education sciences

MDPI

Review

Application of Visual Information in Music Education Digital Technologies: A Scoping Review

Bahareh Behzadaval *, **Laura Serra Marin** and **Luc Nijs**

Department of Education and Social Work, Faculty of Humanities, Education and Social Sciences (FHSE), University of Luxembourg, Belval, 4366 Esch-sur-Alzette, Luxembourg; laura.serramarin@uni.lu (L.S.M.); luc.nijs@uni.lu (L.N.)

* Correspondence: bahareh.behzadaval@uni.lu

Abstract

The relationship between sound and visual representation has long intrigued artists and educators, with historical explorations ranging from colour–music correspondence to alternative notations and graphic visualisations of music. Recent advances in digital technologies have significantly expanded the pedagogical potential of visual information in music education. However, there is still no comprehensive review mapping how visual information is applied in digital music education tools. This scoping review maps the application of visual modalities in original digital tools for music teaching and learning, drawing on 63 studies published between 2014 and 2024. Following Arksey and O'Malley's five-stage framework and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) reporting guidelines, this review analyses the methodological characteristics, pedagogical foundations, and design features of these tools. Findings reveal a dominant focus on performance skills and individual learning, often supported by visual feedback and interactivity. However, other aspects of learning such as creativity, responsiveness, and collaboration remain underexplored. While references to concepts such as multimodality and embodied learning are common, a robust theoretical grounding is frequently lacking or implicit. This review calls for a shift from technology-driven innovation toward pedagogy-led design, advocating for a more holistic educational approach and more rigorous empirical research. Implications highlight the potential of visual information not only to support performance skill acquisition but also to foster creative, expressive, and collaborative dimensions of music learning.

Keywords: music education; visual information; visualisation; educational technology; scoping review

Academic Editor: Sergio Miranda

Received: 19 January 2026

Revised: 9 February 2026

Accepted: 9 February 2026

Published: 13 February 2026

1. Introduction

In recent years, the rapid advancement of digital technologies and the possibilities they create for music educational purposes have led to a continually expanding array of digital tools to innovate and enhance the processes of music learning and teaching (King et al., 2017). One such possibility involves the integration of visual information about music to guide the learning process. Research in cross-modal perception has revealed systematic correspondences between visual and auditory stimuli, demonstrating that structural and aesthetic qualities in visual art can be meaningfully associated with musical dimensions, and vice versa. These correspondences manifest not only at the basic sensory level but also at higher levels of complex, culturally meaningful stimuli, such as works of visual art

 https://doi.org/10.3390/educsci16020309

paired with musical compositions (Spence & Di Stefano, 2025). The visual modality can not only complement auditory input by rendering sounds visible but can also function as pedagogical scaffolds that support learners in, for example, understanding abstract and complex musical concepts.

The idea of making music visible has deep historical roots. Throughout history, artists and musicians have extensively explored the visualisation of music, transferring musical concepts such as harmony, rhythm, and counterpoint to their visual counterparts (Novotná & Ašenbrenerová, 2021).

An important visual counterpart to music is *colour*. As early as Ancient Greece, scholars and artists looked for correspondences between the spectrum of colours and tonal systems, harmonic structures, moods, and emotions. Over time, a variety of representational strategies have been developed to visualise these correspondences, such as colour wheels to map tonal organisation, variations in colour saturation to indicate loudness or intensity, and specific colour mappings to illustrate harmonic proximity or spectral energy distribution (Lima et al., 2021). A notable historical example is the eighteenth-century invention of the *Ocular Harpsichord* by Louis Bertrand Castel, which linked individual tones to corresponding colours through a mechanism of sixty coloured glass panes.

Beyond the use of colours, *geometric and abstract figures* have also served as powerful tools for visualising musical ideas (Hope, 2017). Notable examples include the circular notation system introduced by the music theorist Safi al-Din al-Urmawi (Sethares & Bañuelos, 2007), as well as grid-based representations used to depict musical events over time and to emphasize the cyclical nature of rhythmic structures (Sköld, 2019). Music has further been connected to visible movement through approaches such as the motion curves of Alexander Truslit and Gustav Becking (Nettheim & Becking, 1996). Finally, notation represents one of the most established forms of music visualisation. Evolving from neumes that conveyed only melodic contours, music notation gradually developed into conventional staff notation, which provides detailed visual information about pitch, rhythm, loudness, and expression (Strayer, 2013). In addition, composers and educators explored alternative graphic approaches to music notation, using new symbols and alternative modes for representing pitch and rhythm (Pace, 2007). Graphic notation emerged prominently within the experimental music scene of the 1950s as a response to the perceived constraints of traditional Western staff notation, particularly its limited capacity in representing musical gesture, space, and indeterminacy (Hills, 2011). Composers such as John Cage pioneered non-standard visual representations to expand the role of interpretation, indeterminacy, and performer agency in music making (Cage, 1969; Pace, 2007).

The emergence of new technologies, particularly film and animation, has played a decisive role in the development of music visualisation, giving rise to the domain of *visual* music. Early twentieth-century pioneers such as Mary Ellen Bute and Oskar Fischinger explored the connections between image and sound through analogue techniques such as frame-by-frame animation and celluloid manipulation (Hope, 2017). Subsequent technological innovations further expanded the possibilities for visualising music. For example, Robert Brown's *Atari Video Music* system, incorporating Jeff Minter's *Virtual Light Machine*, represented an early attempt to generate real-time animated graphics synchronized with musical input. The later transition to digital visualisation, exemplified by Stephen Malinowski's *Music Animation Machine*, marked a transformative shift. Novel software enabled the creation of increasingly complex, customizable visualisations and interactive experiences.

https://doi.org/10.3390/educsci16020309

1.1. Music Education and Visualisation

Not surprisingly, the rich history of making music visible has inspired music educators to explore its potential to support music learning. For example, the Tobin Music System (see https://www.tobinmusic.co.uk/content/index.htm (accessed on 8 February 2026)) uses colours to promote the development of music literacy through playful, colour-guided learning. Similarly, Géza Szilvay's Colourstrings approach assigns colours to violin strings to support early instrumental instruction (Sanzone, 2018). Jos Wuytack's Musicogram represents musical elements through colours and geometric figures (Boal-Palheiros & Wuytack, 2006), inviting learners to follow them with their fingers in order to grasp the structure and expression of the music.

Music maps are another category of visual tools, ranging from abstract representations (e.g., lines and dots) to figurative images depicting realistic objects, animals, or humans (Şen, 2021). As creative listening activities, these maps support deeper analysis of musical works and enhance the perception of musical form (Dunbar, 2017). In some teaching and learning contexts, visuals corresponding to music are created by the learners themselves through drawing activities. Such approaches have been shown to improve enjoyment of the learning process and foster a better understanding of musical concepts (Roels & Van Petegem, 2014; Soysal, 2012).

Musicograms and music maps are forms of graphic notation, i.e., music notation systems that use visual symbols, shapes, and pictorial elements to represent musical parameters such as pitch, duration, and dynamics. Such systems often seek to provide an alternative to the symbolic conventions and the literacy they require, offering flexibility and openness for interpretation (Zhou et al., 2026). Graphic notation not only appeals to children's natural inclination to produce graphical representations of music (e.g., Verschaffel et al., 2009), it also supports multimodal collaborative learning (H. H. Ng, 2019), memory (P. N. Lee, 2013), and active listening (Boal-Palheiros & Wuytack, 2006).

