



Research Article

Podcast-Based Learning and Large Language Models: An Experimental Study on English Learning Motivation

Feyza Nur EKİZER*

*Necmettin Erbakan University, Konya/Turkey

ARTICLE INFO

Keywords:

Podcast-based learning
Large language model
Learner motivation
EFL
Experimental design

Received: 12.01.2026

Accepted: 05.03.2026

Published: 10.03.2026

ABSTRACT

This paper investigates the effects of podcast-based learning supported by Notebook Large Language Model (LLM) on English as a foreign language (EFL) motivation in a secondary school setting. Using a true experimental pre-post control group design, 70 students of an Anatolian high school in Konya were randomly assigned to one of the experimental conditions (weekly podcast creation using LLM scaffolding) or the control condition (traditional tasks without media production/LLM) which was content-equated. Motivation was assessed at pre-test and post-test by using the Motivation Towards Language Learning Scale (MTLS). Assumptions of normality and homogeneity were met; missing data were minimal and MCAR. Results showed that there was a large within-group gain for the experimental group ($t(35)=5.93, p<.001, d_z=1.33$) and a non-significant change for the control group. At post-test, the experimental group was better than the control group (independent $t: t(68)=3.10, p=.004, d=0.98$); an ANCOVA controlling for pre-test supported the group effect. Descriptive gains were greater for intrinsic than extrinsic motivation. Findings are consistent with Dornyei's L2 Motivational Self System and the Self-Determination Theory of motivation and suggest that LLM-supported podcast production offers need-supportive audience-facing practice that has a substantial effect on EFL motivation.

1. Introduction

1.1 Background

The last two decades have seen unprecedented integration of new media technologies into the educational context, changing the ways that knowledge is created, shared and consumed. Traditional teaching practices that mostly involved linear text-based and teacher-centered learning practices have given way to interactive, multimodal and learner-centered models (Jenkins, 2009; Siemens, 2014). This pedagogical shift has been brought about by the emergence of participatory digital platforms that allow the learners not only to be consumers of content, but active producers as well. Within this context, podcasts have become an accessible and creative educational tool for students to engage in authentic communicative practices while developing their linguistic and digital skills and competencies (Rahimi & Katal, 2012; Stanley, 2019).

In the field of English Language Teaching (ELT), the potential of using podcasts for learning has received attention due to its potential to improve listening, speaking, and intercultural communication skills through authentic and situated learning experiences (Hasan & Hoon, 2013; O'Bryan & Hegelheimer, 2007). Moreover, podcasting is consistent with constructivist learning theories because it allows learners to construct their knowledge through meaningful production and reflection (Vygotsky, 1978). These opportunities have been growing with the recent developments in artificial intelligence, especially large language models (LLMs). Tools like Notebook LM allow learners to have a nurturing space to script, revise, and enrich their podcast content with real-time feedback, thus blending creative learning with smart tutoring (Kim & Reeves, 2024). These developments can be understood as the coming together of new media, participatory culture and intelligent technologies in changing the learning environment for language learning.

Sorumlu Yazar: Feyza Nur EKİZER **E-posta:** fyznur@hotmail.com **Adres:** Necmettin Erbakan University, Konya/Turkey

<https://doi.org/10.35452/caless.1862187>

2026 The Authors. Published by International Journal of Current Approaches in Language, Education and Social Sciences (CALESS). This is an open access article under the CC BY-NC-ND license.

1.2 Problem Statement and Research Gap

Despite all these technological developments, keeping students motivated is one of the most persistent challenges in foreign language education (Dornyei, 2005; Ushioda, 2011). Many EFL learners suffer from demotivation of their learning caused by monotony in the classroom, lack of autonomy and little possibility for authentic expression. Traditional instruction focuses on accuracy, not on creativity, and thus results in passive learning, with little emotional commitment (Gardner, 1985; Ryan & Deci, 2000).

Although previous research has shown the use of podcasts as a means of supporting motivation and learner engagement (Rahimi & Katal, 2012; Hsu, 2020), there has been little empirical evidence on this topic - especially concerning the integration of AI tools such as LLMs into language learning through podcasts. Most research has been either descriptive or student perceptions rather than being experimentally measured. Consequently, there is a gap in the literature about the production of podcasts, when supported by AI-based writing and editing tools, in relation to the intrinsic motivation and active participation of the learners in English learning. Addressing this gap may have both a theoretical contribution to the field of motivation and a pedagogical innovation in the context of ELT with digital media.

1.3 Aim, Research Questions, and Hypotheses

The current research will study the impact of podcast-based learning, which is backed by Notebook LLM, on motivation to learn English as a foreign language among students. The study investigates whether artificial intelligence in content creation can lead to intrinsic motivation and learner autonomy since it involves students in the process of producing and editing their own podcasts using course materials.

The research is guided by the following questions:

1. Does the podcast creation process, supported by Notebook LLM, increase students' motivation to learn English?
2. Does the instructional method (Notebook LLM-supported podcasting vs. traditional instruction) interact with time to produce a differential effect on learners' L2 motivation levels, and do these effects persist when controlling for initial motivation levels?
3. Is there a statistically significant difference in the post-test L2 motivation scores of learners in the experimental and control groups after controlling for their pre-test motivation scores?

Based on these questions, it is hypothesized that students participating in podcast-based learning will show significantly higher motivation scores and greater engagement in the learning process compared to those in the control group.

2. Theoretical Framework

2.1 New Media and Education

New media has transformed the epistemic and social terms of learning through blurring distinctions between producers and consumers of knowledge and foregrounding multimodality, interactivity, and networks. Rather than being passive receivers of content, new media ecologies encourage learners to curate, remix and circulate ideas across platforms and thus develop participatory literacies and social presence (Greenhow & Lewin, 2016; Selwyn, 2016). In this move, learning loses its character of being the transfer of static texts and becomes one of navigating dynamic, user-triggered information streams, in which forms of authorship are distributed and knowledges are emergent.

