

Original scientific paper

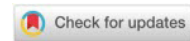
UDC:  
37.091.3.:78

Received: February 23, 2025.

Revised: April 15, 2025.

Accepted: April 16, 2025.

 10.23947/2334-8496-2025-13-1-75-81



# Cross-modal Priming of a Music Education Event in a Digital Environment

Diana Petkova\* 

<sup>1</sup>Trakia University, Faculty of Education, Stara Zagora, Bulgaria, e-mail: [diana.petkova@uni-plovdiv.bg](mailto:diana.petkova@uni-plovdiv.bg)

**Abstract:** This study aims to explore the potential of the digital environment for implementing a multimodal approach in music education. The effectiveness of information received through a combination of sensory stimuli demonstrates a higher coefficient of educational efficiency and is examined as cross-modal priming. Digital technologies: including specialized and educational software, virtual instruments, and artificial intelligence (AI), transform the music education experience into an accessible resource for individuals with limited musical abilities or non-professional knowledge in the field of art. This justifies their consideration as tools for general music education. The study presents a model for applying specialized music software in the perception of a musical piece by students (aged 12–13), as well as a methodological framework among university students, future kindergarten and primary school teachers. The findings indicate improved musical-cognitive outcomes and a high evaluation of specialized software as a didactic tool among university students. Additionally, the study discusses the role of AI in music education.

**Keywords:** *cross-modal, priming, music, education, specialized software, AI.*

## Introduction

In the era of digital transformation, education and the arts are undergoing significant changes, creating opportunities to optimize aesthetic education through interactivity and accessibility. The correlation *music-education-digital technologies* expands the practical and applied aspects of music learning, providing a rich toolkit for artistic interaction (Falkner (1995); Pastarmadzhiev, 2021; Bačlija and Mičija, 2022; Rexhepi, Breznica and Rexhepi, 2024). The complex of sensory stimulus in perception of an object in time-space behavior, organizes the cognitive activity. In general educational musical interaction, conditions for functioning of such a complex of stimulus, is provided in the digital environment. Digital technologies visualize interactive musical text, individualizing the interaction, providing tactile activity.

A study based on the VARK model (Fleming and Baume, 2006) was conducted by Mishra, 2007, analyzing educational strategies for learning to play a musical instrument. The process of performing a musical piece involves reading notation, motor cognition related to the instrument's sound production, systematic praxis for memorization, and artistic interpretation.

Traditional forms of musical engagement are realized through three primary activities: perception, performance, and composition. These activities are interrelated, with perception forming the foundation of music production. This underpins extensive research on multisensory reactivity to sonic structures. Even since 1994 Robert Dunn (Mishra, 2007, p.5) assumed that the musical perception responds to sound, motorial, and visual stimuli, regardless of individual modality. His study reports the following findings:

- 19% respond only to auditory stimuli;
- 50% to both visual and auditory stimuli;
- 6% to auditory and motor stimuli;
- 25% to a combination of all.

<sup>1</sup>Corresponding author: [diana.petkova@uni-plovdiv.bg](mailto:diana.petkova@uni-plovdiv.bg)



Unlike Dunn's study, Falkner's findings indicate that kinesthetic learning plays the most significant role - 50%, compared to 22%- auditory and 28%- visual. These differences are substantial, as interpretative musicianship belongs to the realm of performance, even though perception remains a reflective process. Furthermore, Jennifer Mishra (Mishra, 2007) explores the relationship between perception and memory encoding of auditory information among instrumental musicians. Initially, she analyzes 121 studies related to the process of musical memorization. In 60% of them, she identifies a scientific focus on auditory, visual, and kinesthetic styles, while 51% - emphasize their combination to optimize educational outcomes. These studies highlight the complexity of musical events and the unique nature of musical communication within the framework of multisensory interactions, functioning as cross-modal priming.

