhibited a significant performance advantage

over those in the text-based instruction group at a two-week maintenance probe (Kennedy & Thomas, 2012). Concurrently, a series of randomized controlled trials within the domain of medical education have substantiated the non-inferiority of podcasts with respect to 30-day retention. A randomized controlled trial (RCT) involving 129 participants investigated the comparative effectiveness of podcasts and blog posts in emergency medicine training, but found no significant difference in knowledge retention at 30 days ($p = 0.43$) (Lien et al., 2018). A subsequent study demonstrated that individuals who engaged in podcast listening while exercising, as opposed to in a sedentary setting, exhibited nearly identical retention rates after 30 days (78% vs. 81%) (Gottlieb et al., 2023).

In the context of longer-term retention periods, defined as those exceeding 30 days, the extant evidence base becomes increasingly sparse and fragmented. A preliminary investigation employing electroencephalography (EEG)-based attention metrics ($n = 24$) reported the retention of learning benefits after a 60-day period (Wolpaw et al., 2022). In contrast, a study in physical education revealed that a CAPs group, despite exhibiting robust immediate learning gains, was unable to retain these advantages in a 6-week retention test. Conversely, an open-access podcast group demonstrated the capacity to sustain its performance over time (McNamara et al., 2023).

The certainty of the evidence regarding the retention period can be considered moderate over the 2-week to 30-day timeframe and low for extended follow-up periods. While several well-designed randomized controlled trials (RCTs) support the short-term non-inferiority of podcasts compared with other instructional modalities, the limited number of studies and their mixed results for longer time frames prevent strong conclusions about the durability of podcast-based learning over time.

3.4 Comparative Effectiveness vs. Traditional Methods (RQ3)

A comparative analysis of podcasts and traditional instructional methods has revealed that the evidence generally points to non-inferiority, with superiority manifesting only in specific, well-defined contexts.

Numerous studies and systematic reviews corroborate this phenomenon. A substantial systematic review in the field of medical education has revealed that podcasts are comparable to lectures in terms of effectiveness. A pooled effect size of 0.12 (95% CI: –0.08 to 0.32) supports this finding (Wang et al., 2023). A randomized controlled trial by O'Bannon et al. (2011) found no significant differences in student achievement between podcast and lecture groups. A review of patient education research reached similar conclusions: among 21 studies, only 7 reported improvements in knowledge retention, while the majority showed no measurable difference between podcast-based and traditional approaches (Kakhki et al., 2025).

Superiority effects are often observed in specific instructional designs and targeted applications. For instance, Computer-Assisted Instruction (CAI) has repeatedly demonstrated clear advantages over traditional text-based instruction (Kennedy & Thomas, 2012). In the context of language learning, the incorporation of podcasts into structured learning methodologies has been shown to result in enhanced gains in speaking proficiency and learner engagement when compared to conventional approaches (Hosseini & Ghaemi, 2025). Evidence also suggests that podcasts may be more effective when used as supplemental learning tools rather than as fully integrated replacements. A substantial investigation ($n = 337$) revealed that students who utilized podcasts as supplementary materials attained notably higher final grades compared to those enrolled in integrated podcast curricula (Abdous et al., 2009). In a similar vein, studies have demonstrated that AI-generated, personalized podcasts have been shown to yield superior learning outcomes in comparison to generic content ($d = 0.6$) (Do et al., 2024).

Collectively, the comparative evidence has a moderate degree of certainty. A substantial body of research, including multiple systematic reviews and randomized trials (Hew, 2009; Kay, 2012; O'Bannon et al., 2011; Wang et al., 2023), has consistently demonstrated that podcast-based instruction is generally com-

parable to traditional teaching methods. However, podcasts have been found to be more effective in theory-driven interventions (such as CAPs and structured language learning) and specific implementation strategies, particularly when they are utilized to enhance existing instruction rather than to substitute for it.

3.5 Implementation Factors and Contextual Moderators (RQ4)

The substantial variability in learning outcomes indicates that podcast effectiveness depends heavily on their design and implementation, not just the medium itself.

