Soundtracks for educational games with Generative Artificial Intelligence: an empirical study

Fabrizio Honda[1], **Jeniffer Macena**[1], **Isabela Amora**[1],
Endrenrique Soares[1], **Marcela Pessoa**[1], **Fernanda Pires**[1]

[1]School of Technology – State University of Amazonas (EST-UEA)
ThinkTEd Lab – Research, Development, and Innovation in Emerging Technologies

{icabe.lic24, ecso.lic24,fpires,msppessoa}@uea.edu.br,

{fabrizio.honda,jeniffer.souza}@icomp.ufam.edu.br

Abstract. Introduction*: Soundtracks (music) play an important role in educational games, helping to keep the player in a state of flow. However, many educational game developers do not have the expertise to create soundtracks, and it is not always possible to find free and non-copyrighted materials that are suitable for the artifacts created.* ***Objective:*** *This paper presents an empirical study on using Generative Artificial Intelligence to create soundtracks for educational games. Three AI models were used to achieve this goal (MusicFX, Riffusion, and Suno AI), which generated music for three different educational games.* ***Steps****: The research was carried out through steps that involve selecting models and games, construction of the prompt, and analysis/collection of data.* ***Results:*** *The results of the evaluation with students/computer professionals and educational game developers indicate promising results in the generation of music, mainly regarding the ability to create quality soundtracks that match the games, but also reservations regarding repetitive elements, generic sounds, slow rhythm, etc.*
Keywords *Music, Education, Games, Artificial Intelligence, Soundtracks.*

1. Introduction

The use of games in the educational context has grown gradually, presenting itself as an attractive alternative to traditional methods [Genesio et al. 2024, Guedes e Mello 2024].Games structured with interactive mechanics, engaging narratives, and clear objectives promote active knowledge construction while stimulating students' cognitive skills [Alves 2010, Fernandes 2010, Plass et al. 2015]. In this scenario, educational games not only generate attention and engagement but also act as environments where content is explored in a dynamic and relevant way [Silva 2025].

In these games, soundtracks (music) play an important role in player engagement, contributing to the game experience and promoting flow, a state of concentration in which the individual is not interrupted by external events [Levy 2015, Csikszentmihalyi 1990].Studies indicate that music in games contributes to player immersion and enjoyment, with significant statistical effects when compared to scenarios where music is absent [Caroux e Pujol 2024, Evans 2018]. Furthermore, sounds in games have a motivational function, signal information, and direct the player's attention. They also act as design elements capable of evoking emotions, influencing learning through emotional design [Plass e Kaplan 2016].

However, despite the contribution of soundtracks in games, in academic settings where games are designed by students/teachers and in independent studios, it is common for these individuals to have little or no training in musical composition [Lima et al. 2021]. In addition to artistic knowledge, creating games can require interdisciplinarity from their developers, encompassing aspects such as engineering, pedagogy, psychology, and elements of entertainment. [Battistella et al. 2014, Plass et al. 2015]. In addition to being game designers, they also take on multiple roles: developers, testers, designers, sound designers, etc. Therefore, most of them do not have expertise in composing music for games, and they have difficulty finding materials online that are free, non-copyrighted, and match the game's theme. Furthermore, although they share common goals with entertainment games, educational games need to consider the cognitive load and rhythm of the music because, if the soundtrack is poorly applied, it can be perceived as redundant information and distract players, which can influence the learning process [Linek et al. 2011].

An emerging area that can help in this scenario is Generative Artificial Intelligence (GAI), which involves techniques for generative modeling and deep learning to generate content from texts, graphics, audio, and videos [Jovanovic e Campbell 2022, Sengar et al. 2024]. These models have been applied to different domains, demonstrating satisfactory performance in solving complex problems, such as natural language processing, image translation, medical diagnostics, image and text fusion, among others [Sengar et al. 2024]. Given these capabilities, the use of Generative AI models has also encouraged research focused on generating music for games [Marrinan et al. 2024, Huang et al. 2023, Donahue et al. 2019].

In this context, this work presents the following research questions: "What soundtracks are generated by AI models for educational games like?" and "Which model obtains the most appropriate results in generating these soundtracks?". In this regard, an empirical study was conducted using the MusicFX models (public and refined version of MusicLM) [1], Riffusion[2] and Suno AI[3], Generating music for three different educational games from the literature, which were evaluated by computer science students and educational game developers. The work is organized in: Section 2 (Theoretical Foundation and Related Works), Section 3 (Methodology), Section 4 (Results), and Section 5 (Final Considerations).