Music is structured within strict algorithmic dependencies, defined by pitch and rhythmic relationships. It carries semantic markers with emotional and conceptual projections that shape the perception process. This distinctive dual structure positions music as a system of modes within the broader field of information transfer. The integration of musical activity into a digital environment, driven by the pursuit of singularity, establishes an accessible model that seamlessly incorporates both educational and specialized music software, as well as the increasingly advanced frameworks of artificial intelligence (AI). The synergy between learning styles, embedded within various digital productions, influences perception by delivering information as a complex interplay of sensory stimuli within a given time frame. The study in this paper is dedicated to the crossmodally in perception of sonorous flow.

## Materials and Methods

Learning through computing technologies eliminates intellectual passivity. This can be explained by the way the system itself is structured, based on a database and models for constructing cognitive activity. Computing technologies possess a set of capabilities:

- Diverse and large volumes of information
- Goal setting and behavior planning following the sequence: goal – plan – actions
- Partial selection of the necessary database
- Utilizing available knowledge and reasoning results corresponding to set goals
- Justifying decisions in practice and realizing them in achieved results
- Conducting reflection – evaluation of knowledge and actions
- Stimulating cognitive curiosity – encouraging learners to ask questions independently
- Data monitoring
- Supporting the rationalization of ideas and striving for conceptual clarification
- Interpreting a comprehensive picture of the subject of cognitive activity, integrating knowledge relevant to the set goal
- Correcting and adapting the final result according to changes in the cognitive situation
- The cognitive activity carried out through them is organized into resource modules:
  - Data and knowledge representation
  - Reasoning and computation
  - Multisensory communication – ensuring accessibility and convenience in interaction with the computer

One of the most innovative approaches in this field is cross-modal priming—a method that employs multisensory stimulation to enhance learning and memory retention. When applied to music education events in a digital environment, cross-modal priming can play a pivotal role. Multisensory communication in musical art involves a stimulus (prime), that aligns with the recipient's cognitive predispositions. Subsequently, this stimulus in one modality (visual, auditory, tactile, or semantic) influences the processing or perception of another stimulus in a different modality. This influence may manifest as a faster response, easier recognition, or improved retention of the second stimulus. The network of stimuli that conditions this response is defined as cross-modal priming.

Figure 1 presents a universal model of cross-modal priming in an educational environment, using specialized software as a tool.

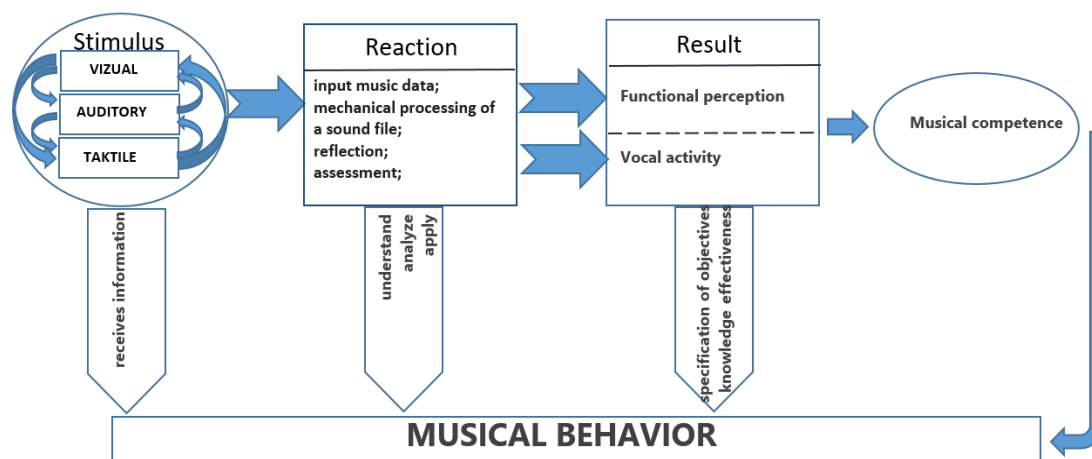


Figure 1. Correlation cross-modal priming-musical behavior

An original methodology for perceiving a musical piece through specialized music software (Sibelius and FL Studio), developed for students in the primary stage of secondary school, includes the following activities:

1. Notation of the theme in interactive mode (each entered written symbol is played back)
2. Reflection on the recording – listening, analyzing, and correcting the notated theme
3. Listening to the musical piece and auditory recognition of the notated theme
4. Creating an audio file by extracting the notated theme from the musical piece
5. Audio processing and arranging the theme from the piece using the respective software tools

Each stage involves visual, auditory, and tactile stimuli. Method for performance of a song by students, the future teachers also include five stages:

1. Notation of the song using the MuseScore software
2. Listening to the notated recording
3. Learning the song
4. Planning activities related to artistic performance and making corrections to the score
5. Artistic performance of the song

### *Research question goals and objectives*

Investigation is based on the role of digital environment for improvement of cognitive-educational musical activity. The goals of the current investigation are to be presenting models for musical educational activities in digital environment and proving the effectiveness of the cognitive activity, based on cross-modal technology for perception of sonorous information flow.

The transfer of a music education event to a digital environment optimizes performance and provides a complex set of sensory activities in the perception process. The application of music software in a music education event offers an innovative perception model that corresponds to both listening to music and performing it. Learning with the help of music computing technologies redistributes the involvement of visual and auditory analyzers by incorporating the mechanical process of data input. The transition from sensation and perception to thinking is mediated by mental representations, depending on their type (visual, auditory, motor) and the content of the sonorous product. The current investigation involves the following computing capabilities (see, chapter Materials and methods): (1) data and knowledge representation; (2) reasoning and computation which I assume to become my independent variables and is defined, based on the intelligence of the computing system—the ability to manage new knowledge through the use of available database information as reasoning, defined to be independent variables, expressed through their corresponding indicators:

1. Writing of a music text with the help of a notation software (Sibelius 6 and MuseScore 3.6)
2. Listening to the notated recording

3. Memorizing of a musical composition
4. Analysis of a sonorial structure
5. Development of a creativity product

The management of the information flow is based on *deductive reasoning*, as the algorithm of activity follows the principle of transferring correctly embedded information to conclusions or inferences. Along with the database, computing system software includes rules and axioms that guide and revise the results of the activity.

*Multisensory communication*, (see, chapter Materials and methods): supports the development of musical hearing and musical behavior through logical-perceptual forms and emotional-reactive expressions in actions and appears to be the dependent variable through its indicators of statistical significance. It is “effective transformation in the course of multimodal education fulfilled in experience” (Dermendzhieva and Tsankov, 2022, p.168). Understanding musical expressive means is facilitated by the cross-modality of priming. When working in a music software environment, the stimulus combines visual, auditory, and motor sensibilities.

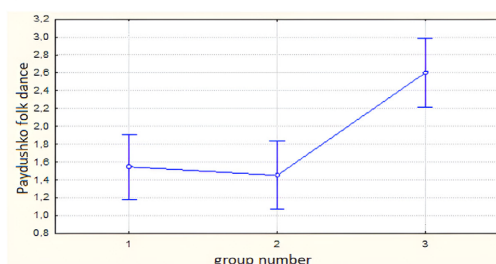
The controlled variable of my investigation is the selected software MuseScore, version.3.6.1 for notation in the performance; Sibelius 6 for perception; FLStudio 9.6 for audio-processing and arranging.

### Experimental data

The experiment perception of musical work was conducted with two control groups and one experimental group (students aged 12–13). The obtained results (Table 1) for auditory recognition of a musical piece, at a 5% significance level ( $F = 11.259$ ,  $p < 0.0007$ ), indicate that the experimental group achieved higher performance, with the statistical difference being significant (Petkova, 2023, p.108).

**Table 1.** Statistical analysis of data and study performance

ANOVA	Sum of Squares	df	Mean Square	F	Sig.
Result "Paydushko hora"					
Between Groups	16.469	2	8.245	11.259,000	
Within Groups	43.205	59	.732		
Total	59.694	61			



The model for vocal performance was conducted with students, future teachers in kindergartens and the primary stage of general education schools (Petkova, 2022). The software used was MuseScore, with the goal of mastering a model for organizing vocal activities and learning the notated song as an object of music-pedagogical communication. The developed methodology aligns with regulatory documents that assign arts classes to all certified teachers: kindergarten teachers (ages 3–7) and primary school teachers (ages 7–11). In the current study, the presented indicators, (see chapter Research question goals and objectives) as input characteristics were applied.