Content Acquisition Podcasts (CAPs) provide one of the clearest illustrations of this, with their success intimately tied to their foundation in cognitive load theory. CAPs are intentionally concise, generally ranging from 2 to 10 minutes in duration. A particular study identified an optimal duration of 7 minutes and 29 seconds (Kennedy & Thomas, 2012). These materials align with established principles of multimedia learning, integrating explicit instruction with audio narration and still images to minimize extraneous cognitive load. Pairing explicit instruction with a keyword mnemonic strategy produced the most significant effects, resulting in substantial learning gains (Kennedy, 2020).

A similar pattern emerges in the context of structured language learning, where the efficacy of podcasts stems not from passive listening but from a deliberate, three-phase structure: pre-listening preparation, structured note-taking during listening, and active speaking practice afterward (Hosseini & Ghaemi, 2025).

Beyond the structural elements of content, the production quality, and design choices, other factors must be considered. Podcasts that employ prosodic signaling – such as strategic emphasis and pacing – yield superior learning outcomes in comparison to monotone delivery (d = 0.45) (Désiron & Schneider, 2024). Learners have been found to express a consistent preference for podcasts that are enriched with visual aids, such as still images or simple graphics (Cho et al., 2017; Kay, 2012). These findings are indicative of an improved comprehension level among learners when exposed to such podcasts. The duration of podcasts is also a salient factor in this regard. While CAPs are brief by design, optimal podcast length more broadly appears to fall between 7 and 25 minutes (Cho et al., 2017). Kay (2014) discovered that video podcasts with a duration of less than 10 minutes exhibited higher rates of completion and greater perceived usefulness in comparison to those that were longer.

The implementation strategy has also proven to be of equal importance. The extant research indicates that the implementation of podcasts as a supplement to conventional instructional methods tends to yield superior outcomes in comparison to their utilization as a substitute. Abdous et al. (2009) discovered that the implementation of supplemental podcasts resulted in superior learning outcomes when compared with the substitution of integrated curriculum. This finding aligns with the conclusions of O'Bannon et al. (2011), who reported that the substitution of lectures with podcasts in an RCT resulted in no significant learning effects and a substantial decline in engagement. Specifically, student listening decreased from 67% at the commencement of the semester to 30% by its conclusion. These findings underscore the potential repercussions of imposing a replacement model, which may result in the erosion of both engagement and efficacy.

Another salient factor pertains to technology adoption, a process that is not automatic. A substantial study (n = 352) determined that students' adoption of podcasts was significantly predicted by their perceived usefulness (= 0.57) and compatibility with existing study practices (Merhi, 2015).

4. Discussion

4.1 Integrative Synthesis

This review synthesizes evidence from 42 primary studies and 12 systematic or scoping reviews, highlighting a clear tension in the existing literature. On the one hand, targeted, well-designed interventions – such as Content Acquisition Podcasts (CAPs) and structured language learning programs – have been shown to result in robust and statistically significant learning gains. Conversely, a substantial corpus of research, encompassing numerous large-scale reviews and meticulously executed randomized controlled trials, suggests that podcasts elicit modest, null, or non-inferior effects when juxtaposed with conventional teaching methods.

A review of the extant literature reveals several convergent findings. Research shows that immediate short-term knowledge gains ranging from 6% to 30% are consistently achievable. Short-term retention (up to 30 days) is non-inferior to traditional instructional formats. Learner satisfaction is consistently high, indicating the medium's accessibility and perceived usefulness. Furthermore, theory-grounded interventions – such as CAPs or structured language integration – outperform generic podcast use, and supplemental use yields stronger outcomes than mandatory replacement of existing instructional approaches.