2. Foundations and Related Works

The emergence of the GAI area has provided advances in the construction of models capable of generating audio and music [Yuan et al. 2024, Schneider et al. 2023, Huang et al. 2023].These models can be based on different architectures, such as: (i) Large Language Models (LLMs), pre-trained on large volumes of data, with tens to hundreds of billions of parameters and capable of performing numerous activities like human beings [Minaee et al. 2024, Kasneci et al. 2023]; (ii) Diffusion models – add progressive noise to the data and then learn to reverse this process to generate new samples [Cao et al. 2024]; among others. These advances have encouraged research

[1]https://labs.google/fx/tools/music-fx
[2]https://www.riffusion.com/
[3]https://suno.com/

in specific contexts, such as the generation of soundtracks for games [Liu et al. 2024, Marrinan et al. 2024], as described in the following works.

In Amaral et al. [2022], for example, a music generation architecture for games based on the Transformer model is proposed. This architecture allows music to be customized from a corpus chosen by the player. Generation occurs in instrumental layers, activated according to the player's emotional state, using the arousal-valence model for control. In addition, an LLM is employed to ensure structural coherence and adapt the music to the game context in real time.

The work of Donahue et al. [2019] presents *LakhNES*, a model based on the Transformer-XL architecture, aimed at generating multi-instrumental music in the chiptune style. Training is performed on the NES-MDB dataset, which contains soundtracks extracted from Nintendo Entertainment System (NES) games, characterized by their retro style and fixed set of four instruments. To increase the generalization capacity of the model, the authors perform pre-training using the Lakh MIDI dataset, an extensive collection of popular songs in MIDI format. With this, the model learns general musical patterns before being adjusted to the specific domain of game soundtracks. Evaluation was performed based on quantitative metrics, such as perplexity, and user tests, including Turing and musical preference tests.

The work of Marrinan et al. [2024] investigates the use of generative AI to create soundtracks and sound effects in games with user-generated content, focusing on the automatic adaptation of audio to personalized environments and objects. The study presents two prototype games in which sounds are dynamically generated through the MusicGen and AudioGen models, both based on LLMs developed by Meta. To generate the background music or effects, the system constructs textual descriptions of the content created by the player, either through color analysis of the environment or through image-to-text techniques (such as the BLIP model). The AI then generates audio from these prompts, allowing the audio to match the aesthetics and functionality of the created content.

The works in the literature demonstrate advances in the use of Generative AI models for generating soundtracks applied to games. However, some gaps are noted, which this work presents as innovation: (i) focus on generating soundtracks for three games with educational purposes; (ii) evaluation from the perspective of educational game developers; (iii) involve both LLM-based models and diffusion models, specifically MusicFX, Riffusion, and SunoAI.

3. Methods

In this work, an empirical study was conducted to investigate how the soundtracks generated by AI models for educational games sound and which model makes the best music. We used the methodology illustrated in Figure 1, which contains the stages of selecting models and games, constructing the prompt, collecting data, drawing up an evaluation form, and analyzing the data.

Model Selection: In order to select the models to be used in the research, an in-depth analysis of the literature was carried out on the Scopus database and Google Scholar. Using surveys and exploratory research, a total of 10 models were mapped and

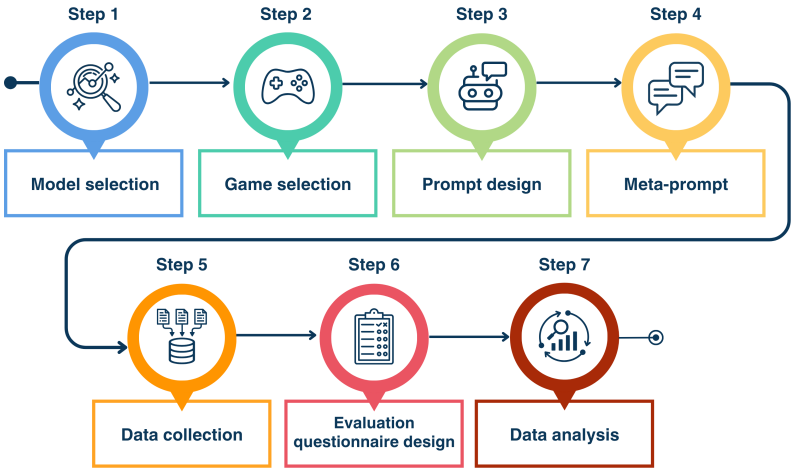

Figure 1. Study steps.