Visual information is also occasionally used to complement traditional notation. For instance, Fürst et al. (2020) juxtaposed conventional sheet music with colourful circular shapes, termed *rhythmic fingerprints*, to facilitate novice learners' perception of complex rhythmic patterns. In contrast to systems that merely extend traditional notation, some approaches fundamentally redesign it for accessibility. For example, Kivijärvi (2019) introduced the Figurenotes system as an alternative to standard Western notation, using colours and shape-based symbols to indicate pitch levels on instruments. The system simplifies music reading and assists students who struggle with traditional notation.

1.2. Music Educational Technologies and Visual Information

Digital technologies have substantially expanded the potential of visualising music for educational purposes, shifting from static notation toward dynamic, interactive, and immersive forms of visually augmented musical learning experiences that can support, for example, the development of music perception, creativity, improvisation, and practice (León-Garrido et al., 2022). Through real-time signal processing and advanced graphical interfaces, complex auditory information can now be translated into colours, shapes, and interactive visual elements, enabling the creation of rich multimodal learning environments. Such environments have been shown to promote engaging learning experiences (Gorbunova & Hiner, 2019) and to foster a deeper understanding of music (T. Wang & Yu, 2024).

While traditional music notation has been used in a static form for centuries, digital environments enable dynamic visualisations that unfold in time. Given the inherently temporal nature of music, such visualisations can enhance the perception of its temporal features, such as rhythm, pulse, and phrasing (Grahn, 2012). Moreover, the real-time and data-based visualisation of music can function as formative feedback for learning, providing information about, for example, performance accuracy, including musical aspects such as intonation, dynamics, and pitch (Aksoy, 2023; Hamond et al., 2019; Muresan & Pop, 2017).

In some cases, adopting a student-centred and creative approach, visual representations extend beyond feedback on performance accuracy to support users' creative expression. For instance, the "music paint machine" developed by Nijs and Leman (2013) allowed learners to generate a digital painting by playing music, thus integrating creative visual and sonic output.

Arguably, the rapid proliferation of digital applications that integrate visual information into music learning and teaching, together with the pace of technological innovation, which may outpace the systematic evaluation of their educational effectiveness, underscores the need for a comprehensive examination of why and how digitised visual information is integrated in music learning.

1.3. Previous Reviews and the Present Study

Several reviews have explored the role of technology in music education, either from a broad perspective or focusing on specific tools such as the metaverse or artificial intelligence. Kardeş (2022), for example, analysed 213 studies published between 2000 and 2022, finding that most of them relied on convenience sampling, survey methods, and basic statistical analyses, indicating limited methodological rigour. Similarly, Caamaño Liñares et al. (2023), in their review of 252 articles published between 2011 and 2020, reported small sample sizes and vague research procedures. Despite these methodological issues, León-Garrido et al. (2022) concluded that ICT can positively affect motivation, creativity, and critical thinking, although much of the existing research remains descriptive and lacks pedagogical depth.

More recent reviews have targeted specific technologies. Lam (2024) identified digital tools that support creativity in K–12 settings, although evidence for causal effects remained limited. C. Liu et al. (2023) highlighted the growing interest in mobile technology–supported music education, with a particular emphasis on learners' perceptions. Lin and Yunus (2024) mapped metaverse-based tools, including virtual and augmented reality, artificial intelligence, and speech recognition. Merchán Sánchez-Jara et al. (2024) identified nine applications areas of AI in music education, while Y. Zhang et al. (2024) described AI use in Chinese higher education as still emerging but promising. Finally, Weatherly et al. (2024) examined digital game-based learning, emphasizing persistent issues of cost and access and calling for more experimental research designs.

Although previous reviews highlight the growing implementation of music educational technologies that integrate visual information, such as mobile devices, virtual reality, and augmented reality, they do not offer a comprehensive analysis of why and how *visual* modalities are used in these technologies. This scoping review addresses this gap by mapping the landscape of visual elements in music education technologies, identifies prevalent design approaches and pedagogical frameworks, and synthesizes existing findings to inform the future development of effective visual learning environments in music education. In this review, the term *music education* is used in a broad sense, encompassing learning across formal, non-formal, and informal contexts.

Our research questions were as follows:

- Q1. What are the general characteristics of the included studies in terms of publication year, geographical location, and type of document?
- Q2. What are the methodological characteristics of empirical studies, including sample size, participants' age groups, data collection methods, and methodological approaches?
- Q3. What are the defining characteristics of the original digital tools developed for teaching and/or learning music?

https://doi.org/10.3390/educsci16020309

- Q3.1. What types of technology and display devices are used?
- Q3.2. What conceptual and pedagogical foundations underpin the design of these tools?
- Q3.3. What music education objectives do the tools aim to support?
- Q3.4. What types of visual information are used?
- Q3.5. What functions does the visual information serve?

2. Methods

This scoping review aimed to provide a comprehensive overview of the application of visual information in digital technologies for music education, based on literature published between 2014 and 2024. This period is characterised by rapid advances in visual display technologies (e.g., high-resolution touchscreens, head-mounted displays), the widespread adoption of mobile and touch-based devices (e.g., tablets and smartphones enabling ubiquitous access and gesture-based interaction), and the integration of rich multimedia features. Together, these developments have expanded the pedagogical potential of digitised visual information in music education.

To structure the review, we followed Arksey and O'Malley (2005) five-stage methodological framework: (1) Identifying the research question, (2) Identifying relevant studies, (3) Study selection, (4) Charting the data, (5) Collating, summarising, and reporting the results. To guide transparent and systematic reporting, we adopted the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) (Tricco et al., 2018). No protocol was registered for this review, since protocol registration is optional for scoping reviews.

2.1. Eligibility Criteria

We included English-language publications, encompassing journal articles, conference papers, and book chapters, that introduced original digital tools designed for music teaching and learning. To achieve a comprehensive overview of digital technologies in music education, we did not restrict studies based on educational context. Studies situated in formal, non-formal, and informal music learning settings were all eligible for inclusion, provided they met the remaining inclusion criteria. We considered a digital tool to be original when it was neither commercially available nor previously introduced in the scientific literature. To enable an in-depth examination of the role of visual information, we only considered studies that provided explicit explanations and/or images of the tools' interfaces of the tools. We excluded publications that solely presented methodological or technical aspects (e.g., algorithmic descriptions) of presented digital tools as well as those without implementing a complete development process for an original tool. Furthermore, studies investigating music visualisations for purposes unrelated to education, such as psychological, cognitive, or therapeutic studies, were also excluded. Only full-text publications were considered eligible.

2.2. Information Sources and Search Strategy

We conducted systematic searches on 22 May 2025 in Scopus and Web of Science (WoS), widely regarded as standard sources in music education and educational technology research, and frequently used in bibliometric studies and systematic reviews in the field (Li et al., 2025; Ma & Wang, 2025; Sánchez-Marroquí & Vicente-Nicolás, 2024). In Scopus, the search was limited to journal articles, conference papers, and book chapters. In WoS, it was restricted to journal articles and conference proceeding papers. We used three groups of keywords in the search strategy to search for titles, abstracts, and keywords (see Table 1).

Table 1. Search strategy.

No.	Searched Items in the Titles, Abstracts, and Keywords
1	"visual*" OR "graphic*" OR "animat*" OR "augmented reality" OR "virtual reality" OR "extended reality" OR "mixed reality"
2	"music education" OR "music learn*" OR "music teach*" OR "learning music" OR "teaching music" OR "music instruction"
3	"technolog*" OR "digital*" OR "multimedia" OR interface* OR interact* OR "software" OR "gam*" OR "prototype" OR computer*
4	1 AND 2 AND 3

Note. The asterisk (*) indicates truncation (wildcard) to retrieve multiple word endings (e.g., visual*, graphic*, animat*).