However, critical divergent findings counterbalance these results. Large-scale reviews encompassing 30 to 53 studies each have reported predominantly modest or null effects (Hew, 2009; Kay, 2012; Heilesen, 2010). Several well-designed RCTs (e.g., O'Bannon et al., 2011; Abt & Barry, 2007) have also found no substantial discrepancies between podcast-based instruction and conventional teaching methods. Learning engagement can also decline precipitously when podcasts replace lectures rather than serving as a supplemental resource. For instance, O'Bannon et al. (2011) documented a significant decrease in engagement, from 67% to 30% over a semester, when podcasts replaced lectures. Finally, the findings show that retention effects beyond 30 days remain inconsistent and are only marginally supported (McNamara et al., 2023).

The educational impact of podcasts is not inherent to the medium itself; rather, it is shaped by instructional design, implementation strategy, and pedagogical context. This discrepancy stems from design differences: targeted, theory-driven applications yield significant impacts, whereas generic implementations show modest or nonexistent effects. This phenomenon, which has been documented in numerous studies, is a significant source of variability in the extant literature.

4.2 Interpretation of Main Findings

4.2.1 Immediate Learning Effectiveness

The central finding of this review is that podcasts are not a monolithic instructional tool. The substantial variability in outcomes (ranging from $d = 0.19$ to $\eta^2 = 0.26$) is most adequately explained by discrepancies in design and implementation. Generic "lecture-capture" podcasts, as reviewed by Hew (2009) and tested by O'Bannon et al. (2011), appear to offer little or no advantage over text or live lectures. In contrast, podcasts that are explicitly designed according to cognitive load theory and multimedia learning principles (i.e., CAPs) or those integrated into a structured pedagogical sequence (e.g., language learning) have been demonstrated to have robust positive effects. A considerable portion of the enthusiasm associated with podcasts appears to be contingent upon the potential of the medium. However, their practical efficacy is predominantly influenced by instructional design.

4.2.2 Knowledge Retention

The phenomenon of knowledge retention is of particular interest in this study. The extant evidence supporting non-inferior 2- to 30-day retention is encouraging. The findings of this study suggest that podcasts are a viable alternative for short-term learning, providing flexibility without compromising immediate recall. Nevertheless, the absence of compelling evidence beyond 30 days, in conjunction with equivocal results (McNamara et al., 2023), hinders the ability to draw definitive conclusions regarding their involvement in long-term, enduring learning.

4.3 Theoretical Foundations of Podcast Effectiveness

4.3.1 Cognitive Load and Multimedia Learning

The superior performance of CAPs (Kennedy, 2020; Kennedy & Thomas, 2012) and the positive impact of prosodic signaling (Désiron & Schneider, 2024) provide strong empirical support for cognitive load theory and multimedia learning principles. The failure of generic lecture-capture podcasts to achieve success often stems from their tendency to impose high cognitive load, characterized by extended duration and the absence of visual anchors. Conversely, the efficacy of CAPs is attributed to their adept management of cognitive load through brevity, segmentation, and the integration of audio with visual elements such as still images. The finding that video podcasts are not universally more effective than audio alone (Kay, 2012) aligns with the prevailing theory, which posits that redundant or irrelevant visuals (e.g., a "talking head") may increase extraneous cognitive load and impede learning rather than enhance it.

4.3.2 The Implementation and Self-Determination Theory

The observation that supplemental podcast use (high autonomy) was more effective than mandatory replacement (low autonomy) (Abdous et al., 2009) represents a key theoretical contribution. This result, in conjunction with the documented decline in engagement as reported in O'Bannon et al.'s (2011) replacement-model RCT, provides substantial support for the principles of self-determination theory. When podcasts are offered as optional resources, they enhance learner autonomy, which in turn fosters motivation and sustained engagement. Conversely, when podcasts are mandated as replacements, there is a risk that they will be perceived as an additional obligation, thereby undermining the flexibility that is pedagogically valuable in the first place.

4.3.3 Technology Acceptance Model (TAM).

Merhi's (2015) findings provide clear validation of the Technology Acceptance Model in the podcasting context. The strongest predictors of student adoption were found to be perceived usefulness and compatibility with existing study habits. The implication is straightforward: if learners do not perceive a direct benefit for their academic goals or find that podcast use disrupts their established learning routines, they are unlikely to engage with the resource–regardless of its objective quality.