A total of 102 students participated in the study. Among them, 70 students (68.63%) wrote the notes *without errors*. Ten students (9.80%) made *one mistake* each, and 22 students (21.57%) made *more than one mistake*. The calculation of the average number of mistakes is as follows:

$$(10 \times 1) + (22 \times 2) = 10 + 44 = 54 \text{ mistakes}; \text{ thus, the average number of mistakes per student is: } 54 / 102 \approx 0.53 \text{ mistakes per student.}$$

This result indicates that each student made fewer than one mistake on average, which is a good indicator of success.

## Results and Discussions

The initial stimuli are maintained through necessary actions to reveal the educational value of the subject of the educational event: either an instrumental piece or a song. The response is algorithmically

controlled by the training parameters embedded in computer programs, meaning that errors can only occur in the pitch of the musical sound. This characteristic requires systematic auditory reflection, while simultaneously involving the input of musical notation, which is linked to observation and its transfer into an interactive digital environment. The outcome in both models is musical activity— perception and performance. Cross-modal priming also ensures the cross-modality of musical activities. During the perception phase, auditory representations of the theme from the instrumental piece are formed. Auditory observation is supported by a visualization of the sound structure, which also enables solfège (performance). The result of notating a song is its performance, which combines both perception and execution. This standardizes the process of developing skills necessary for engaging in musical activities.

This structured approach integrates into an operational model that transforms into musical competence. Musical competence is equivalent to musical behavior, as all independently initiated forms of musical activity, resulting from the internalization of the educational event, stem from competencies acquired through cross-modal priming. A conducted survey on the attitudes toward applying specialized software in music education in kindergartens and grades 1–4 found that 84% of students indicated they would use digital resources to support educational activities—76% for perception activities and 84% for song performance.

Artificial Intelligence (AI) emerges as a natural extension of the development of digital technologies. An increasing number of algorithms are embedded in technology to ease and optimize human activity. A new, rapidly growing industry is being built around various AI services offered worldwide for the creation, processing, and analysis of music.

### *Future in cognitive studies*

Virtual models of musical activity are supported by specialized educational software products and AI. Musical communication is built on the interaction between sensory stimuli and the analysis of the sonorous structure in terms of pitch, meter-rhythm, timbre, and dynamic responsiveness, synchronized with the semantic value of the musical piece for the recipient.

Platforms such as OpenAI, ChatGPT, and ChatGPT Plus are based on GPT models available at [chat.openai.com](https://chat.openai.com) and generate information based on strategically formulated questions (Holster, J. 2024). Applications have been developed that analyze musical structures by recognizing audio pieces (Shazam), support sound environments to enhance intellectual activity (Brain FM), or analyze users' aesthetic preferences (Spotify). A major focus of many projects is exploring AI's ability to autonomously generate music or collaborate with composers. An example is the European Research Council-funded project *Flow Machines* by Sony CSL (Sony CSL, 1993–2025). *Google Magenta* (Magenta.js, 2023) is a research project launched by Google that investigates the role of machine learning in creating synthetic music products. Algorithms for generating songs have been developed. *Aiva*, created by AIVA Technologies, has an AI script capable of composing “emotional” soundtracks for advertisements, video games, or films, as well as creating variations of existing songs. The first generator to create a musical structure based on emotional analysis is *Melodrive* (Melodrive, 2025), while *IBM Watson Beat* is a project with the ability to harmonize melodies.

When analyzing the digital environment, it is important to consider that AI is rapidly gaining popularity. It learns communication models through text and integrates into specialized music software. The integrative potential of AI can be explored in the fields of text processing and sound generation using code in various programming languages. From the perspective of the syntax of a musical event, this shifts the focus to an alternative symbolic system, distinct from traditional musical notation. In Figure 2, an example code snippet is presented for generating a musical tone with a frequency of 441 Hz corresponding to the note “a<sup>1</sup>” with a duration equivalent to a quarter note, alongside a notation of the same event in the specialized scoring software *MuseScore*.