2.3. Selection of Sources of Evidence

We initially identified 642 records across the two databases (Scopus: n = 398; WoS: n = 244). Using Rayyan, an AI-powered systematic review management platform, the first author removed duplicate records (n = 182) and performed screening by title and abstract, resulting in 152 inclusions and 308 exclusions. The first and second authors independently conducted full-text assessment. Based on the predefined eligible criteria, 63 publications were ultimately included (see Figure 1).

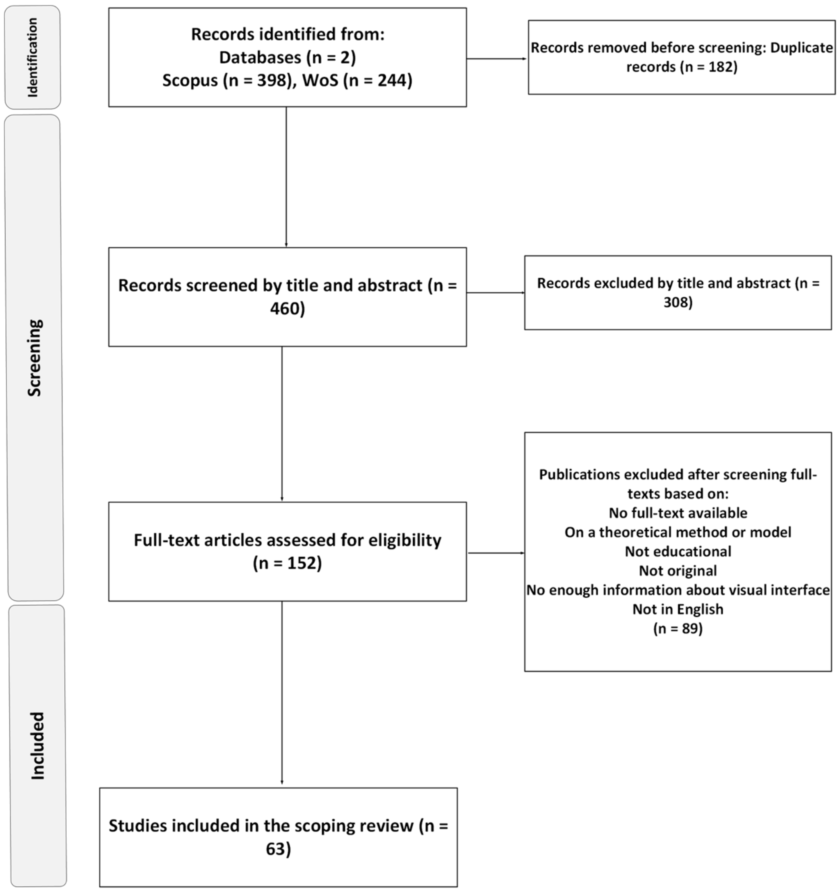


Figure 1. Flowchart of article selection process.

2.4. Data Charting Process and Data Items

Data extraction addressed the research questions following two complementary approaches. For general and methodological aspects (Q1 and 2), we used a *top-down* approach (Hu et al., 2024). For methodological classification, only empirical studies were examined and categorised by sample size, age of participants, data collection methods, and methodological approaches. The first and third authors discussed and verified the extracted data.

As this review represents the first comprehensive analysis in this domain, a *bottom-up* approach classified the data to answer the third research question concerning the features of music education technologies integrating visual information (Hu et al., 2024). We adopted thematic analysis approach (Braun & Clarke, 2006) during coding and synthesis. The first author systematically analysed the data to identify initial codes and emerging themes. The first and third authors discussed the relationships among these themes, which led to the formulation of overarching categories.

The first author conducted the primary data extraction according to the research questions and established categories. To reduce potential bias, one third of the data set (21 of 63 publications) was randomly reviewed by the third author. The first and third authors then engaged in detailed discussion of the data to verify consistency, resolve discrepancies, and ensure the reliability and accuracy of the extracted information.

We used SciSpace (SciSpace, 2024–2025) to support the data extraction process through AI-based summarisation and organisational features that help streamline the review process. However, all final decisions regarding article inclusion, data extraction, and analysis were made manually by the researchers to ensure methodological rigour. Table 2 provides an overview of the main characteristics of the included publications.

Table 2. Overview of study characteristics.

No.	Author(s)/Year/Country	Name of the Tool	Type of Technology and Display Device	Types of Visuals	Music Education Objectives	Empirical Study
1	(Dascălu et al., 2014)/Romania	Touch & Listen	Screen_based/Computer	Text/Colours/shapes/2D or 3D images	Performing/Creating	Yes
2	(Schönfeld et al., 2014)/Germany	Vocalmetrics	Screen_based/Web-based	Plots and/or spectrogram	Performing/Responding/Connecting	No
3	(Perdana, 2014)/Malaysia	Not given	Screen_based/Computer	Text/Colours/Shapes/2D or 3D images	Performing/Responding	Yes
4	(X. Xiao et al., 2014)/USA	Andante	AR/Projected light	Colours/2D or 3D images	Performing/Creating/Responding	No
5	(Franceschini et al., 2014)/UK	Not given	Screen_based/Tabletop	Colours/Shapes	Creating/Generic knowledge	Yes
6	(Keebler et al., 2014)/USA	Fretlight®	AR/Installed light	Colours	Performing	Yes
7	(Martins et al., 2015)/Brasil	Music-AR	AR/Computer	Text/Colours/Shapes/2D or 3D images	Responding/Generic knowledge	Yes
8	(T. K.-L. Ho et al., 2015)/Taiwan	VQVA-Sys	Screen_based/Computer	Text/Shapes/Traditional notation/Plots and/or spectrogram	Performing	Yes
9	(Oestermeier et al., 2015)/Germany	LEGO music	Screen_based/Tabletop	Shapes	Creating/Generic knowledge	Yes
10	(X. Xiao et al., 2016)/USA	Andantino	AR/Projected light	Colours/Shapes/2D or 3D images/Traditional notation	Performing/Creating/Responding/Generic knowledge	Yes
11	(Baldassarri et al., 2016)/Spain	ImmertableApp	Screen_based/Multi-interface	Text/Colours/Shapes/2D or 3D images	Creating	Yes
12	(Ling, 2016)/Austria	Not given	Screen_based/Computer	Colours/Shapes/2D or 3D images/Traditional notation	Generic knowledge	No
13	(Huang & Chu, 2016)/Taiwan	Not given	Screen_based/Web-based	Text/Shapes/Plots and/or spectrogram/sound wave form	Performing	Yes
14	(L. Liu et al., 2016)/China	Not given	Screen_based/Computer	Text/Colours/Shapes/2D or 3D images/Plots and/or spectrogram/Sound wave form	Performing	Yes