4.4 Practical Implications for Educators and Institutions

Based on this synthesis, we offer a set of evidence-based recommendations for both educators and institutions to optimize the design, implementation, and impact of educational podcasts.

For Educators (Design and Implementation):

1. Prioritize Supplemental Use. Empirical evidence strongly supports using podcasts as optional, supplemental resources for review and reinforcement rather than as mandatory replacements for lectures or core readings (Abdous et al., 2009; O'Bannon et al., 2011). Supplemental use sustains engagement and enhances learner autonomy, while compulsory replacement may undermine both.
2. Design with Cognitive Load Principles in Mind. Podcast effectiveness is tied to how well they are designed, not simply to their availability. Keep them short, aiming for focused episodes of 7–15 minutes that address clearly defined learning objectives (Kennedy & Thomas, 2012; Kay, 2014). Use visual anchors by integrating simple, relevant still images or graphics to reduce cognitive load and enhance understanding, particularly for complex topics (Cho et al., 2017). Signal key concepts by incorporating prosodic cues (e.g., emphasis, pacing) and explicit verbal signposting (e.g., "The most important point is…") to direct learner attention effectively (Désiron & Schneider, 2024).
3. Integrate, Don't Just Add. In contexts such as language learning, passive listening is insufficient. Build structured pedagogical activities around the podcast, including pre-listening preparation, guided listening, and post-listening practice (Hosseini & Ghaemi, 2025).
4. Align with Assessment. Link podcast content explicitly to exam questions and course learning objectives. Perceived usefulness is the strongest predictor of student adoption and engagement (Merhi, 2015).

For Institutions (Policy and Support):

5. Invest in Pedagogical Design, Not Just Technology. Providing recording equipment alone is not enough. Institutions should offer faculty training and instructional design support grounded in cognitive load theory and multimedia learning principles to ensure pedagogically sound podcast production.
6. Ensure Accessibility. Transcripts should be available for all educational podcasts to support learners who are hearing impaired, non-native speakers, or those who prefer searchable text formats for review and note-taking.
7. Monitor Engagement and Outcomes. Podcast popularity is not a proxy for effectiveness. Institutions should track meaningful analytics–such as completion rates, not just downloads–and link them to learning outcomes. Continuous evaluation allows for iterative improvement and, when necessary, the discontinuation of ineffective implementations.

4.5 Strengths and Limitations of the Evidence

The extant body of literature is fortified by the incorporation of diverse study designs, encompassing numerous well-controlled randomized controlled trials (Lien et al., 2018; Gottlieb et al., 2023; O'Bannon et al., 2011) and multiple large-scale systematic reviews (Hew, 2009; Kay, 2012). The development of theoretically grounded frameworks, most notably Content Acquisition Podcasts (CAPs), and the cross-cultural validation of podcast-based interventions (e.g., in language learning) offer promising and replicable models for educational practice. The integration of these elements fosters a more mature and structured

research landscape, representing a progression from the early descriptive studies that characterized the field in its earlier stages.

Notwithstanding these strengths, the extant literature exhibits several critical weaknesses. Methodological quality remains inconsistent, with substantial variability in study designs and outcome measures. The field remains predominantly characterized by Kirkpatrick Level 1–2 evaluations, with a paucity of studies that assess higher-level learning or behavioral outcomes. Short retention windows, which typically last no more than 30 days, represent a significant gap in the research, as they leave unresolved questions regarding long-term learning effects. The heterogeneity of interventions is a significant complicating factor in cross-study comparisons. It is imperative to acknowledge that null findings from extensive reviews (Hew, 2009) and meticulously designed RCTs (O'Bannon et al., 2011) should be accorded equal interpretive weight to positive results. This is crucial to dispel any simplistic narrative that portrays podcasts as universally effective.

It must be noted that this integrative review is subject to several limitations. The search strategy employed was purposive rather than exhaustive, the synthesis relied on a narrative approach rather than a quantitative one, and the inclusion criteria were limited to publications in English. These factors may have constrained the extent and generalizability of the findings.