included in an online collaborative spreadsheet. Tests were then carried out on each of the models, and a benchmark was built, comparing them in terms of free access, platform (online or software), audio quality, character limits, export types, and generation/time limits. Based on this, the models defined were: (i) MusicFX, developed by Google DeepMind, which generates music from natural language texts (prompts) and uses LLMs to interpret the inputs; (ii) Riffusion, which uses diffusion models to generate spectogram images from text, which are then converted into audio; and (iii) Suno AI, which uses LLMs to interpret the prompts and generate lyrics with sound synthesis models, and is one of the most popular and advanced tools for generating audio/music. These templates were chosen mainly because: (i) they were accessible online. Others were only accessible on HuggingFace and required a higher level of knowledge about infrastructure and environment configuration; (ii) character limitation, since some models did not allow long prompts; and (iii) free use, considering that some models could only be used partially or completely for a fee.

Game selection: We went on to select the educational games for which the models would generate soundtracks. This process considered the following requirements: (i) educational games in Portuguese, being in the authors' native language; (ii) simple mechanics, facilitating the construction of prompts; (iii) familiarity and experience of the authors with the games; and (iv) publication available in scientific databases to facilitate the collection of information. The games selected were: (i) "A Viagem" (The trip) [Pires et al. 2018], (ii) "Automigos" [Honda et al. 2023] and (iii) "Cadê minha pizza?"(Where's my pizza?) [Honda et al. 2022]. Game (i) focuses on geography, in which the player has to steer a helicopter that flies over the map of Brazil to the corresponding state, city, or coat of arms. Game (ii) aims to exercise the content of Deterministic Finite Automata, in which you have to build food systems from platforms, conveyors, and space beings (automigos). In turn, game (iii) focuses on implicit learning of the minimum path, where the player must deliver pizzas in the city that considers the best route and the most efficient strategy for allocating delivery men.

Prompt design: this stage is divided into two parts: (i) defining the game elements to make up the prompt and (ii) building the prompts to be sent to the models. In (i), it was

necessary to define which game elements would be sent, so that the model would be able to abstract the game and propose the most appropriate music. Therefore, by analyzing the game and learning elements that make up the Educational Game Design Document (EGDD) [Pires 2021], we listed: name, story, gameplay, learning mechanics, and setting of the phases - this data is considered to be the basis for an educational game. It was decided not to add any more elements because of the models' hallucinations, which could generate meaningless and/or inaccurate answers [IBM 2023] – and to keep the prompt short, since the model can ignore information when it is long. Thus, for each educational game selected, the elements were located and described in a text file. Next, for item (ii), ChatGPT was used to help build the prompt: the information on the educational games from the elements was sent, asking it to generate a prompt for a model to generate soundtracks. As a result, the model generated a prompt for each game, which can be consulted at the following link[4]

Meta-prompt: in order to add an extra layer of refinement to the prompts generated by ChatGPT, meta-prompt was used. This is a prompt engineering technique that consists of guiding a model to analyze a prompt, send feedback and update it, resulting in an improved prompt [Ye et al. 2023]. The model chosen for the meta-prompt was DeepSeek, considering its capabilities and relevance in Generative AI. In this way, the prompts created by ChatGPT were sent to DeepSeek, asking it to limit them to 200-220 characters - the maximum allowed by Suno AI and with the aim of standardizing the prompts in a compact way. The prompts were then updated, with the main changes referring to: clarity and objectivity, emphasis on technical musical details, softening or omitting educational aspects, reducing descriptions of game mechanics, summarizing the environment, etc. The resulting prompts can be consulted at the link[5].

Data collection: then the prompts were sent to each model, which generated two songs per game. Considering that the study aimed to evaluate only one song per model for each game and that the evaluation form could be lengthy, causing tester bias, it was decided to evaluate only the first song generated by each model – totaling 9 soundtracks. As a result, each model generated three songs for the games "A Viagem", "Automigos" and "Cadê minha Pizza?", which can be accessed at the link[6].