Table 2. *Cont.*

No.	Author(s)/Year/Country	Name of the Tool	Type of Technology and Display Device	Types of Visuals	Music Education Objectives	Empirical Study
15	(Pérez-Gil et al., 2016)/Spain	Cantus	Screen_based/Web-based	Text/Colours/Shapes/Traditional notation/Plots and/or spectrogram	Performing/Generic knowledge	Yes
16	(Kumar et al., 2016)/India	Not given	Screen_based/Computer	Text/Shapes/Plots and/or spectrogram/Sound wave form	Performing	No
17	(Hong et al., 2016)/USA	MiLa	Screen_based/Web-based	Text/Colours/2D or 3D images	Performing/Generic knowledge	Yes
18	(Fernandez et al., 2016)/Japan	Not given	AR/HMD	Text/Colours/Shapes/2D or 3D images	Performing	No
19	(Paule-Ruiz et al., 2017)/Spain	SAMI	Screen_based/Mobile	Text/Colours/Shapes/2D or 3D images	Creating/Responding/Generic knowledge	Yes
20	(M. Ng & Schutz, 2017)/Canada	MAESTRO	Screen_based/Computer	Text/Colours/Shapes/2D or 3D images/Plots and/or spectrogram/Sound wave form	Responding	No
21	(Das et al., 2017)/USA	Music Everywhere	AR/HMD	Text/Colours/Shapes/2D or 3D images	Performing/Creating	No
22	(Johnson & Tzanetakis, 2017)/Canada	VRMin	MR/HMD	Text/Colours/Shapes/2D or 3D images	Performing	No
23	(Trujano et al., 2018)/USA	ARPiano	AR/HMD	Text/Colours/Shapes/2D or 3D images	Performing/Generic knowledge	Yes
24	(Cheng, 2018)/USA	Not given	Screen_based/Computer	Colours/Shapes/2D or 3D images/Plots and/or spectrogram	Performing	Yes
25	(Alvarez-Molina et al., 2018)/Germany	Not given	Screen_based/Computer	Text/Colours/Shapes/2D or 3D images/Traditional notation	Responding/Generic knowledge	Yes
26	(Tan & Lim, 2019)/Malaysia	MMAR	AR/Mobile	Text/Shapes/2D or 3D images	Connecting/Generic knowledge	Yes
27	(Fletcher et al., 2019)/UK	Virtual Reality Ear Training System	VR/HMD	Text/Colours/Shapes/2D or 3D images	Responding/Generic knowledge	Yes
28	(Klamka et al., 2019)/Germany	ScaleDial	Other/Installed light	Text/Colours/Shapes	Generic knowledge	No
29	(Del Rio-Guerra et al., 2019)/Mexico	Not given	Screen_based/Computer	Colours/Shapes/2D or 3D images	Performing	Yes
30	(Nasrifan & Saidon, 2019)/Malaysia	Asas Teori Muzik	Screen_based/Web-based	Text/Colours/Shapes/2D or 3D images/Traditional notation	Generic knowledge	Yes
31	(Micheloni et al., 2019)/Italy	Musa	Screen_based/Multi-interface	Colours/Shapes/2D or 3D images	Performing	Yes
32	(Johnson et al., 2020)/Canada	MR:emin	MR/HMD	Text/Colours/Shapes/2D or 3D images	Performing	Yes
33	(Pedersen et al., 2020)/UK	Virtual Reality Ear Training System	VR/HMD	Text/Colours/Shapes/2D or 3D images	Responding/Generic knowledge	Yes
34	(C. Xiao, 2020)/USA	MADSA	Screen_based/Computer	Text/Shapes/Plots and/or spectrogram	Performing	Yes
35	(Rigby et al., 2020)/New Zealand	piARno	AR/Computer	Text/Colours/Traditional notation	Performing/Generic knowledge	Yes
36	(Blanco et al., 2021)/Spain	SkyNote	Screen_based/Computer	Text/Colours/Shapes/Traditional notation/Plots and/or spectrogram	Performing	Yes
37	(B. Wang et al., 2021)/Canada	Soloist	Screen_based/Web-based	Text/Colours/Shapes/Plots and/or spectrogram/Sound wave form/Video	Performing	Yes
38	(Haferkamp & Dengel, 2021)/Germany	Not given	Screen_based/Computer	Text/Shapes/2D or 3D images	Responding/Generic knowledge	Yes
39	(Y. Lu et al., 2021)/China	Not given	AR/Mobile	Text/Colours/Shapes/2D or 3D images	Generic knowledge	Yes
40	(Guclu et al., 2021)/Turkey	AR-Flute App	AR/Mobile	Text/Colours/Shapes/2D or 3D images/Traditional notation	Performing/Generic knowledge	No

Table 2. *Cont.*

No.	Author(s)/Year/Country	Name of the Tool	Type of Technology and Display Device	Types of Visuals	Music Education Objectives	Empirical Study
41	(Molero et al., 2021)/Spain	HoloMusic XP	MR/HMD	Text/Colours/Shapes/2D or 3D images	Performing	Yes
42	(Corintha et al., 2021)/Portugal	AM-I-BLUES	Other/Installed light	Colours	Performing/Creating	Yes
43	(Feng et al., 2022)/China	Music Playground	Screen_based/Computer	Text/Colours/2D or 3D images	Generic knowledge	No
44	(Chin & Xia, 2022)/China	Not given	Screen_based/Computer	Colours/Shapes/Traditional notation	Performing/Generic knowledge	Yes
45	(Avanzini et al., 2022)/Italy	AREmbody	AR/Mobile	Text/Colours/Shapes/2D or 3D images	Responding/Generic knowledge	No
46	(K. Lee, 2022)/Korea	Not given	Screen_based/Extra screen	Colours/Shapes	Performing	Yes
47	(Blewett & Gerhard, 2023)/Canada	Brass Haptics	VR/HMD	Colours/Shapes/2D or 3D images/Traditional notation	Performing	No
48	(C.-L. Ho et al., 2023)/Taiwan	Let's go to the Magic Forest concert!	AR/Mobile	Text/2D or 3D images	Connecting/Generic knowledge	Yes
49	(Y.-H. Wang, 2023)/Taiwan	Not given	Screen_based/Computer	Text/Colours/Shapes/2D or 3D images/Traditional notation	Responding/Generic knowledge	Yes
50	(X. Wei et al., 2023)/China	Forest Kingdom Concert	AR/Mobile	Text/Colours/Shapes/2D or 3D images	Performing/Generic knowledge	Yes
51	(López-Calatayud & Tejada, 2023)/Spain	Plectrus	Screen_based/Web-based	Text/Colours/Shapes/2D or 3D images/Traditional notation/Plots and/or spectrogram	Performing	Yes
52	(Pinkl & Cohen, 2023)/Japan	Halvatar	VR/HMD	Text/Colours/Shapes/2D or 3D images	Performing	Yes
53	(Pinkl et al., 2024)/Japan	Not given	MR/HMD	Text/Colours/2D or 3D images	performing	Yes
54	(Meneses Rodríguez & Sánchez, 2024)/Chile	Tone Cluster	Screen_based/Mobile	Text/Colours/Shapes/2D or 3D images/Traditional notation	Generic knowledge	Yes
55	(Amm et al., 2024)/Germany	Piano Theory Hub	MR/HMD	Text/Colours/Shapes/2D or 3D images/Traditional notation	Performing/Generic knowledge	Yes
56	(Pesek et al., 2024)/Slovenia	Steady the Drums!	VR/HMD	Text/Colours/Shapes/2D or 3D images	Performing/Responding	Yes
57	(D. Lu, 2024)/China	Not given	AR/Multi-interface	Text/Colours/Shapes	Performing	Yes
58	(O'Toole et al., 2024)/Ireland	Chromaesthetica	VR/HMD	Text/Colours/Shapes/2D or 3D images	Generic knowledge	No
59	(Stachurska et al., 2024)/Poland	HarmosphereVR	VR/HMD	Text/Colours/Shapes/2D or 3D images/Traditional notation	Generic knowledge	Yes
60	(Deja et al., 2024)/Slovenia, Philippines, South Africa	ImproVISe	AR/Projected light	Text/Colours/Shapes	Creating	No
61	(B. T. Wei et al., 2024)/USA	Sympathetic Orchestra	Screen_based/Computer	Text/Colours/Shapes	Conducting	No
62	(Safian et al., 2024)/Malaysia	e-MARZ	VR/HMD	Text/Shapes/2D or 3D images/Video	Connecting	Yes
63	(Delegido et al., 2024)/Spain	Not given	Screen_based/Multi-interface	Text/2D or 3D images	Responding/Generic knowledge	Yes

3. Results

3.1. General Characteristics of the Included Publications

Sixty-three publications introducing original digital tools in the field of music education were examined in this review. Table 3 summarizes the distribution of publication types across categories. Conference papers constituted the majority of the included publications (n = 35; 55.56%), followed by journal articles (n = 26; 41.27%). Only two book chapters met the eligibility criteria (3.17%). Across the 11-year period, no consistent upward or downward trend in publication output was observed. Nevertheless, the year 2024 stood out with a notably higher number of publications (*n* = 11; 17.46%) compared to other years.