4.6 Research Priorities and Future Directions

In order to fortify causal inference and address the present tensions in the literature, future research should pursue a more rigorous and targeted agenda.

1. Conduct Long-Term Retention Studies. To address this knowledge gap, there is an urgent need for adequately powered randomized controlled trials with retention intervals extending beyond 30 days–ideally at 3, 6, and 12 months–to determine the longevity of podcast-based learning effects over time.
2. The measurement of higher-level outcomes is imperative. Future studies should move beyond short-term knowledge assessments and examine outcomes at Kirkpatrick Levels 3 and 4, including behavioral change, performance improvement, and patient or client outcomes.
3. An investigation into engagement dynamics is warranted. The engagement decline observed by O'-Bannon et al. (2011) underscores a critical gap in understanding how learners interact with podcast-based materials over time. It is imperative that research explore the psychological, instructional, and contextual factors that influence sustained engagement.
4. Optimization of Design Through Factorial Studies. The necessity of factorial design studies is evident in the deconstruction of complex interventions, such as CAPs, to identify the specific contributions of key design features to learning outcomes. These features include, but are not limited to, duration, mnemonic strategies, and visual supports.
5. It is imperative to broaden the scope of research contexts. The extant evidence is scant in K–12, vocational, and corporate training settings. Subsequent research endeavors must systematically assess the efficacy of podcasts in these under-explored contexts, which are imperative for ensuring real-world applicability.
6. Conduct Replication Studies. The creation of high-quality replications of CAP and structured language learning interventions is imperative to ascertain their robustness and generalizability across disparate educational settings and learner populations.

5. Conclusions

Based on moderate-certainty evidence, podcasts can produce immediate knowledge gains of approximately 6–30% and demonstrate non-inferior short-term retention (″ 30 days) when used in specific, well-designed instructional contexts. This positive evidence base, however, must be interpreted cautiously in light of several critical counter-findings.

First, multiple large-scale systematic reviews examining 30–53 studies each (Hew, 2009; Kay, 2012; Heilesen, 2010) consistently reported modest, inconsistent, or null effects on learning outcomes. Second, well-designed randomized controlled trials have yielded mixed results, with several finding no significant differences between podcast-based and traditional instruction (O'Bannon et al., 2011; Abt & Barry, 2007). Third, when podcasts are deployed as a mandatory replacement for lectures, student engagement can decline sharply over time (O'Bannon et al., 2011).

The evidence is clear that podcasts are not universally superior to traditional methods. Their effectiveness depends heavily on instructional design quality and implementation strategy. The most consistent and substantial positive effects appear in two specific domains:

1. Theory-Driven Design. Content Acquisition Podcasts (CAPs), explicitly designed using cognitive learning principles, have demonstrated robust and replicable effects (Kennedy, 2020; Kennedy & Thomas, 2012).
2. Structured Integration. Podcast-based language learning interventions that employ structured three-phase pedagogical frameworks–pre-listening, guided listening, and post-listening practice–have shown significantly better outcomes than generic listening approaches (Hosseini & Ghaemi, 2025).

By contrast, evidence for long-term retention (beyond 30 days) and transfer of learning to real-world performance remains low in both quantity and quality. Methodological heterogeneity, variable effect sizes, and a significant number of null-effect studies underline the need for cautious interpretation and targeted application.

For educators, the implication is strategic: podcasts should be adopted as evidence-informed supplements, not as wholesale replacements for existing instructional modalities. The strongest results are achieved when podcasts are short, multimedia-enhanced, and explicitly aligned with cognitive learning principles, rather than when they simply replicate classroom lectures.

This leads to what can be termed the "podcast paradox": despite modest aggregate effects, podcasts have been widely adopted. This paradox reflects a fundamental reality of educational technology–tools do not teach; pedagogies do. The stark divergence between generic lecture-capture failures and CAP-driven successes highlights that podcasts operate not as autonomous instructional agents but as amplifiers of pedagogical quality. Poor design, when podcasted, remains poor; excellent design, when mediated through audio, can transcend the temporal and spatial boundaries of the classroom.