Evaluation questionnaire design: The objective of this introductory study is to analyze the possibilities of using Generative AI in the creation of music for games. Considering that this is the first study in the literature at the intersection of AI models, soundtracks, and educational games, there is no validated instrument or questionnaire for this type of evaluation. Therefore, two study authors designed a form to evaluate the music generated by the models from the perspective of computer science students and educational game developers. The form is organized into three pillars: (i) personal information – with questions about name, age, gender, and academic data; (ii) specific sections for each educational game – containing the music generated by the three models and questions to evaluate them, two of which are quantitative on a 5-point Likert scale and two of which are qualitative. In addition, at the end of each section, there is a

[4]https://drive.google.com/file/d/17hz8UA-ssHutSuNzYt-kWog2fa0W8yhc/view
[5]https://drive.google.com/file/d/1-Vh4AdViKQPjGUIAPtX1XD8s5XCVVr5o/view
[6]https://drive.google.com/drive/folders/1sqfjQu86Nr-gfGb1FgFzZTKfBKhJ1HE6?usp=sharing

quantitative question to determine which model generated the most appropriate music[7] and a qualitative question to collect the testers' opinions on the music generated for the specific game; (iii) general evaluation section – with links to access the generated soundtracks, three quantitative questions about the testers' opinion on the generation of music by AIs, concerning the difference between the songs and whether they used Generative AI to generate music for educational games, along with qualitative questions for justification. Questions about "What did you like most" and "What did you like least" about the AI-generated music were also included. It should be noted that the focus of this research is on the perspective of learning designers, to include sound designers in the future to develop a new evaluation form and to evaluate the music in technical terms.

Data analysis: To evaluate the songs generated by the AIs, computer science students and teachers were invited to fill out the form, emphasizing those who already had experience building educational games. As a result, 22 people participated in the study, of which: (i) 59% were male and 41% were female; (ii) aged between 19 and 27, except teachers aged 34 and 42; (iii) most were currently pursuing a degree (68.2%), were affiliated with institutions in the state of [omitted for review] (86.3%), and were enrolled in computer science courses (95.5%), either undergraduate or graduate; (iv) 72.7% of testers with less than 1 to 3 years of experience in building educational games, of which 68.2% participated in the development of these artifacts; and (v) only 9.1% composed soundtracks for educational games, listing software such as Audacity, Encore, and Bandlab. Thus, participants used Google Forms to listen to the songs and evaluate them, and their responses were collected and recorded in a collaborative online spreadsheet. Google Colab was used to generate a boxplot graph to compare the evaluations and a descriptive statistics table for a thorough analysis of the results described in the following section.

4. Results and discussions

Figure 2 contains a boxplot graph illustrating the testers' evaluations of the songs generated by each model. To facilitate individual analysis, the scores for quantitative questions Q1 (consider the song good for use in the game) and Q2 (song appropriate to the style and theme of the game) were added together. Therefore, while the X-axis refers to the games, the Y-axis refers to the accumulated scores for the questions (Q1 and Q2), which can range from 2 (both questions scored 1) to 10 (questions scored the maximum).

Overall, the graph shows that the ratings for each model were similar in the three games evaluated. In "Cadê minha pizza?", Suno AI and MusicFX had similar scores, with a high dispersion. On the other hand, Riffusion showed a lower dispersion with scores more concentrated between 8 and 9. In "Automigos", Suno AI was the lowest rated, varying mostly between scores of 6 and 8. MusicFX and Riffusion had the same scores. As for the game "A Viagem", Suno AI stood out from the rest with higher scores and less dispersion. In this case, MusicFX and Riffusion had similar scores. Therefore, from the testers' perspectives, the performance of each model was similar. Below is an individual analysis of the songs generated for each educational game.

Songs for "Cadê minha pizza?": 45.5% of testers considered the music generated by MusicFX to be good. On the other hand, the same percentage disagreed

[7]The names of the models were anonymized to avoid bias.