Table 3. Type of publications.

Type of Document	Number of Documents	%
Conference papers	35	55.56%
Journal articles	26	41.27%
Book chapters	2	3.17%

As several of the included studies did not involve empirical research, sufficient data on study locations were not consistently available. Therefore, we classified the geographical distribution of the studies based on the affiliation of the first authors (see Figure 2). Europe accounts for the highest number of studies (*n* = 26, 41.3%), followed by Asia (*n* = 19, 30.2%) and North America (*n* = 14, 22.2%). South America and Oceania contributed three (4.8%) and one (1.6%) study, respectively.

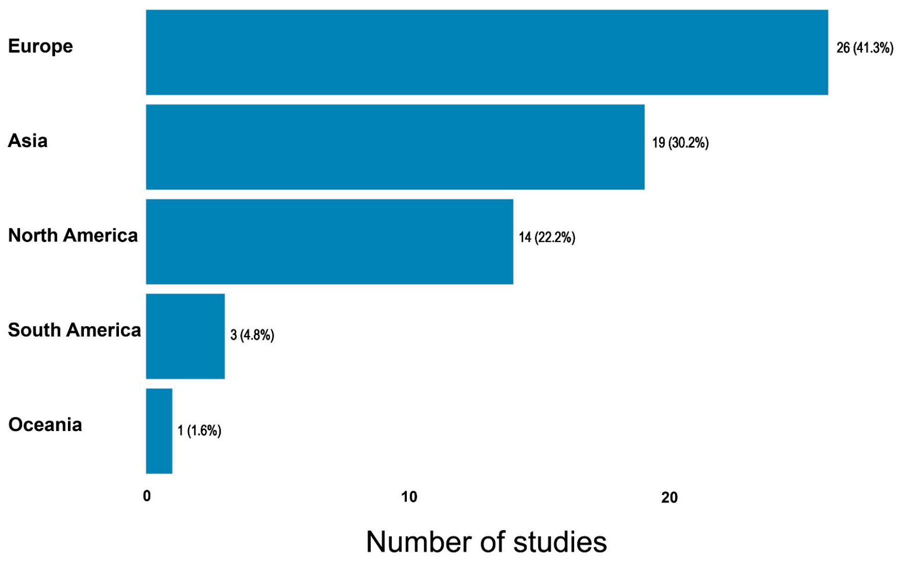


Figure 2. Locations of the studies.

3.2. Methodological Characteristics in Empirical Studies

Of the 63 included publications, 47 studies conducted empirical research to evaluate the designed and developed digital tools. The results indicated that more than 31% (*n* = 15) involve small sample sizes, ranging from 2 to 18 participants. A comparable proportion, approximately 31% (*n* = 15), included medium-sized samples, ranging from 20 to 47 participants. In addition, more than 27% (*n* = 13) of the studies involved large samples, ranging from 51 to 497 participants. Four studies (8.51%) reported the involvement of participants but did not provide information on sample size.

To classify participants' age groups, we used the INTEGRIS Health's Life Stages (2015). As shown in Figure 3, 18 studies did not report specific age details, although four referred to participants as students, graduate students, or staff members at schools and/or universities. A total of 15 studies focused on children aged 5–12 years, with most of them (*n* = 10) examining this group exclusively. Adolescents aged 13–19 were the exclusive focus in two studies and were included with other age groups in nine additional cases. Adults aged 20–39 years were included in 14 studies, most often alongside other age groups. Middle-aged adults (40–59 years) appeared in four studies, always in combination with three other age groups.

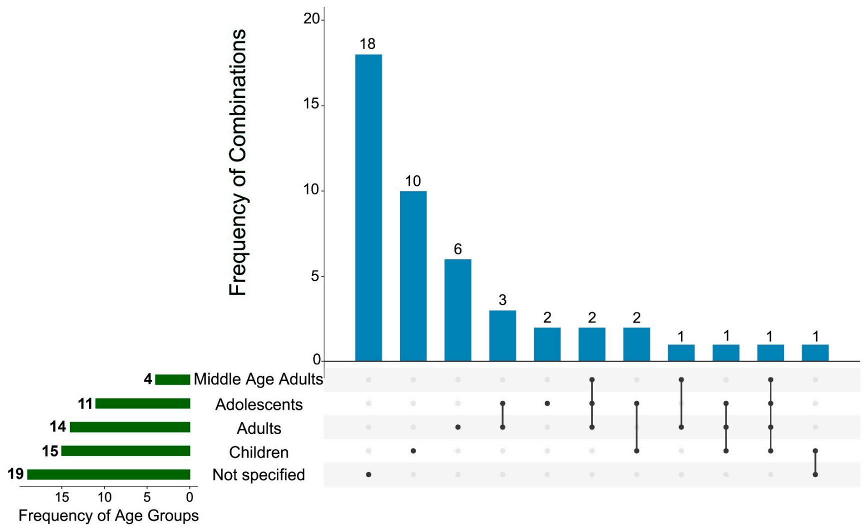


Figure 3. Age of participants in empirical studies (47 studies out of 63). Green horizontal bars indicate the total frequency of each age group across all studies. Blue vertical bars show the frequency of each age-group combination. Filled dots identify which age groups are included in a given combination, and vertical connecting lines indicate intersections between those groups. Numbers above the bars represent the count of studies for each combination.

Studies used a variety of methods, sometimes combined (see Figure 4). *Questionnaires* and surveys (n = 29; 61.70) were the most frequently used data collection methods, focusing on user feedback. Questionnaires probed usability, user experience, appreciation, and game experience. Only two studies used questionnaires to assess the effectiveness of their digital tools through pre- and post-test measures of participants acquired knowledge. Some questionnaires also included demographic and open-ended questions. Nine studies relied solely on questionnaires, while 20 studies combined them with other methods.

Performance metrics were the second most frequently used method (n = 23; 48.93%), typically collected directly from the digital tools. Some studies used standardized or researcher-designed tests for performance evaluation. This method was usually combined with other methods.

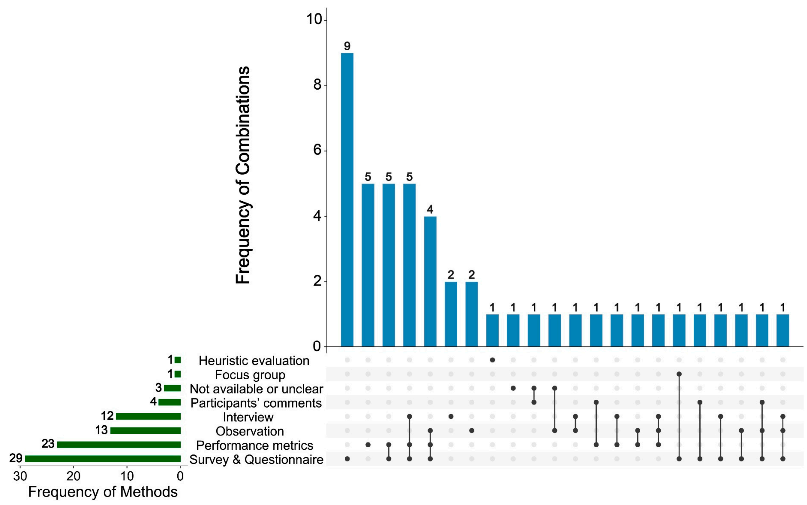


Figure 4. Data collection methods in empirical studies (47 studies out of 63). Green horizontal bars indicate the total frequency of each data collection method across studies. Blue vertical bars show the frequency of specific combinations of methods. Filled dots identify the methods included in each combination, and vertical connecting lines indicate intersections between methods. Numbers above the bars represent the number of studies corresponding to each combination.