The crucial question is therefore not "Do podcasts work?" but "Under what pedagogical conditions, for which learners, and for what kinds of knowledge do audio-mediated learning experiences foster genuine understanding rather than merely transmit information?"

Future research must move beyond outcome equivalence studies to investigate the unique affordances of the audio medium–its potential for intimate, reflective, mobile, and emotionally resonant learning experiences. This shift would allow researchers and practitioners alike to better understand and harness what podcasts can do that other instructional modalities cannot.

References

Abdous, M., Camarena, M. M., & Facer, B. R. (2009). MALL technology: Use of academic podcasting in the foreign language classroom. *ReCALL*, *21*(1), 76–95. https://doi.org/10.1017/S0958344009000020

Abt, G., & Barry, T. (2007). The quantitative effect of students using podcasts in a first year undergraduate exercise physiology module. *Bioscience Education*, *10*(1). https://doi.org/10.3108/beej.10.1

Baker, J. W., Green, S., & Friedman-Wheeler, D. (2007). Podcasting for continuing education: A case study. *Educational Technology Research and Development*, *55*(5), 453–470. https://doi.org/10.1007/s11423-007-9047-y

Cho, D., Cosimini, M., & Espinoza, J. (2017). Podcasting in medical education: A review of the literature. *Korean Journal of Medical Education*, *29*(4), 229–239. https://doi.org/10.3946/kjme.2017.69

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340. https://doi.org/10.2307/249008

Désiron, J. C., & Schneider, S. (2024). Listen closely: Prosodic signals in podcast support learning. *Computers & Education*, *218*, 105051. https://doi.org/10.1016/j.compedu.2024.105051

Do, T. D., Shafqat, U. B., Ling, E., & Sarda, N. (2024). PAIGE: Examining learning outcomes and experiences with personalized AI-generated educational podcasts. arXiv preprint. https://doi.org/10.48550/arxiv.2409.04645

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

Gottlieb, M., Cooney, R., Haas, M. R. C., King, A., Fung, C.-C., & Riddell, J. (2023). A randomized trial assessing the effect of exercise on residents' podcast knowledge acquisition and retention. *Academic Medicine*, *99*(3), 312–318. https://doi.org/10.1097/acm.0000000000005592

Gottlieb, M., Fung, C.-C., Haas, M. R. C., Cooney, R., King, A., & Riddell, J. (2024). Educational podcasts: Effect of content delivery timing on knowledge acquisition and retention. *Academic Medicine*, *99*(7), 789–795. https://doi.org/10.1097/acm.0000000000005722

Hasan, M. M., & Hoon, T. B. (2013). Podcast applications in language learning: A review of recent studies. *English Language Teaching*, *6*(2), 128–135. https://doi.org/10.5539/elt.v6n2p128

Heilesen, S. (2010). What is the academic efficacy of podcasting? *Computers & Education*, *55*(3), 1063–1068. https://doi.org/10.1016/j.compedu.2010.05.002

Hew, K. F. (2009). Use of audio podcast in K-12 and higher education: A review of research topics and methodologies. *Educational Technology Research and Development*, *57*(3), 333–357. https://doi.org/10.1007/s11423-008-9108-3

Hosseini, S. E., & Ghaemi, H. (2025). The impact of podcast-integrated speaking instruction on EFL learners' informal digital learning of English, academic engagement, and speaking skills. *Acta Psychologica*, *251*, 104718. https://doi.org/10.1016/j.actpsy.2024.104718

Indahsari, D. (2020). Using podcast for EFL students in language learning. *JELE (Journal of English Language and Education)*, *6*(2), 97–106. https://doi.org/10.26486/jele.v6i2.1403

Kakhki, S. T., Ebadi, A., & Vanaki, Z. (2025). The effect of podcast-based education on health outcomes: A systematic review. *Journal of Education and Health Promotion*, *14*(1), 13. https://doi.org/10.4103/jehp.jehp_1089_23