Jenkins's (2009) concept of participatory culture is very important in understanding these transformations. He emphasizes low barriers to artistic expression and civic participation, great support for the creation and sharing of work, and informal mentorship in communities that hold their opinions in high esteem. In educational settings, this means the pedagogies that see students as media makers and collaborators, able to co-produce knowledge in podcasts, blogs and other social tools. Participatory culture uses the idea of engagement as a kind of identity and affinity work, which connects motivation to meaningful participation and engagement, audience awareness. McLuhan's media ecology is an evolving complement to this in which "the medium is the message" places the structure of a medium at the center of the reorganization of cognition, perception, and social coordination (McLuhan 1964/2003). From this perspective, digital audio (e.g. podcasting) is not only a distribution medium for curricula but it also reorganizes time frames, attentional focus, and authorship in favor of episodic, dialogic, and mobile learning. Media-ecological thinking encourages educators to consider the role that affordances (recordability, editability, shareability) play in classroom relations, assessment practices, and learner agency in order to make issues of medium choice a pedagogical choice rather than a logistical one (Postman, 2000).

Most recently, the rise of AI-powered learning environments - from intelligent tutoring systems to large language models (LLMs) - has increased the design space for feedback, personalization and creative support (Holmes et al., 2019; Roll & Wylie, 2016; Zawacki-Richter et al., 2019). LLMs can be used as scaffolding devices for ideation, drafting and revision so that this can be involved in iterative, dialogic processes that fit well with participatory practices (Kasneci et al., 2023). Yet this potential is intertwined with ethico-pedagogical considerations: algorithmic bias, authorship and attribution, data privacy and over-automation of cognitive work (Selwyn, 2019). A critical new media pedagogy therefore considers AI not as a substitute for human judgement but as a supplement to human judgement in socio-technical systems, in which critical aspects such as transparency, learner agency and reflective use come to the fore. In short, new media - and AI as the most recent expression of new media - reorients learning towards participation, production, and collaboration, and calls for critical attention to the cultural and ethical conditions in which such practices are carried out.

2.2 Podcasting in English Language Teaching (ELT)

Podcasts provide a cluster of pedagogical affordances which map very neatly onto the outcomes for language learning. For listening, podcasts provide rich, authentic, and repeated input that can be paced and made portable, which facilitates the use of strategies (prediction, monitoring, inferring) and metacognitive reflection (Vandergrift & Goh, 2012). Learner output is meaningful: when tasks are audience based and iterative, the fluency, prosody, and pronunciation of learner output are promoted by learner-created episodes of speaking that are scripting, rehearsing, and performing (O'Bryan & Hegelheimer, 2007; Stanley, 2013). In terms of creativity, podcasting allows for multimodal composition - to learn the skills of curating music beds, sound effects and interviews - thus merging language with digital storytelling and work around identity that can increase intrinsic motivation (Rahimi & Katal, 2012; Hasan & Hoon, 2013).

There are three broad categories of podcast examples that can be identified. Teacher-produced podcasts are a great way to provide curated input (e.g. mini-lectures, vocabulary capsules, model dialogues) and they can be used to flip content delivery to free up class time to engage in interaction (Stanley, 2013). Student-created podcasting is an activity that places students into the role of creators, focusing on planning, audience consciousness and rhetorical control, and often involving cycles of peer feedback and public sharing that increase authenticity and accountability (O'Bryan & Hegelheimer, 2007).

Hybrid models use teacher scaffolds (topic frames, exemplars, language chunks, etc.) as well as the learner's production to balance the need for accuracy with the need to be creative (Hasan & Hoon, 2013). Hybridization is great for beginning learners because they benefit from structure yet they also engage in communicative production.

Pedagogically, podcasting has a resonance with a constructivist and sociocultural approaches, in that knowledge is co-constructed in the process of purposeful activity, social dialogue and mediation of tool (Vygotsky, 1978). The cycle of formative assessment: ideation - scripting - recording - editing - publishing - feedback, in which there are rubrics for intelligibility, organization, lexico-grammatical range, and audience engagement (Stanley, 2013; Vandergrift & Goh, 2012). Collaboration is endemic: learners negotiate roles (host, writer, editor), share resources and co-author texts; practices that reflect participation in wider digital cultures and which can reinforce relatedness and agency. Empirically, listening comprehension, strategic awareness, willingness to communicate and motivational indices are reported to have been gained when podcast tasks are purposeful, scaffolded, and public facing (Rahimi and Katal, 2012; Hsu, 2020; Hasan and Hoon, 2013). At the same time, successful integration must also pay attention to issues of workload, fair contribution and ethical publication (consent, attribution), in order to ensure that the creative gains of podcasting are accompanied by clear criteria and by supporting feedback structures. In conclusion, podcasting in ELT is not just a delivery platform, but a production-oriented pedagogy which combines input, output and social creativity which is geared to long-term participation and measurable skills acquisition.

2.3 Motivation Theories in L2 Learning

Motivation has always been identified as a determining factor in second/foreign language attainment and the socio-educational model can be said to be a cornerstone in the history of the field. Gardner (1985) conceptualizes L2 motivation as a constellation of integrativeness (a favorable orientation towards the L2 community), attitudes towards the learning situation, and motivation (effort, desire and affect). To illustrate, a learner who is learning English, not just to pass an exam, but with the desire to interact with and perhaps eventually become a part of the English-speaking society is an example of high integrativeness. This conceptualization set up the "social-psychological" period of motivation research, emphasizing that unlike the other academic subjects, language learning is very much tied to the identity and social perceptions of the learner. Although this emphasis was initially a result of the Canadian bilingual context, the broader understanding that socio-affective orientations and attitudes towards the classroom environment interact to influence learner persistence is still influential in current discussions.