In addition, 13 (27.65%) studies used *participant observation* and 12 studies (24.53%) conducted *interviews* with participants following the use of the digital tools. A few studies used video recording to capture participants behaviour. Four studies collected user comments during or after the experiments but did not specify whether the these were obtained through structured or semi-structured interviews. These studies were classified separately. One study conducted a focus group, in addition to a questionnaire, and one study employed heuristic evaluation.

Finally, three studies lacked sufficient detail regarding their data collection and were therefore classified as *not available or unclear*. Although two of these studies partially referred to observation and participants' comments as data collection methods, they lacked sufficient detail to allow for clear categorisation.

Mixed methods were the most frequently adopted methodological approach in the empirical studies (n = 22; 46.81%), followed by quantitative (n = 14; 29.79%) and qualitative (n = 10; 21.28%) approaches. One publication reported limited evidence of empirical data collection but did not report any methodological details. In addition, two mixed method studies mentioned the use of quantitative measurements but did not report the corresponding data or analyses.

Of the 36 studies that included quantitative analysis, 19 (54.28%) applied inferential statistics, most commonly t-tests (n = 9) and ANOVA tests (n = 7). The remaining studies relied solely on descriptive statistics such as means, medians, standard deviations, or percentages. While several studies detailed their qualitative analysis procedures, others reported outcomes without specifying the data analysis methods used (see Table 4).

Table 4. The methodological approach in empirical studies (46 studies out of 63).

Method	Number of Studies	%
Quantitative	14	29.79%
Qualitative	10	21.28%
Mixed	22	46.81%
Undefined	1	2.13%

3.3. Characteristics of the Original Digital Tools

One of the main objectives of this study was to examine digital tools in music education that incorporate visual interfaces from multiple perspectives. To investigate how these tools were designed and developed, we applied a bottom-up analysis, systematically breaking down and synthesizing data in relation to the third research question and its sub-questions.

3.3.1. Technologies and Display Devices Used in the Digital Tools

We identified five types of technologies among 63 digital tools (see Figure 5): (1) Screen-based (n = 32; 50.79%), (2) Augmented Reality (AR) (n = 16; 25.39%), (3) Virtual Reality (VR) (n = 8; 12.69%), (4) Mixed Reality (MR) (n = 5; 7.93%), and (5) other technologies (n = 2; 3.17%). Display devices were categorised into seven categories: (1) computer (laptop, PC), (2) mobile devices (mobile phones and tablets), (3) web-based, (4) tabletop, (5) Head Mounted Display (HMD), (6) installed lights, and (7) projected lights. Some tools used *multi-interface* configurations (e.g., PC and mobile), observed in screen-based (n = 3) and AR (n = 1) tools.

Screen-based tools primarily used computers (n = 17), followed by web-based (n = 7), mobile (n = 2) and tabletop displays (n = 2). One screen-based tool applied an extra-large display screen as its primary visual interface. Augmented Reality (AR) featured mobile devices (n = 6), HMDs (n = 3), or computers (n = 2).

Several digital tools used lights as visual information, projected onto the surface of a piano (n = 3), or installed on the guitar fretboard or on a tangible device. VR and MR-based tools primarily used HMDs, with a marked increase (n = 6; half of the studies) in 2024.

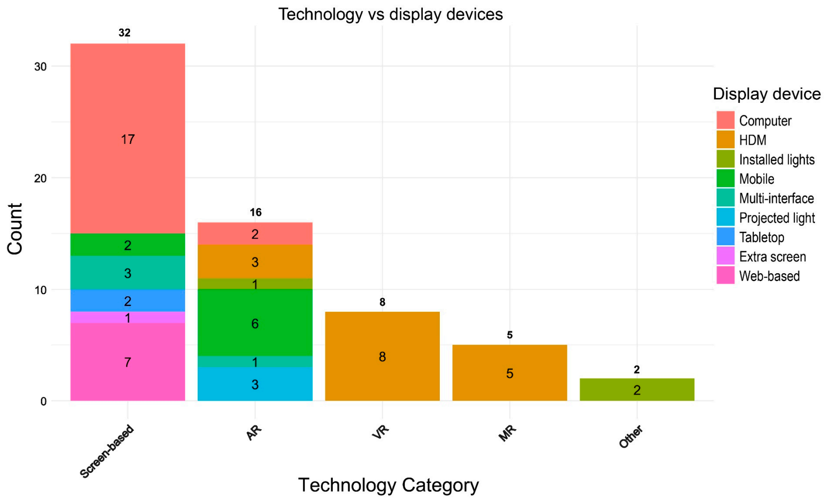


Figure 5. Technologies and display devices. Stacked bars show the number of studies for each technology category by display device type (colours). Numbers within each segment indicate the count of studies using that specific display device, and numbers above each bar indicate the total number of studies within the technology category.

3.3.2. Conceptual and Pedagogical Foundations

We examined the conceptual and pedagogical foundations underlying the design of the original tools described in the included studies, focusing on explicitly stated theories, conceptual frameworks, guiding principles, and underlying assumptions. Through thematic analysis, we grouped the most frequent concepts into 14 categories: (1) Interactive learning (n = 29), (2) Feedback in learning (n = 27), (3) Multimodal learning (n =18), (4) Individual learning (n = 14), (5) Game-based learning (n = 11), (6) Embodied learning (n = 10), (7) Immersive learning (n = 9), (8) Tangible interfaces in learning (n = 5), (9) Music learning theories (n = 4), (10) Collaborative learning (n = 4), (11) Computer-assisted learning (n = 4), (12) Visual learning (n = 3), and (13) Multimedia in learning (n = 3). Appendix A presents the keywords identified in the thematic analysis and their associated categories and corresponding publications.

We classified studies under *interactive learning* when they referred to concepts such as interactive learning environments, interaction techniques, interactive video tutorials, or visual music pedagogy. These studies aimed to enhance interactivity in music learning by leveraging technological potential. While some were informed by established theoretical frameworks and previous works, many lacked a clearly articulated or empirical justification for their design decisions.

Studies focusing on *feedback in learning* used two approaches: real-time feedback and visual feedback. While many studies emphasized their combined use, some focused exclusively on real-time feedback. A few studies also incorporated haptic feedback, often integrated with other modalities, such as visual, and auditory feedback. Overall, the studies underscore the potential of technologies to support learners' independent skill development, especially in the absence of a teacher.

Studies addressing *multimodal learning* drew on theories involving the educational use of multiple sensory channels (e.g., visual, aural, and/or kinaesthetic). While most studies mainly focused on the integration of visual and auditory modalities in learning, some studies additionally integrated sensorimotor and kinaesthetic modalities to support embodied learning.

The studies categorised under *individual learning* aimed to promote various self-related processes in music learning, such as self-learning, self-development, self-regulated learning, self-directed learning, self-assessment, and self-efficacy. The goal was to help students taking charge of their own musical growth by encouraging them to reflect on their progress and independent practice. In many cases, digital tools and personalized learning activities were used to support this autonomy.