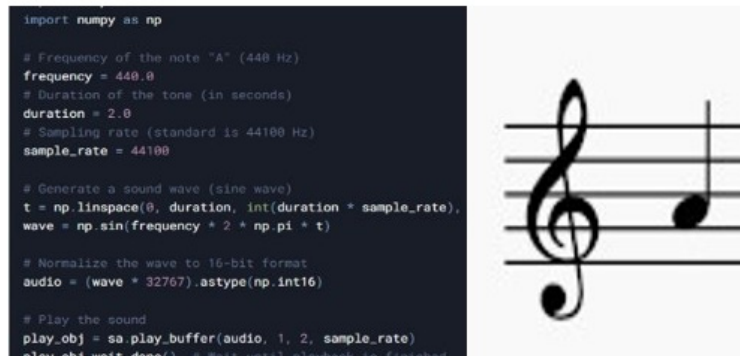


Figure 2. Tone code „a1“

The primary approach to developing an AI skill is the definition of a semantic network, which enables “the process of constructing, linking, and representing generalized concepts” (Kordon, 2023, p.36). At this stage, in a digital environment, crossmodal priming operates more effectively within specialized music software. For example, when entering the note “a1” on an interactive score sheet in notation software (such as MuseScore, Sibelius, etc.), the user hears, sees, acts, and understands (learns) the meaning of this symbol. When working with *AIVA* or *SUNO* (AI for music creation), the user defines the semantic model of the product by specifying the musical style, content, and mood. Once the *sonic model* is generated, some of its parameters can be edited.

The educational benefits of this music-technology activity manifest in a remote projection of the creative process, based on emotional needs. Here, the user does not participate in the spatiotemporal organization of musical activity but acts as a mere consumer (Grigorova, 2018) of a pre-designed model. Even though the user initiates the creation process, they remain in the role of a listener, with only “the personal function of perception” being active (Baleva, p. 66, 2010). The result is a product with emotional, stylistic, and textual parameters tailored to the recipient’s needs. Whether the auditory structure functions as a true creative process or merely as a musical activity is a matter of perspective. From the standpoint of music education, the achievement lies in defining the parameters of the final product, determining its content and style, without direct interaction with the sonic material.

## Conclusions

The analysis of the digital environment in the context of music education reveals various application possibilities. The diversity of resources that enhance multimodal learning through multisensory perception enriches the learning space. The semantic nature of musical art, along with its spatiotemporal functioning, creates a system of modes that interact within musical activity. The trends in digitalization align with the goal of universalizing and optimizing the operational model of musical activities. The visual-sonic-tactile correlation in structuring the audio model within a digital environment ensures information accessibility through crossmodality *in* multisensory arrangement. The greatest advantage is the materialization of the project into a musical-creative activity, presented as an interactive digital product. The influence of AI in education is rapidly growing, but text-based models still dominate. The machine models developed for music have yet to integrate the ability to “read” musical notation. However, for crossmodal priming, it is crucial that visual, auditory, and tactile elements are linked to an understanding of musical syntax. The future of AI in music education remains an open field for further research.

## Acknowledgements

This study is financed by the European Union-NextGenerationEU, in the frames of the National Recovery and Resilience Plan of the Republic of Bulgaria, first pillar “Innovative Bulgaria”, through the Bulgarian Ministry of Education and Science (MES), Project No BG-RRP-2.004-0006-C02 “Development of research and innovation at Trakia University in service of health and sustainable well-being”, subproject “Digital technologies and artificial intelligence for multimodal learning – a transgressive educational perspective for pedagogical specialists” No H001-2023.47/23.01.2024.

## Conflict of interests

The authors declare no conflict of interest.