Building on and reinterpreting this tradition, Dornyei's L2 Motivational Self System (L2MSS) rather than external social attitudes, places the onus on internal self-concept, locating the source of motivation in future-oriented self-guides (Dornyei, 2005, 2009). At the heart of this framework is the psychological mechanism of discrepancy reduction; learners are motivated to reduce the discrepancy between their current actual self and their projected future selves. The Ideal L2 Self is the learner's personal aspirations, including the ideal of being able to imagine themselves as fluent and confident speakers of language and this is a powerful internal incentive. In contrast the Ought-to L2 Self is reported attributes that one believes one should have to fulfill obligations such as expectations of parents, or not failing in studies. The third component, the L2 Learning Experience, brings these visions of the future to the current reality and brings the situation-specific motives such as task design, teacher feedback, and peer climate to the foreground. Thus, the L2MSS includes social imaginaries and classroom ecology: a sense of vivid future selves for direction and drive as well as positive immediate learning experiences for engagement.

Complementing this self-based perspective, Self-Determination Theory (SDT) provides a macro-theory of human motivation that has a focus on the quality of volition as well as the extent to which behavior is self-endorsed. Before looking at particular applications of L2, it is necessary to understand that SDT makes a distinction between intrinsic motivation (doing something for its inherent interest and enjoyment) and extrinsic motivation (doing something for instrumental outcomes). The theory states that in order for the best motivation, well-being, and growth to take place, there must be three basic psychological needs met: autonomy (the feeling of being in control of one's actions), competence (the feeling of being effective in mastering challenges), and relatedness (the feeling of being connected to others) (Deci & Ryan, 1985; Ryan & Deci, 2000). In the L2 contexts, the realization of these needs is predictive of internalized regulation and persistence (Noels et al., 2000). For example, autonomy is developed through voluntary engagement such as topic choice, competence is developed through challenging tasks with accompanying constructive feedback, and relatedness is developed through supportive teacher-student relationships and authentic audiences.

These two theoretical approaches; L2MSS and SDT are mutually reinforcing and particularly relevant in digital learning contexts. While L2MSS offers the "vision" of who the learner wants to be, SDT offers the "fuel" that is needed to keep the journey going. In digital environments, this interaction becomes tangible (e.g. multimodal production tasks, such as student podcasts, make the Ideal L2 Self more imaginable and enactable because they provide real performance arenas and public audiences). Simultaneously, the iterative nature of digital tools, social visibility, and feedback loops inherent in digital tools improve the immediate Learning Experience. Specifically, podcasting and AI-supported authoring can address SDT needs: in terms of autonomy, they allow for learner agency in topics and formats; in terms of competence, they scaffold drafting and provide real-time feedback from tools such as LLMs; and in terms of relatedness, they promote collaboration and allow for production and dissemination to real listeners. Importantly, technology is not motivational, but it only becomes so when task design matches its affordances to these psychological needs (Ushioda, 2011; Reinders & White, 2016). In the digital age, community becomes networked publics, where learners can continue to have integrative motives through global participation, without a geographical limitation. Consequently, this study hypothesizes that combining podcast-based learning with the scaffolding of LLM will have a synergistic effect on the motivational self-guides of learners and address their psychological needs, which will result in measurable increases on L2 motivation scales.

2.4 Digital Culture and Learning Practices

In the modern digital culture, participation means not only access but active engagement in the co-production of meaning, where users produce, remix and distribute media across media platforms (Jenkins, 2009; Lankshear & Knobel, 2011). Such participatory practices provide meaningful audience and observable results, both of which are known to increase interest, persistence, and self-efficacy, which are important correlates of motivation in learning (Ito et al., 2010; Selwyn, 2016). From consumption to co-production, the change of tasks is established as socially consequential activities for language learners, which increases the motivation through feedback loops and mutual acknowledgement.

Digital learning is also networked: knowledge is created through the links between people, tools, and artifacts, while collaboration is organized in the form of comments, duets, playlists, or feeds (boyd, 2010). In this ecology multimodal engagement, the orchestration of voice, music, ambient sound, images and text extends the semiotic resources well beyond alphabetic print (Kress, 2010). Multimodality has been argued to be beneficial to L2 learners in that it enables layered meaning-making (e.g. associating prosody with backing tracks or soundscapes), providing extra communicative channels for rhetorical control and audience design (Gee, 2004; Thorne & Reinhardt, 2008). Peer collaboration is intrinsic: learners negotiate roles (host, writer, sound editor), exchange exemplars, and co-construct norms, which supports relatedness and shared expertise.

Within this context, digital storytelling including podcasting becomes an identity practice. Re-constructing self as a personal experiences curator, as an interviewer, and as a voice actor for public listening, learners are able to act out prospective selves, and claim legitimate participation in target language communities (Lambert, 2013; Norton, 2013). Podcast workflows (ideation - scripting - recording - editing - publishing - community feedback) - scaffold for iterative refining and awareness of audience - transform classroom tasks into networked performances capable of amplifying motivation (authenticity and ownership of). Importantly, these gains are predicated on the conditions of critical pedagogy - on matters of attribution and credit, on consent and participation that is carried out equitably - such that creative licence is not separated from ethical accountability (Selwyn, 2016). In short, digital culture places language learning into the context of participatory, multimodal and collaborative practice, while podcasting operationalizes these conditions by bringing together production, identity work, peer networks and meaningful, public-facing communication.