Kay, R. H. (2012). Exploring the use of video podcasts in education: A comprehensive review of the literature. *Computers in Human Behavior*, *28*(3), 820–831. https://doi.org/10.1016/j.chb.2012.01.011

Kay, R. H. (2014). Developing a framework for creating effective instructional video podcasts. *International Journal of Emerging Technologies in Learning*, *9*(1), 22–30. https://doi.org/10.3991/ijet.v9i1.3335

Kennedy, M. J. (2020). *Content acquisition podcasts to support secondary students with and without learning disabilities in acquiring academic vocabulary* [Doctoral dissertation, University of Kansas]. KU ScholarWorks. https://kuscholarworks.ku.edu/handle/1808/30934

Kennedy, M. J., & Thomas, C. N. (2012). Effects of content acquisition podcasts to develop preservice teachers' knowledge of positive behavioral interventions and supports. *Exceptionality*, *20*(1), 1–19. https://doi.org/10.1080/09362835.2012.640901

Lakhal, S., Khechine, H., & Pascot, D. (2007). Student behavioural intentions to use desktop video conferencing in a distance course: Integration of autonomy to the UTAUT model. *Journal of Computing in Higher Education*, *19*(1), 93–121. https://doi.org/10.1007/BF03033421

Lien, K., Chin, A., Helman, A., & Chan, T. M. (2018). A randomized comparative trial of the knowledge retention and usage conditions in undergraduate medical students using podcasts and blog posts. *Cureus*, *10*(1), e2065. https://doi.org/10.7759/CUREUS.2065

Mayer, R. E. (2014). *The Cambridge handbook of multimedia learning* (2nd ed.). Cambridge University Press.

McNamara, J. M., Sidor, A., & Biber, D. (2023). Content acquisition podcasts vs. open access podcasts: Improving preservice physical educators' ability to teach students with visual impairments. *Physical Education and Sport Pedagogy*, *28*(4), 423–437. https://doi.org/10.1080/17408989.2021.1990249

Merhi, M. I. (2015). Factors influencing higher education students to adopt podcast: An empirical study. *Computers & Education*, *83*, 32–43. https://doi.org/10.1016/j.compedu.2014.12.014

Miller, B., & Uphold, N. (2021). Using content acquisition podcasts to improve preservice teacher use of behavior-specific praise. *Teacher Education and Special Education*, *44*(4), 300–318. https://doi.org/10.1177/-08884064211007134

O'Bannon, B. W., Lubke, J. K., Beard, J. L., & Britt, V. G. (2011). Using podcasts to replace lecture: Effects on student achievement. *Computers & Education*, *57*(3), 1885–1892. https://doi.org/10.1016/j.compedu.20-11.04.001

O'Neill, S., Deegan, J., Ramjit, S., Reynolds, J. V., & Donohoe, C. L. (2023). 'SurgTalk': The educational outcomes associated with development of a surgical podcast for undergraduate medical students. *Journal of Surgical Education*, *81*(2), 298–305. https://doi.org/10.1016/j.jsurg.2023.10.008

Pierce, N. E. (2022). Learning through listening: A scoping review of podcast use in medical education. *Academic Medicine*, *97*(9), 1289–1297. https://doi.org/10.1097/acm.0000000000004565

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285.

Wang, S., Meckling, G., Hanafi, S., Kasbia, G. S., Singhal, N., Zhou, C., & Raben, A. (2023). *The effects of podcasts on medical education: A systematic review*. Preprint. https://doi.org/10.21203/rs.3.rs-3297034/v1

Whittemore, R., & Knafl, K. (2005). The integrative review: Updated methodology. *Journal of Advanced Nursing*, *52*(5), 546–553. https://doi.org/10.1111/j.1365-2648.2005.03621.x

Wolpaw, J., Ozsoy, S., Berenholtz, S., Wright, S., Bowen, K., Gogula, S., Lee, S., & Toy S. (2022). A multimodal evaluation of podcast learning, retention, and electroencephalographically measured attention in medical trainees. *Cureus*, *14*(11), e31289. https://doi.org/10.7759/cureus.31289