Studies involving games and gamified tools applied the key principles of *game-based learning*, including game mechanics and gamification techniques to align musical learning objectives with engaging gameplay elements. Several studies integrated gamification elements, such as progression levels, awards and challenges, to motivate users in the music learning process.

In studies on *embodiment* in music education, Leman's theory of embodied music cognition was the most frequently referenced. Although most of the studies in this category emphasized physical interaction in supporting learning, they approached embodiment from diverse educational perspectives. For example, a few studies addressed the importance of muscle memory and bodily engagement and movements for understanding musical concepts. In addition, some studies integrated embodiment in visual representations to promote musical expressions and interactive or enactive, reduce cognitive learning, and support associations between visual cues, musical notes, and hand postures.

The studies focusing on *immersive learning* applied technologies such as virtual reality (VR), augmented reality (AR), and the metaverse to create interactive and multisensory experiences that allow learners to practice skills, explore complex concepts, and collaborate in engaging ways.

In the *music learning theories* category, the included studies relied on pedagogical methods and approaches that explain how learners develop musical understanding. We identified three specific music learning theories referenced in these studies, including Gordon's, Kodály's, and Dalcroze's.

Despite the use of concepts and terms as foundations for tool design, many studies did not explicitly reference established theoretical frameworks. In fact, authors scarcely provided scientific rationale explaining why these frameworks are important for music learning or how they inform the design of their digital tools. Furthermore, four studies referred to only a single category (e.g., real-time feedback or interactive learning) without citing related theories or previous research. One study did not specify any theoretical or pedagogical framework underlying the design and development of its digital tools.

3.3.3. Music Education Objectives

We classified the music education objectives of the digital tools into six categories:

- Performing: the development of technical skills in playing an instrument or singing, such as fingering techniques, articulation, expression, pitch accuracy, and intonation.
- Creating: learning to improvise, compose, and generate sounds.
- Responding: responsiveness to music, both internally (awareness and recognition of sound characteristics) and externally (interaction with music through physical movement) (Abril & Gault, 2016).
- Connecting: linking music to other disciplines, cultures, histories, and contemporary experiences (Abril & Gault, 2016).
- Generic knowledge: acquiring knowledge about music theory, including traditional notation and musical concepts (rhythm, pitch, melody, intonation, etc.), musical instruments, in addition to developing sight reading skills.

https://doi.org/10.3390/educsci16020309

- Conducting: developing skills in leading ensembles, including gestures, rehearsal techniques, interpretation, score reading, and communication with performers.

Thirty-seven tools were developed to enhance *performance skills*. In 21 studies, performance was the sole educational objective, while the remaining studies addressed it alongside other learning objectives. For example, among the 13 studies on digital tools for piano learning, four also targeted developing improvisation skills (creating). A total of 19 studies focused on developing performance skills on other instruments, such as guitar, violin, trumpet, flute, drum, oboe, and theremin, were addressed in. In addition, six studies focused on developing singing and vocal skills, mainly providing feedback on characteristics of the voice, particularly pitch accuracy. One study introduced a gesture-controlled tool that enabled users to perform melodies via motion tracking, accompanied by real-time visual feedback.

Acquiring *generic knowledge* was the second most frequently addressed learning objective (n = 30). Eight focused exclusively on this objective, while 21 studies combined it, most often, with developing performance skills (n = 8). Three studies supported knowledge acquisition about a variety of musical instruments, two in combination with the *responding* objective and one in combination with the connecting objective.

Of the 15 tools designed to address the *responding* objective, 11 focused on internal responsiveness, such as ear training, and the perception and recognition of pitch, timbre, and harmony. Except for one study on timbre perception, all tools addressed multiple learning objectives. Four tools targeted external responsiveness by using dynamic visual representations to support users' synchronization with music during performance.

A total of 10 studies developed digital tools to promote *creativity*. Five focused on piano improvisation, four on composition, and one on sound generation. While creating was commonly addressed in combination with other learning goals, only two studies centred it as the sole educational objective: one focused on sound generation, and the other on piano improvisation. Finally, four digital tools were classified under the *connecting* category. Three introduced traditional music and instruments from Taiwan and Malaysia, and one targeted a variety of music styles.

One tool targeted conducting skills using a visual interface that represented a symphony orchestra and offered real-time feedback on hand gestures and facial expressions.

3.3.4. Types of the Visual Information and Their Functions

We identified eight categories for types of visual information applied in the included publications: (1) colours, (2) shapes, (3) 2D and 3D images, (4) text, (5) traditional notation, (6) plots, (7) sound waveforms, (8) videos.

In addition, 28 digital tools (44.44%) applied static visualisations, 33 (52.38%) incorporated both static and dynamic visualisations, and two applied dynamic visualisations exclusively (3.18%) (see Table 5). A substantial number of studies (n = 22; 34.92%) used dynamic visualisations to support the development of time-oriented skills such as synchronization and sight-reading. These visualisations took various forms, including piano rolls, moving cursors or trackers on notation or waveforms, rectangles moving toward the piano surface, or visuals that appeared and disappeared in time with the music. Some tools extended beyond synchronization. For example, X. Xiao et al. (2014) and X. Xiao et al. (2016) used human and animal-like figures to convey expressive timing through varied walking styles. Cheng's (2018) tool for piano learning visualised the teacher's hand postures to support precision in rubato. While most dynamic visualisations were system-generated and -controlled, some digital tools enabled learners to manipulate visual objects, such as a vehicle or drumsticks, during rhythm synchronization tasks. Other dynamic visualisations consisted of avatars, animated characters or images, and recorded videos that did not represent time-oriented musical features.

Table 5. Number of studies with static and dynamic visual information.

Visual Information	Number of Studies	%
Only static	28	44.44%
Static and dynamic	33	52.38%
Only dynamic	2	3.18%

Furthermore, we identified four functions of using visual information: instruction, feedback, interactive elements, and contextual elements. Appendix B provides examples of the different types of visuals and their associated functions.

Instruction

We classified the function of visualisations as *instruction* when they provided users with information that directly or indirectly guided task performance, including explicit cues about what, how, and/or when to act, as well as visual familiarization with instruments or interfaces intended to support subsequent interaction.

Colours were the predominant type of visualisation used to convey instructions across instruments, often mapped to individual musical notes (Alvarez-Molina et al., 2018; Amm et al., 2024; Baldassarri et al., 2016; Chin & Xia, 2022; Ling, 2016; Ludovico & Mangione, 2014; O'Toole et al., 2024; Paule-Ruiz et al., 2017; Perdana, 2014; X. Wei et al., 2023) or piano keys (Corintha et al., 2021; Trujano et al., 2018; Zeng et al., 2019). Deja et al.'s (2024) AR system for piano improvisation, for example, used yellow for harmonic grouping and pink to highlight notes that could be involved in improvisation. This colour strategy extended to various instruments, such as the guitar, with colours indicating fingering positions on the fretboard (Del Rio-Guerra et al., 2019); drums, where colours distinguished between right- and left-hand patterns (Pesek et al., 2024; Pinkl et al., 2024; Pinkl & Cohen, 2023); and trumpet, with colours indicating which register to play (Blewett & Gerhard, 2023). In a VR application for learning music theory, Stachurska et al. (2024) applied colour differentiation, assigning four distinct colours to represent the four voices (bass, tenor, alto, and soprano) both on the musical staff and through corresponding interactive cubes.