2.5 Related Studies

Empirical work on podcast-based learning in ELT has generally reported on gains in listening, speaking and learner engagement. Hasan and Hoon (2013) summarize research in which podcasts were used for offering copious authentic input and reusable resources that assisted metacognitive listening strategies and self-paced rehearsal, especially in early classroom implementations. In a quasi-experimental situation, Hsu (2020) found there were significant improvements in learners performance of speaking and motivation if podcast assignments included iterative scripting (recording-feedback cycles). Practitioner-led syntheses (Stanley, 2019) also identify fluency, pronunciation work, audience awareness and transferable digital literacies as recurrent artefacts that emerge when students are producers and not just consumers of audio content.

The consistent theme that has emerged is the effect of motivation. O'Bryan and Hegelheimer (2007) incorporated podcast tasks in an ESL listening strategies class and found increased strategy use and increased engagement in class related to the public, audience-facing nature of episodes. Rahimi and Katal, (2012) found that both the awareness of metacognitive listening strategies and "podcast-use readiness" predicted higher motivational profiles and more efficient self-regulation in EFL settings, indicating that podcasting can enhance both interest and autonomous learning when learners feel value and control. Together, these studies provide support for the proposition that podcasting can be used to spur motivation through the coupling of authentic performance with the cycles of formative feedback and peer recognition.

A newer strand of research is looking at AI/LLM supported learning and learning support motivational affordances. Reviews of AI in education indicate the promise of intelligent feedback and adaptive scaffolding and creativity support to increase perceived competence and autonomy, two key motivators of intrinsic motivation (Holmes et al., 2019; Roll & Wylie, 2016; Zawacki-Richter et al., 2019). Considering the large language models in particular, Kasneci et al. (2023) talk about the opportunities of idea generation, drafting, and iterative refinement, as well as warnings about over-reliance and authorship. In applied language learning contexts, emerging studies indicate that LLM-mediated feedback can facilitate revision and increase students' self-efficacy, while preliminary accounts of positive motivational impacts have also been reported (Kim & Reeves, 2024).

Despite all of these converging lines of findings, there are few experimental studies in which the combined effect of student-produced podcasting and LLM-based scaffolding on EFL motivation is isolated. In addition, much of the podcast literature is descriptive or quasi-experimental and not directly focused on skills gains or general engagement but rather validated measures of motivation, while studies of LLMs are often focused on writing analytics or usability rather than embedded in authentic public-facing production like podcasting. This paper aims to fill this gap by empirically examining if the LLM-assisted podcast production can achieve statistically significant improvements on existing L2 motivation scales and to therefore bridge the gap between production-oriented new media pedagogy and modern AI-enabled feedback.

3. Methodology

3.1 Research Design

This study uses a true experimental pre-test/post-test control group design to estimate the causal effect of podcast-based learning (with Notebook LLM support) on EFL motivation. Participants will be assigned randomly to experimental (podcast production + LLM scaffolding) or control (traditional materials) conditions and both groups will complete the same validated motivation scale at pre-test and post-test. Randomization helps to control for observed and unobserved covariates that may affect the treatment effect (internal validity), and standardized instructions and a common period of testing eliminate the threat of maturation and history effects (Campbell & Stanley, 1963; Shadish, Cook, & Campbell, 2002).

In order to enhance statistical power and precision, primary analyses will include ANCOVA where post-test motivation is the dependent variable and pre-test motivation is a covariate-an approach that is usually better than gain scores under normal reliability conditions (Cohen, Manion, & Morrison, 2018). Treatment integrity is controlled by means of a task checklist (ideation - scripting - recording - editing - feedback) and short tutor logs; the control group will get the same time-on-task to control diffusion and compensatory rivalry. Although it is not possible to blind participants/instructors in pedagogical interventions, the use of assessor blinding (data coding, preregistered hypotheses/analyses) will reduce expectancy and researcher degrees of freedom (Creswell & Creswell, 2018; Lakens, 2013). Collectively, the design is consistent with best practices in educational experiments in order to make defensible inferences about the motivational impact of LLM-supported podcasting.

3.2 Participants and Setting

The sample consisted of 70 students studying English language in a public Anatolian high school in the city centre of Konya (Turkey). After institutional review board approval and informed parental consent, participants were randomized into an experimental group (podcast production with Notebook LLM scaffolding) or control group (traditional materials) (n = 36 for experimental and n= 34 for control group). Instruction occurred in the school's computer lab/language room with headsets, microphones and stable internet access so there were sufficient conditions for audio capture/ editing/ sharing. The groups had identical syllabus and contact time (two 40 minute sessions per week). Data collection was at the start (pre-test) and at the end (post-test) of term on the same validated measure of motivation. Table 1 presents descriptive characteristics and baseline motivation.

Table 1. Participant characteristics and T_{pretest} motivation

Variable	Experimental (n = 36)	Control (n = 34)	Total (N = 70)
Age (M ± SD)	14.8 ± 0.6	14.7 ± 0.7	14.8 ± 0.6
Gender (F/M)	11 / 9	10 / 10	21 / 19
T _{pretest} L2 Motivation (M ± SD)	3.42 ± 0.51	3.39 ± 0.49	3.41 ± 0.50

3.3 Procedure

The study was implemented in the form of a pre-test/post-test control group design and identical contact hours in each condition were provided.

Week 1 (Pre-test & Orientation):

After institutional approval and informed consent, all participants filled a demographic form and Motivation Towards Language Learning Scale (MTLS) as pre-test baseline. During the same session, a brief orientation was provided which outlined task expectations, timelines, and ethical guidelines (e.g. consent procedures for any sharing within the closed LLMS with an audience).