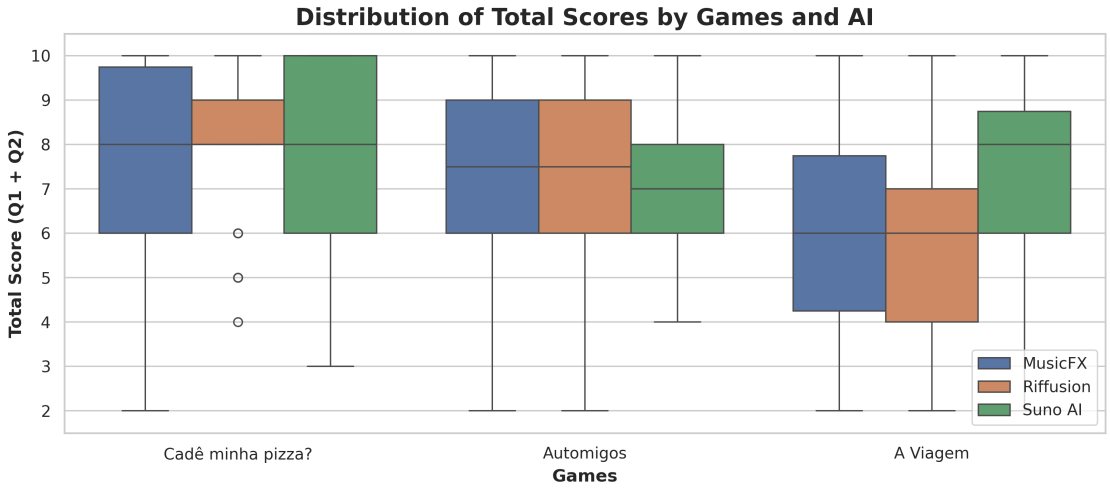

Figure 2. Boxplot with music ratings for each model.

that the music seemed appropriate for the style and theme of the game. When asked why, testers pointed out: generic, elevator music, artificial, cool and engaging, fun, relaxing, arrhythmic, etc. As for the music generated by Riffusion: 40.9% agreed that the music is good for the game, and the same percentage disagreed. 45.4% found it unsuitable for the game, and 36.4% remained neutral. Among the comments, the main ones refer to the slowness, which resembles the jazz style and reminds one of elevator music. As for Suno AI, 59.1% considered the music suitable for the style and theme of the game, obtaining the best evaluation among the models. Some comments highlight it as the one that best suited the game, being more lively. Regarding which AI generated the most appropriate music for this game, 54.5% indicated Suno AI and 31.8% indicated Riffusion. Responding to the qualitative question, most of the testers' answers pointed out that the songs were cool and good, highlighting that some specific ones were not generic and did not fit the context of the game.

Songs for "Automigos": 63.7% considered the music generated by MusicFX to be suitable for the style and theme of the game (68.2%). Most testers' comments indicate that it is compatible and fits the space theme. Rifussion's evaluations were also positive, with 77.3% of testers considering the music good for use in the game and appropriate for its context. Testers reported that it evokes a futuristic theme, bringing a more mysterious and tense aspect, with one tester commenting that "The music made me want to play." Regarding Suno AI, 63.7% and 63.6% of testers agreed that the music was good and appropriate for the game, respectively, describing it as calm and fitting for the theme. On the other hand, some testers did not like it, especially the beats of the music. In this game, the AIs that generated the most suitable music were also Suno AI (40.9%) and Riffusion (36.4%). Testers said they liked the music, which "brought more variety" and was more fitting than the music generated in the previous game – a sentiment corroborated by the higher ratings.

Songs for "A Viagem": Regarding MusicFX, 50% consider the music to be good and 63.7% consider it to be appropriate for the style/theme of the game. Testers' comments say that the music reminds them of the beach and Rio de Janeiro, highlighting

styles such as bossa nova and samba. As for Riffusion, agreement on both questions was 63.6% and 54.6%, respectively. Testers found the music to be good and more lively, while some disagreed about its suitability for the game. Evaluations for Suno AI were mixed on the first question: 40.9% neutral and 45.4% in agreement. On the second question, 54.6% agreed. The testers' comments called it "elevator music," with a slower rhythm similar to MusicFX music. Viagem was the game where testers most disagreed when listing the AI that generated the most appropriate music, with 36.4% for MusicFX, 36.4% for Riffusion, and 27.3% for Suno AI. For this game, testers felt that the songs referred to Rio de Janeiro; some liked them, while others did not.