## References

- Bačlija Sušić, B., & Mičija Palić, M. (2022). Digital competencies in the context of preschool music education. *International Journal of Cognitive Research in Science, Engineering and Education (IJCRSEE)*, 10(2), 77–87. <https://doi.org/10.23947/2334-8496-2022-10-2-77-87>
- Baleva, A. (2010). *Musical perception: Historical and theoretical aspects*. University Publishing House „Paisii Hilendarski“. ISBN 978-954-423-640-3
- Chen, F.-H., Looi, C.-K., & Chen, W. (2009). Integrating technology in the classroom: A visual conceptualization of teachers' knowledge, goals and beliefs. *Journal of Computer Assisted Learning*, 25(5), 470–488. <https://doi.org/10.1111/j.1365-2729.2009.00323.x>
- Dermendzhieva, S., & Tsankov, N. (2022). Transgressive core of multi-modal education. *International Journal of Cognitive Research in Science, Engineering and Education (IJCRSEE)*, 10(3), 167-175. <https://doi.org/10.23947/2334-8496-2022-10-3-167-175>
- Falkner, D. L. (1995). An investigation of modality preferences... [Doctoral dissertation, University of Mississippi]. ProQuest. (AAI9523456). [https://repository.arizona.edu/bitstream/handle/10150/288775/azu\\_td\\_9817346\\_sip1\\_m.pdf?sequence=1&isAllowed=y](https://repository.arizona.edu/bitstream/handle/10150/288775/azu_td_9817346_sip1_m.pdf?sequence=1&isAllowed=y)
- Fleming, N., & Baume, D. (2006). Learning styles again: VARKing up the right tree! *Educational Developments*, 7(4), 4-7. <https://www.vark-learn.com/wp-content/uploads/2014/08/Educational-Developments.pdf>
- Grigorova, Hr. (2019). *Homo sapiens or Homo digitalis? Is pedagogy facing a choice?* In *Информационни технологии в образованието – предизвикателства и възможности* [Information technologies in education – challenges and opportunities], 85–103. University Publishing House “Paisii Hilendarski“. ISBN 978-619-202-437-6
- Holster, J. (2024). Augmenting music education through AI: Practical applications of ChatGP Falkner, D. (1994). *An investigation of modality preferences, musical aptitude, and attitude toward music at the third-grade level* [Doctoral dissertation, University of Mississippi]. ProQuest Dissertations and Theses. (Accession No. AAT 9509516)T. *ResearchGate*. <https://doi.org/10.13140/RG.2.2.35958.57923/1>
- Kordon, A. (2023). *Перспективата изкуствен интелект* [The perspective of artificial intelligence]. Iztok-Zapad. ISBN 978-619-01-1251-8
- Magenta.js. (2023). JavaScript API for music and art generation with machine learning (Version 2.0) [Computer software]. TensorFlow.js. <https://magenta.tensorflow.org/js>
- Melodrive. (2025). *Melodrive: AI-generated adaptive music for games and VR* [Computer software]. itch.io. <https://melodrive.itch.io/>
- Mishra, J. (2007). Correlating musical memorization styles and perceptual learning modalities. *Visions of Research in Music Education*, 9, Article 4. <https://digitalcommons.lib.uconn.edu/vrme/vol9/iss1/4>
- Pastarmadzhiev, V. (2021). *Ispolzuwane na kompiuterni tehnologii pri obuchenieto po muzika v III i IV klas na obshtoobrazovatelnoto uchilishte* [Using computer technologies in music education in grades 3 and 4 of secondary school]. Fast Print Books. ISBN 978-619-236-248-5
- Petkova, D. (2022). Design of musical and performing activities in an electronic environment in the training of students from primary and preschool pedagogy. *Yearbook of the Faculty of Education – Stara Zagora*, 23, 249–260. <https://doi.org/10.15547/YbFE.2022.23>
- Petkova, D. (2023). *Musical computer technologies in education: A model for communicative interaction*. Stara Zagora. ISBN 978-954-314-111-1
- Rexhepi, F. G., Breznica, R. K., & Rexhepi, B. R. (2024). Evaluating the effectiveness of using digital technologies in music education. *Journal of Educational Technology Development and Exchange*, 7(1), 273-289. <https://doi.org/10.18785/jetde.1701.16>
- Sony Computer Science Laboratories, Inc. (1993–2025). *Flow Machines: AI for musical creativity* [Technology description]. <https://www.flow-machines.com/>