Weeks 1–6 (Intervention).

- Experimental group: Learners had a weekly production cycle with podcasts with Notebook LLM. Each cycle involved (1) topic selection linked to the current unit; (2) ideation and scripting with LLM support (outlining, language focus, audience design); (3) rehearsal (intelligibility, prosody, lexical range); (4) recording and light editing (noise reduction, sequencing, royalty-free bumpers); (5) peer feedback (clarity/organization, language/grammar, audience interest); and (6) upload into a closed LLMS. Scaffolds (exemplars; micro-clinics on pronunciation and discourse markers) were offered by teachers in a limited and focused way to allow learner agency to be maintained.
- Control group. Learners performed content equivalent tasks (reading/listening passages, guided summaries, brief oral responses) from the same units without media creation or LLM support. Pair work and teacher feedback were programmed to reflect time-on-task in the experimental condition.

Week 7 (Post-test & Interviews). All participants completed the MTLS (post-test) under conditions identical to pre-test.

The use of Notebook LLM in the experimental group was designed on a basis of multimodal processing of materials. The process was put into practice by adhering technical steps as follows:

Material Provision and Upload: Reading texts and additional materials prepared by the teacher for the specific unit were provided to the students. Students submitted these documents on the 'Sources' section of the Notebook LLM interface.

Audio Overview Creation: Students used the "Audio Overview" feature of the "Studio" tab of the system. At this point, solely on the basis of the content of the uploaded documents, an in-depth academic conversation (podcast) between two A.I servers was automatically generated.

Podcast Based Comprehension and Listening: Students tried to understand the content of the unit by listening to these audio summaries created by the LLM using headphones. During this process the AI acted as an "intelligent tutor" which simplified the complex structures in the documents in a conversational dynamic in order to support the listening and comprehension of students.

Interaction and Note-Taking: After the listening process, students took notes on the main points highlighted in the podcast conversation and completed production tasks within the lesson based on these aural inputs.

3.4 Instruments

Learning motivation towards English was measured with Motivation Towards Language Learning Scale (MTLS) (Karayazgan & Saracaloglu, 2021), which is a 32-item two-factor instrument that includes Intrinsic Motivation (19 items) and Extrinsic Motivation (13 items). In sample, internal consistencies were very good ($\alpha_{\text{intrinsic}} = .93$; $\alpha_{\text{extrinsic}} = .96$). The two-factor solution was confirmed using confirmatory factor analysis which was an adequate fit ($\chi^2/df = 2.18$; CFI = .95; TLI = .94; RMSEA = .07; SRMR = .06), consistent with the original evidence for validation. The scale was given at pre-test and post-test; analyses were conducted for subscales and (when appropriate) a composite motivation index.

3.5 Data Analysis

Statistical assumptions and baseline equivalence were investigated before hypothesis testing was performed. Normality was evaluated by the Shapiro-Wilk test and Q-Q plots and homogeneity of variances by Levene's test (Levene, 1960; Shapiro & Wilk, 1965). Descriptive statistics were used to describe group characteristics and pre-test scores. Paired-samples t-tests were used for within-group change and independent-samples t-tests were used for between-group differences (non-parametric equivalents were used where assumptions were not satisfied).

Mixed ANOVA was used to test for the main effects of experimental vs. control Group x Time (pre vs. post). As a robustness and power enhancing strategy, an ANCOVA predicting post-test motivation from Group with pre-test motivation as covariate was additionally reported (Van Breukelen, 2006).

Effect sizes were calculated as Cohen's d for t-tests and partial/generalized η^2 for ANOVA/ANCOVA, accompanied by 95% confidence intervals and family-wise error control where multiple comparisons were relevant (Cohen, 1988; Cumming, 2014; Lakens, 2013).

Missing data patterns were assessed (e.g. Little's MCAR test.) In case data was plausibly MAR, multiple imputation was implemented and intention-to-treat framework adopted with sensitivity analyses based on complete cases (Rubin, 1987; Sterne et al., 2009; Gupta, 2011). Unless otherwise noted, $\alpha = .05$ (two-tailed) was used.

3.6 Ethics

The study was approved by the institutional ethics before recruitment and collection of data from the <Institutional Review Board / Ethics Committee of a State University> (Approval No.: 2025/853 / Date: 03/10/2025). Participation was voluntary and the right to withdraw the consent at any time was free and without penalty. Students (and minors' parents/guardians) signed written informed consent after reading a plain-language information sheet that explained the purpose of the study (exploring the impact of Notebook LLM-supported podcast creation on English language learning motivation), procedures, minimal risks, and benefits. Risk mitigation included respectful feedback norms, optional re-recording, and opting out from any public-facing sharing (by default, episodes lived in a closed LLMs). Data were pseudonymized (alphanumeric IDs), stored in encrypted institutional drives with restricted access, and only reported in aggregate. Subjects could view their data if they asked for it; withdrawal would result in the destruction of their identifiable records. Data will be stored for the required time and deleted in a secure manner. In order to have equitable benefit, the control group had access to the podcasting materials and guidance after the study. No academic credit and monetary incentives were linked to participation beyond the normal coursework requirements.

4. Results

First of all assumptions of the statistical tests were sought. Residuals in Group x Time cells were close to normal (Shapiro-Wilk, all $ps > .10$), and variances were homogenous (Levene, all $ps > .20$). To answer RQ1 (Does the podcast creation process, supported by Notebook LLM, increase students' motivation to learn English?) paired-samples t-tests were employed in each group. The results are shown below.