To corroborate the individual analyses of each game, Table 1 was constructed, containing descriptive statistics (Mean – M, Median – Md, and Standard Deviation – SD) for the questions evaluated for each game-model combination. It can be observed that the game "A Viagem" has the lowest scores – mainly Riffusion, for which the models presented songs that were not so appropriate. This factor may be related to the game dealing with Brazil as a whole, making it challenging to represent the entire territory and its diversity in a single song. The best songs were from the game "Cadê minha Pizza?" (Where's my Pizza?), which received satisfactory evaluations in all three models. "Automigos" also received a good evaluation and good receptivity of the songs in the opinion of the testers, with the exception of Suno AI, which performed below expectations. Regarding the models, there was no preference, but Suno AI (40.9%) and Riffusion (34.8%) were the ones that testers most often pointed out as adequate.

Table 1. Descriptive statistics of the questions evaluated about the songs.

Game	Model	Q1 M	Q1 Md	Q1 SD	Q2 M	Q2 Md	Q2 SD
A Viagem	MusicFX	3.09	3.00	1.23	2.82	3.00	1.14
A Viagem	Riffusion	2.95	3.00	1.36	2.68	3.00	1.13
A Viagem	Suno AI	3.64	4.00	1.09	3.55	4.00	1.41
Automigos	MusicFX	3.64	3.50	1.18	3.59	4.00	1.30
Automigos	Riffusion	3.55	4.00	1.06	3.64	4.00	1.29
Automigos	Suno AI	3.55	3.00	1.01	3.59	4.00	1.14
Cadê minha pizza?	MusicFX	3.68	4.00	1.21	3.91	4.00	1.31
Cadê minha pizza?	Riffusion	3.91	4.00	0.75	3.82	4.00	0.91
Cadê minha pizza?	Suno AI	3.82	4.00	1.14	3.91	4.00	1.15

The analysis of quantitative and qualitative issues suggests a high disparity in the testers' evaluations. This point can be observed: (i) in the testers' comments, pointing out negative and positive aspects for each song – sometimes at opposite extremes; (ii) in the boxplot, where most evaluations show a wide dispersion of data; and (iii) in the descriptive statistics, where no question had an average of 4 or higher. These points may suggest: (i) the difficulty of testers in evaluating the songs, most of whom (91%) have no experience in composing soundtracks for educational games; (ii) diverse opinions, since liking a song is subjective and depends on a number of factors; (iii) the diversity of the models, which have different knowledge bases and can generate completely different results; (iv) the complexity of generating music by the models, which need to abstract

the game to understand it and return corresponding results; and (v) the elaborate prompts, which may not have been constructed most appropriately.

Regarding using these models, 81.9% of testers indicated that they generate music suitable for educational games. The reasons for this include the appropriate atmosphere that these songs created, a sense of tranquility, practical application, and harmony. On the other hand, the incompatibility of some songs was also highlighted, with some being considered generic. Most testers (68.2%) reported noticing differences between the music from the different models, such as: number of instruments, volume, rhythm, speed, etc. 81.8% would use Generative AI to generate music for educational games, mainly because of: ease, customization, context suitability, and creativity. Regarding what they liked most about the songs, the following stand out: versatility, variety of sounds and instruments, overall sound, etc. On the other hand, the testers did not like: generic songs, strong beats, repetition, continuity (songs that end without a closing), etc.

Thus, answering the research questions: the models generate distinct soundtracks for educational games, differing in several aspects (rhythm, speed, style, etc.). No model is more suitable than the others evaluated, but Suno AI and Riffusion were identified as the ones that generated the most compatible music for the games. From the testers' perspective, using these models is promising, resulting in appropriate and interesting music. However, some negative points were also pointed out, expressing that the music still needs further refinement.

5. Conclusions

This work conducted an empirical study to answer the research questions: "What soundtracks are generated by AI models for educational games?" and "Which model obtains the most appropriate results in generating these soundtracks?". The adopted methodology followed a series of steps, including: (i) model selection – which performed a benchmark and chose the MusicFX, Riffusion and Suno AI models; (ii) Game selection – which chose three games from literature with simple mechanics and distinct contexts; (iii) Prompt design – which involved selecting the attributes of educational games that would be sent to the models and creating the prompt, using ChatGPT; (iv) Meta-prompt – DeepSeek was used to refine the created prompts; (v) Data collection – the prompts were sent to the models, resulting in 9 distinct soundtracks, where each model generated three songs for each game; (vi) Design of the evaluation questionnaire – design of a form to collect participants' profiles, music ratings and general perceptions about the use of AI to generate soundtracks for games; and (vii) Data analysis – administering the tests and collecting the data, involving students and teachers of computing, most of whom have experience with the creation of educational games.