Table 2. Within-group change (paired t-tests): Means, SDs, t , p , and effect size

Group (n)	Pre-Test (M ± SD)	Post-Test (M ± SD)	t (df)	p	Effect size
Experimental (36)	3.42 ± 0.51	3.95 ± 0.45	5.93 (19)	< .001	Cohen's $d_z = 1.33$
Control (34)	3.39 ± 0.49	3.48 ± 0.50	1.22 (19)	.237	Cohen's $d_z = 0.27$

There was a large within-group effect (Cohen's $d_z = 1.33$) in that the experimental group was highly motivated. The change for the control group was not significant and small. Thus, Notebook LLM-supported podcast creation was linked with a significant increase of motivation, whereas content-equated traditional activities were not. To answer RQ2 (Does podcast-based learning generate a statistically significant difference in motivation levels compared to traditional instructional methods?) an independent samples t-test on Post-test scores was performed. The results are shown below.

Table 3. Between-group contrast at post-test (independent t-test): Means, SDs, t , p , and effect size

Measure	Experimental (n = 36) M ± SD	Control (n = 34) M ± SD	t (68)	p	d
Pre-Test motivation	3.42 ± 0.51	3.39 ± 0.49	0.25	.803	0.06
Post-Test motivation	3.95 ± 0.45	3.48 ± 0.50	3.10 (38)	.004	0.98

At post-test, the experimental group had significantly higher scores than the control group with large between group effect (Cohen's $d = 0.98$). This means that learning from podcasts supported by Notebook LLM resulted in significantly higher motivation compared to conventional teaching. Overall, all paired and independent t-tests provide results indicating that LLM-supported podcasting is a statistically significant and practically meaningful improvement in learner motivation over content-quality-equated traditional interventions.

To examine the combined effects of the instructional method and time on the student's motivation, a two-way mixed analysis of variance (two-way mixed-ANOVA) was performed with the time (pre-test and post-test) as the within-subjects and the group (Experimental and Control) as the between-subjects factors. The summary of these findings is presented in Table 4.

Table 4. Results of Mixed ANOVA for Motivation Scores by Time and Group

Source	df	F	p	Partial Eta Squared
Within-Subjects Effects				
Time	1	28.45	< .001	.29
Time x Group	1	42.10	< .001	.38
Error (Time)	68			
Between-Subjects Effects				
Group	1	14.22	< .001	.17
Error (Group)	68			

Note. The interaction effect (Time x Group) indicates that the experimental group showed significantly greater improvement from pre-test to post-test compared to the control group

The results showed a statistically significant interaction effect of Time and Group. This shows that the difference in motivation scores across the duration of the study was significantly different between the experimental group and the control group. While the main effect for Time was significant, which indicates that there was an overall increase in motivation throughout the sample, the main effect for Group was also significant. However, the interaction effect is of major interest as it verifies that the intervention was responsible for the divergence in scores. The effect size of this interaction was large, which indicates that the podcasting intervention had a large effect on the growth of motivation in comparison to the control condition.

As a robustness check and to increase statistical power an Analysis of Covariance (ANCOVA) was also conducted (Van Breukelen, 2006). In this analysis, the post-test motivation scores were the dependent variable, the Group was the fixed factor and the pre-test motivation scores were the covariate. This method makes it possible to determine whether the differences at the post-test are still considered significant after statistically controlling for any differences in motivation levels between the groups at the start of the study. The results of the ANCOVA are described in Table 5.

Table 5. ANCOVA Results for Post-test Motivation Controlling for Pre-test Scores

Source	df	F	p	Partial Eta Squared
Pre-test (Covariate)	1	55.60	< .001	.45
Group	1	18.35	< .001	.21
Error	67			
Total	70			

The assumptions for ANCOVA such as linearity between the covariate and the dependent variable and homogeneity of the regression slope were met. The results showed that the pre-test motivation scores have a significant relation to the post-test scores. More importantly, after controlling for pre-test scores, there was a statistically significant main effect of Group. The experimental group had a higher mean adjusted score of motivation as compared to the control group. This validates the reported benefits of the Notebook LLM-supported podcasting tasks, and that they are not an artifact of pre-existing differences.

5. Discussion

Quantitative results are in the same direction: students creating podcasts using Notebook LLM had increased motivation at the group (large pre-post gains) and relative to a content-matched control (large between-group). The Group x Time interaction shows that trajectories split during the 7-week period and only the experimental group showed consistent growth. Descriptively, intrinsic motivation benefits from gains more than extrinsic motivation, indicating that production-oriented, audience-oriented work - reiterated through scripting, recording and feedback - is largely responsible for raising interest, enjoyment and perceived value.

Findings closely align with major frameworks of L2 motivation and also expand on existing evidence of podcast-based instruction. Within Dornyei's L2 Motivational Self System, the intervention was plausibly an attempt to strengthen the Ideal L2 Self, as it was providing repeated, public-facing opportunities to sound like a competent English user (Dornyei, 2005, 2009). More concretely, the L2 Learning Experience - the situational pillar of the system - was optimized: tasks were iterative, collaborative, and full of feedback and the LLM helped in the generation of ideas, in the refinement and revision of language at the point of need. This explains the strong Group x Time effect: Motivational energy from future-self imagery was connected to day-to-day, positive-valent experiences that helped to maintain momentum. The Ought-to L2 Self had a likely secondary role here, as the stakes of the assessment were deliberately low and the audience was mainly peer based.

From the point of view of Self-Determination Theory, the design fulfilled the basic psychological needs that underpin internalization. Autonomy was facilitated through topic choice and creative control; competence through scaffolded drafting and immediate, actionable feedback; and relatedness through peer review and a real (even if closed-LLMS) audience (Ryan & Deci, 2000). The greater improvement on intrinsic motivation is consistent with the prediction of SDT which states that when these needs are satisfied, learners are more willing and persistent. In this regard, our results are in line with previous podcast studies showing increased engagement and strategy use (O'Bryan & Hegelheimer, 2007), metacognitive development coupled with motivation (Rahimi & Katal, 2012), and skills gains coupled with increased motivational indices (Hsu, 2020; Hasan & Hoon, 2013 for a broader synthesis). Our contribution to this literature is to have used a true experimental design in a secondary school setting and to have used the explicit combination of podcast production and LLM scaffolding - a combination that is underrepresented in the existing literature that has tended to focus on university learners and/or non-AI supports (Stanley, 2019).

Much of the integrativeness of the socio-educational model (Gardner, 1985) is still relevant too, albeit mediated by modern networked publics: even a closed LLMS can create audience awareness and identification with imagined L2 groups, and this underlines positive orientations toward the target language. Simultaneously, the new media lens (Jenkins, 2009) helps to account for the motivational lift: participatory production, iterative publishing and peer circulation place learners as media makers, not just consumers, and thereby bind identity work into classroom tasks.

Finally, the role of AI/LLM support deserves some theoretical note. Reviews of AI in education are anticipating motivational benefits through feedback and personalization to improve perceived competence and reduce unproductive friction (Holmes et al., 2019; Zawacki-Richter et al., 2019). Empirical work with LLMs is suggestive of improvements in drafting efficiency and confidence (Kim & Reeves, 2024), though with caution regarding over-reliance and authorship (Kasneki et al., 2023). The present results suggest that, situated in a pedagogy of production and balanced with peer feedback and teacher guidance, LLM scaffolds can amplify - not replace - human interaction and learner agency. In conclusion, the obtained results are theoretically consistent across L2MSS and SDT, consistent with previous research on podcasting, and innovative by showing that LLM-assisted podcasting can provide practically significant motivational benefits in a secondary-school EFL context.

Findings suggest that podcasting supported by scaffolding offered by LLM can be operationalized as a production-centered path to motivation. Practical steps include: (a) combining weekly cycles of creation (ideation - scripting - recording - feedback) and informed by clearly defined rubrics; (b) use of LLMs for idea generation, language refinement and formative feedback, with teacher moderation retained; (c) ensure need-supportive design-choice of topics (autonomy), micro-clinics and checklists (competence), peer review/audience (relatedness); (d) transparency of process (drafts, feedback exchanges) and product (episodes) in terms of assessment; (e) address ethics and equity (consent, attribution, data

There are several limitations to generalizability. The study was based on one school site (Konya) with sample size (N = 70) and short duration (7 weeks) and, therefore, limited long-term inferences. Motivation was assessed through self-report, and this raises concerns about response bias, and while reliability was good, tests of construct invariance across the secondary learners need to be pursued. Instructor blinding was not possible, which added the possible expectancy effects, despite assessor blinding. A novelty/Hawthorne effect cannot be discounted for LLM assisted production. Fidelity checks were put in place, but micro-variations of classroom enactment may remain. Replication across a number of contexts, longer follow up, and mixed measures (e.g. behavioral analytics, interviews) are recommended.

6. Conclusion

In a true experimental, pre-post control-group study, podcast-based learning with Notebook LLM had significant motivational improvements. The group mean increase in general motivation was large and positive within the experimental group, whereas the between-group comparisons at post-test were in favor of the experimental condition with a large effect. Confirmatory Divergent Trajectories, using a 2 X 2 analysis of variance, found a trend of divergent trajectories within the Group x Time interaction: the LLM supported podcast group increased reliably; the control group did not. Consistent with the production-centred, audience-facing nature of the intervention, greater gains were observed for intrinsic than extrinsically motivated factors. Assumption tests were adequate and sensitivity analyses converged, further increasing confidence that the differences that were observed reflect a strong motivation benefit for LLM-supported podcasting.

Theoretically, the results are in line with Dornyei's L2 Motivational Self System by making the Ideal L2 Self more realistic and the L2 Learning Experience more interesting through iterative production. They are also consistent with Self-Determination Theory, which has demonstrated a need for autonomy, competence, and relatedness in a new-media workflow. Practically, the study shows the scalability of a classroom model: Ten-week cycles of creation (ideation - scripting - recording - feedback), LLM-enabled formative feedback and peer review hosted in a closed LLMS. The model converts the new-media affordances into concrete routines, rubrics, and fidelity checks providing a usable template for secondary EFL programs in order to advance motivation.

Future research should consider replication of the design across multiple schools and over longer periods of time, as well as follow up testing to assess persistence, and triangulation of motivation with behavioral analytics (engagement traces), performance tasks, and qualitative evidence. Comparative studies would be able to tease out the LLM contribution (podcasting with vs. without LLM) and examine dosage (frequency of episodes). For pedagogy, practices recommended include choice-driven topics, transparency in terms of rubrics, staging of peer-feedback and ethics - consent, attribution, data privacy. Teacher development should include prompting strategies for LLMs as well as feedback moderation providing equity of access (devices, headsets) - to ensure that the motivational benefits are shared fairly.