The results were positive, indicating that testers liked the music generated by the models. In response to the research questions, the generated soundtracks are distinct in terms of number of instruments, volume, rhythm, speed, etc., and, despite similar evaluations for the three models, testers preferred the music generated by Suno AI and Riffusion. The game that had the most suitable music from the testers' perspective was "Cadê minha pizza?", while the songs for the game "A Viagem" They were the least well-rated. Positive aspects of the songs include versatility, variety of sounds and instruments, compatibility with games, etc. On the other hand, the following stand out as negative

aspects: generic songs, strong beats, repetition, among others. Most testers indicated that they would use these models to generate songs for educational games, highlighting their potential for this type of activity. However, a series of refinements is still needed in the songs to make them suitable for games. The contributions of this work demonstrate the ability of Generative AI models to compose songs for educational games, and can be an alternative to the lack of free materials, qualified professionals, and expertise in the area.

The limitations of this research include: (i) the prompts constructed, which may have influenced the quality of the music generated; (ii) the perspective, since the focus of this introductory study was only on educational game designers – with the intention of expanding it to music professionals in the future; (iii) the subjectivity in the evaluations, since the perception of music quality depends on the participants' opinion and personal taste. This may have varied the responses and made it difficult to generalize the results, not necessarily indicating that the best-rated models are the most suitable for the activity; (iv) the evaluation form, which may not have been sufficient to evaluate the music accurately; and (v) the diversity of the testers, most of whom do not have experience in composing soundtracks for games. These items will be minimized in future research, the objective of which in this work was an initial analysis of the use of these models to generate music for educational games, from the perspective of those responsible for creating these objects – who are often students, do not have expertise in this context and are unable to find free/non-copyrighted materials online.

As future work, we intend to: (i) expand the scope of the study, analyzing other models; (ii) bring the perspective of sound designers, both in the evaluation and in the construction of the form; (iii) increase the sample of testers, disseminating to a larger number of universities; and (iv) use a more rigorous process of validating the prompts.

References

Alves, L. R. G. e. o. (2010). *Jogos eletrônicos e educações: como os games podem ajudar no processo de aprendizagem*. Editora Papirus.

Amaral, G., Baffa, A., Briot, J.-P., Feijó, B., e Furtado, A. (2022). An adaptive music generation architecture for games based on the deep learning transformer model. In *2022 21st Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)*, pages 1–6. IEEE.

Battistella, P., von Wangenheim, C., e Fernandes, J. M. (2014). Como jogos educacionais são desenvolvidos? uma revisão sistemática da literatura. In *Anais do XXII Workshop sobre Educação em Computação*, pages 159–168. SBC.

Cao, H., Tan, C., Gao, Z., Xu, Y., Chen, G., Heng, P.-A., e Li, S. Z. (2024). A survey on generative diffusion models. *IEEE Transactions on Knowledge and Data Engineering*.

Caroux, L. e Pujol, M. (2024). Player enjoyment in video games: A systematic review and meta-analysis of the effects of game design choices. *International Journal of Human–Computer Interaction*, 40(16):4227–4238.

Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. Harper & Row, New York.

Donahue, C., Mao, H. H., Li, Y. E., Cottrell, G. W., e McAuley, J. (2019). Lakhnes: Improving multi-instrumental music generation with cross-domain pre-training. *arXiv preprint arXiv:1907.04868*.

Evans, R. (2018). An investigation into the effect of music on immersion in video games.

Fernandes, R. (2010). *Brincar na pré-escola: o jogo como proposta de trabalho pedagógico*. Editora Vozes.

Genesio, N. O. S., de Oliveira, A. M., Oliveira, E. W., e Valle, P. H. D. (2024). Panorama de estudos sobre jogos educacionais digitais em educação em computação. In *Workshop sobre Educação em Computação (WEI)*, pages 737–749. SBC.

Guedes, U. e Mello, A. (2024). Conflito de estrelas: Adaptando um jogo educativo analógico para o meio digital. In *Anais do XXXV Simpósio Brasileiro de Informática na Educação*, pages 870–883, Porto Alegre, RS, Brasil. SBC.

Honda, F., Pires, F., Pessoa, M., e Maia, J. (2022). Cadê minha pizza? um jogo para exercitar matemática e pensamento computacional através de grafos. In *Simpósio Brasileiro de Jogos e Entretenimento Digital (SBGames)*, pages 876–885. SBC.

Honda, F., Pires, F., Pessoa, M., e Oliveira, E. H. (2023). Automigos: learning design para ludificação de autômatos finitos determinísticos. In *Workshop sobre Educação em Computação (WEI)*, pages 545–556. SBC.

Huang, Q., Park, D. S., Wang, T., Denk, T. I., Ly, A., Chen, N., Zhang, Z., Zhang, Z., Yu, J., Frank, C., et al. (2023). Noise2music: Text-conditioned music generation with diffusion models. *arXiv preprint arXiv:2302.03917*.

IBM (2023). Ai hallucinations. https://www.ibm.com/topics/ai-hallucinations. Accessed: 2025-04-20.

Jovanovic, M. e Campbell, M. (2022). Generative artificial intelligence: Trends and prospects. *Computer*, 55(10):107–112.

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., et al. (2023). Chatgpt for good? on opportunities and challenges of large language models for education. *Learning and individual differences*, 103:102274.

Levy, L. (2015). The effects of background music on video game play performance, behavior, and experience in extraverts and introverts. Accessed on: April 21, 2025.

Lima, W. G. d., de Freitas, A. R. R., e de Lima, T. F. M. (2021). Design e desenvolvimento de um jogo sério para educaçao musical. In *Simpósio Brasileiro de Jogos e Entretenimento Digital (SBGames)*, pages 417–426. SBC.

Linek, S. B., Marte, B., e Albert, D. (2011). Background music in educational games: Motivational appeal and cognitive impact. In Ifenthaler, D., Eseryel, D. S., e Ge, X., editors, *Computer Games and Instruction*, pages 259–270. IGI Global.

Liu, H., Wang, Z., Hong, H., Feng, Y., Yu, J., Diao, H., Xu, Y., e Zhang, K. (2024). Metabgm: Dynamic soundtrack transformation for continuous multi-scene experiences with ambient awareness and personalization. *arXiv preprint arXiv:2409.03844*.

Marrinan, T., Akram, P., Gurmessa, O., e Shishkin, A. (2024). Leveraging ai to generate audio for user-generated content in video games. *arXiv preprint arXiv:2404.17018*.

Minaee, S., Mikolov, T., Nikzad, N., Chenaghlu, M., Socher, R., Amatriain, X., e Gao, J. (2024). Large language models: A survey. *arXiv preprint arXiv:2402.06196*.

Pires, F. G. d. S. (2021). Thinkted lab, um caso de aprendizagem criativa em computação no nível superior.

Pires, F. G. d. S., Queroga, J. d. S., Pessoa, M. S. P., e Melo, R. (2018). A viagem: Um jogo para explorar o território brasileiro. In *Brazilian Symposium on Computers in Education (Simposio Brasileiro de Informatica na Educacao-SBIE)*, volume 29, page 705.

Plass, J. L., Homer, B. D., e Kinzer, C. K. (2015). Foundations of game-based learning. *Educational Psychologist*, 50(4):258–283.

Plass, J. L. e Kaplan, U. (2016). Emotional design in digital media for learning. In *Emotions, technology, design, and learning*, pages 131–161. Elsevier.

Schneider, F., Kamal, O., Jin, Z., e Schölkopf, B. (2023). Mo\^ usai: Text-to-music generation with long-context latent diffusion. *arXiv preprint arXiv:2301.11757*.

Sengar, S. S., Hasan, A. B., Kumar, S., e Carroll, F. (2024). Generative artificial intelligence: a systematic review and applications. *Multimedia Tools and Applications*, pages 1–40.

Silva, P. d. S. M. d. (2025). Playful approach: stimuli to mental learning processes in early years. Accessed on: April 21, 2025.

Ye, Q., Axmed, M., Pryzant, R., e Khani, F. (2023). Prompt engineering a prompt engineer. *arXiv preprint arXiv:2311.05661*.

Yuan, R., Lin, H., Wang, Y., Tian, Z., Wu, S., Shen, T., Zhang, G., Wu, Y., Liu, C., Zhou, Z., et al. (2024). Chatmusician: Understanding and generating music intrinsically with llm. *arXiv preprint arXiv:2402.16153*.