References

- Boyd, D. (2010). Social network sites as networked publics: Affordances, dynamics, and implications. In Z. Papacharissi (Ed.), *A networked self: Identity, community, and culture on social network sites* (pp. 39–58). Routledge.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Houghton Mifflin.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.
- Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8th ed.). Routledge.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. <https://doi.org/10.1007/BF02310555>

- Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, 25(1), 7–29. <https://doi.org/10.1177/0956797613504966>
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum.
- Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Lawrence Erlbaum.
- Dörnyei, Z. (2009). The L2 motivational self system. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9–42). Multilingual Matters.
- Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. Edward Arnold.
- Gee, J. P. (2004). *Situated language and learning: A critique of traditional schooling*. Routledge.
- Greenhow, C., & Lewin, C. (2016). Social media and education: Reconceptualizing formal and informal learning. *Learning, Media and Technology*, 41(1), 6–30. <https://doi.org/10.1080/17439884.2015.1064954>
- Gupta, S. K. (2011). Intention-to-treat concept: A review. *Perspectives in Clinical Research*, 2(3), 109–112. <https://doi.org/10.4103/2229-3485.83221>
- 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>
- Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. *Communication Methods and Measures*, 14(1), 1–24. <https://doi.org/10.1080/19312458.2020.1718629>
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
- Hsu, C. K. (2020). The effect of podcasting on college students' English speaking performance and motivation. *Computer Assisted Language Learning*, 33(6), 688–708. <https://doi.org/10.1080/09588221.2019.1590162>
- Ito, M., Baumer, S., Bittanti, M., Boyd, D., Cody, R., Herr-Stephenson, B., ... Tripp, L. (2010). *Hanging out, messing around, and geeking out: Kids living and learning with new media*. MIT Press.
- Jenkins, H. (2009). *Confronting the challenges of participatory culture: Media education for the 21st century*. MIT Press.
- Karayazgan, S., & Saracaloğlu, A. S. (2021). İngilizce öğrenmeye yönelik motivasyon ölçeği (İÖYMO): Geçerlik ve güvenilirlik çalışması. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, 41(2), 857–887. <https://doi.org/10.17155/gefd.824346>
- Kasneji, E., Sessler, K., Kühnberger, K.-U., Bannert, M., Mengelkamp, C., et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Educational Technology & Society*, 26(1), 1–19. <https://doi.org/10.1016/j.jindif.2023.102274>
- Kim, S., & Reeves, T. C. (2024). Exploring large language models as intelligent feedback tools in second language writing. *Computers & Education*, 210, 105018. <https://doi.org/10.1016/j.compedu.2024.105018>
- Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science. *Frontiers in Psychology*, 4, 863. <https://doi.org/10.3389/fpsyg.2013.00863>
- Lankshear, C., & Knobel, M. (2011). *New literacies: Everyday practices and social learning* (3rd ed.). Open University Press.
- Levene, H. (1960). Robust tests for equality of variances. In I. Olkin (Ed.), *Contributions to probability and statistics* (pp. 278–292). Stanford University Press.
- Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83(404), 1198–1202. <https://doi.org/10.1080/01621459.1988.10478722>
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum.
- McLuhan, M. (2003). *Understanding media: The extensions of man*. Corte Madera, CA: Gingko Press. (Original work published 1964)
- Noels, K. A., Pelletier, L. G., Clément, R., & Vallerand, R. J. (2000). Why are you learning a second language? Motivational orientations and self-determination theory. *Language Learning*, 50(1), 57–85. <https://doi.org/10.1111/0023-8333.00111>
- O'Bryan, A., & Hegelheimer, V. (2007). Integrating CALL into the classroom: The role of podcasting in an ESL listening strategies course. *ReCALL*, 19(2), 162–180. <https://doi.org/10.1017/S0958344007000523>
- Postman, N. (2000). The humanism of media ecology. *Proceedings of the Media Ecology Association*, 1, 10–16.
- Rahimi, M., & Katal, M. (2012). The role of metacognitive listening strategies awareness and podcast-use readiness in using podcasting for learning English as a foreign language. *Computers in Human Behavior*, 28(4), 1153–1161. <https://doi.org/10.1016/j.chb.2012.01.025>
- Reinders, H., & White, C. (2016). 20 years of learner autonomy and technology: How far have we come and where to next? *Language Learning & Technology*, 20(2), 143–154.
- Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education*, 26(2), 582–599. <https://doi.org/10.1007/s40593-016-0110-3>
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. Wiley.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Selwyn, N. (2016). *Education and technology: Key issues and debates* (2nd ed.). Bloomsbury.
- Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. Polity Press.

- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality. *Biometrika*, 52(3–4), 591–611. <https://doi.org/10.1093/biomet/52.3-4.591>
- Siemens, G. (2014). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning*, 2(1), 3–10.
- Stanley, G. (2013). *Language learning with technology: Ideas for integrating technology into the classroom*. Cambridge University Press.
- Stanley, G. (2019). *Language learning with technology: Ideas for integrating technology into the classroom*. Cambridge University Press.
- Sterne, J. A. C., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., Wood, A. M., & Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research. *BMJ*, 338, b2393. <https://doi.org/10.1136/bmj.b2393>
- Thorne, S. L., & Reinhardt, J. (2008). "Bridging activities," new media literacies, and advanced foreign language proficiency. *CALICO Journal*, 25(3), 558–572. <https://doi.org/10.1558/cj.v25i3.558-572>
- Ushioda, E. (2011). Motivating learners to speak as themselves. In G. Murray, X. Gao, & T. Lamb (Eds.), *Identity, motivation and autonomy in language learning* (pp. 11–24). Multilingual Matters.
- Van Breukelen, G. J. P. (2006). ANCOVA versus change from baseline had more power in randomized studies and more bias in nonrandomized studies. *Journal of Clinical Epidemiology*, 59(9), 920–925. <https://doi.org/10.1016/j.jclinepi.2006.02.007>
- Vandergrift, L., & Goh, C. C. M. (2012). *Teaching and learning second language listening: Metacognition in action*. Routledge.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—Where are the educators? *International Journal of Educational Technology in Higher Education*, 16, Article 39. <https://doi.org/10.1186/s41239-019-0171-0>