Geometric *shapes*, such as 2D or 3D rectangles, circles, cubes, and spheres, served diverse instructional functions across applications. Rectangles were the most prevalent, particularly in piano learning tools, where their position indicated which key to play and when (Amm et al., 2024; D. Lu, 2024), their length represented note duration, and their movement supported synchronization during performance (Fernandez et al., 2016; D. Lu, 2024; Trujano et al., 2018). Grid-based interfaces used square position and number to represent pitch (vertical) and duration (horizontal) (Franceschini et al., 2014; Oestermeier et al., 2015), or to support chord construction by placing notes within squares (Meneses Rodríguez & Sánchez, 2024). Other shape-based applications included rectangular shapes mirroring instrumental sections of an orchestra in a conducting tool (B. T. Wei et al., 2024); 3D shapes guiding pitch identification and navigation for the theremin (Johnson et al., 2020; Johnson & Tzanetakis, 2017); varied shapes such as squares and ovals for identifying different intervals (Ling, 2016); flattened spheres and squares indicating which valves to press, or whether no valves should be pressed (Blewett & Gerhard, 2023); and vertical and horizontal lines, similar to staff notation, illustrating pitch concepts, with vertical position representing pitch height (Paule-Ruiz et al., 2017).

The phrases *2D and 3D images* refer to visualisations that are more iconic than geometric shapes or that represent elements from the real world, such as humans, animals, or physical objects. In some cases, figurative images provided cues for learners to perform specific tasks. For instance, some tools used gesture-based imagery, such as virtual hands demonstrating teachers' and/or students' hand gestures (Amm et al., 2024; Cheng, 2018), or Curwen–Kodály hand signs (Hong et al., 2016). Game-based tools incorporated a narrative framework with animated cartoon characters (C.-L. Ho et al., 2023; Y. Lu et al., 2021; Martins et al., 2015; Micheloni et al., 2019). Martins et al. (2015), for instance, integrated images of bees and ropes to teach about duration and pitch. Furthermore, two studies used pictures of musical instruments to support children's development of knowledge about the instruments (C.-L. Ho et al., 2023; Tan & Lim, 2019).

Textual information was frequently used for instructional purposes. In several AR applications, note names (e.g., C and D) were overlaid on piano keys (Amm et al., 2024; Rigby et al., 2020; Trujano et al., 2018) or displayed within a virtual environment (Dascălu et al., 2014; Johnson et al., 2020; Johnson & Tzanetakis, 2017). Some gamified interfaces integrated text, such as note names, to teach music theory (Meneses Rodríguez & Sánchez, 2024; Paule-Ruiz et al., 2017; Perdana, 2014; Y.-H. Wang, 2023) or to help users understand what and how to play and, as such, enhance performance skills (Guclu et al., 2021). For example, one AR tool projected large 3D chord symbols in the classroom to teach harmony (Avanzini et al., 2022), while another used illuminated note names on a tangible device for teaching scales and triads (Klamka et al., 2019). In an MR-based drumming tool, the letters L and R indicated left- and right-hand movements (Pinkl et al., 2024). Beyond music, textual instructions were used for system navigation (Alvarez-Molina et al., 2018; Blanco et al., 2021; Y.-H. Wang, 2023; B. T. Wei et al., 2024) and for delivering instructional content (Dascălu et al., 2014; Nasrifan & Saidon, 2019).

Traditional notation was primarily used to support sight-reading in singing and instrumental performance (Amm et al., 2024; Chin & Xia, 2022; López-Calatayud & Tejada, 2023; Pérez-Gil et al., 2016). Most tools allowed learners to visually follow the notes in real time using a cursor or tracking elements. Two game-based applications supported users in learning both pitch and note placement on a staff (Alvarez-Molina et al., 2018; Y.-H. Wang, 2023). One interactive online tool used traditional notation as the main visual information for teaching music theory (Nasrifan & Saidon, 2019).

We defined a distinct category for *plots*, such as graphs, spectrograms, and spectra, as they convey information by combining text, shapes, and colours. We also specified a separate category for *sound waveforms*, which visually represent how a sound's amplitude changes over time. M. Ng and Schutz (2017) used both plots and waveforms to help students compare instrumental timbres. Similarly, L. Liu et al. (2016) integrated dynamic spectrograms into their digital tool to visually analyse the singing teacher's sound, allowing learners to compare their vocal output with the teacher's performance and imitate it as a reference model.

Finally, *videos*, defined as recorded visual materials of real-world scenes, were used for instructional purposes. Safian et al. (2024) integrated 360-degree performance videos into a VR setting to offer a first-person perspective and introduce Malaysian instruments and music. Similarly, B. Wang et al. (2021) incorporated instructional videos in a guitar learning system.

Feedback

The *feedback* function refers to digital tools' responses to users' actions by displaying performance-related information. Similar to instructional visuals, colours played a key role in visual feedback, often used to indicate correctness or precision during performance or other learning tasks (K. Lee, 2022; Molero et al., 2021; Pinkl et al., 2024; Rigby et al., 2020; Stachurska et al., 2024; Trujano et al., 2018).

Although *rectangular* shapes were primarily instructional, in one application they provided feedback, displaying the duration of the sound after being played (Chin & Xia, 2022).

Several tools used *plots*, *graphs*, and *spectrograms* to provide detailed visual feedback. For example, pitch accuracy and intonation were visualised in singing tools (Huang & Chu, 2016; Kumar et al., 2016; Pérez-Gil et al., 2016; X. Zhang, 2023), while spectrograms supported breath control, resonance, and articulation (L. Liu et al., 2016). Visual feedback on violin and guitar performance included pitch accuracy (T. K.-L. Ho et al., 2015), hand postures (Cheng, 2018), kinematics (Blanco et al., 2021), and progression (B. Wang et al., 2021). Some digital tools integrated waveforms, helping learners by visualising the acoustic features of their guitar playing (B. Wang et al., 2021) or singing (Kumar et al., 2016). C. Xiao (2020) combined spectrograms and waveforms to provide real-time feedback on performance accuracy across various instruments.

In some gamified digital tools, performance feedback was delivered through *textual information*, such as numerical scores or motivating phrases (e.g., "congratulations!") (Alvarez-Molina et al., 2018; Molero et al., 2021; O'Toole et al., 2024; Perdana, 2014; Pesek et al., 2024). *Figurative imagery* was also used. For example, a virtual 3D character began dancing when achieving a high score (Fernandez et al., 2016).

Interactive Elements

We classified visualisations as *interactive elements* when learners could use them to control the system, including manipulating or clicking objects or guiding an avatar.

Geometric shapes (e.g., rectangles, circles, cubes, and spheres), often combined with textual information or iconic images, mainly served as clickable buttons in user interfaces (e.g., Fletcher et al., 2019; Haferkamp & Dengel, 2021; López-Calatayud & Tejada, 2023; Meneses Rodríguez & Sánchez, 2024; Pérez-Gil et al., 2016; Tan & Lim, 2019; X. Wei et al., 2023). In immersive environments, virtual coloured cubes could be manipulated to place notes on a virtual staff (Amm et al., 2024; Stachurska et al., 2024).

The *2D and 3D images* often included avatars or other virtual objects (e.g., drumsticks) that learners controlled for gaming or for exploring a virtual environment (Alvarez-Molina et al., 2018; Delegido et al., 2024; Feng et al., 2022; Haferkamp & Dengel, 2021; Paule-Ruiz et al., 2017; Pesek et al., 2024; Pinkl & Cohen, 2023; Y.-H. Wang, 2023). Some tools enabled users to play virtual instruments (Blewett & Gerhard, 2023) or manipulate virtual objects (Dascălu et al., 2014).

In one digital tool, sound waveforms allowed learners to scroll through a performance timeline (B. Wang et al., 2021), while another let users manipulate spectral representations to visually explore sound characteristics of different instruments (M. Ng & Schutz, 2017).

In some cases, *colours* played a role in the interaction. For example, in a collaborative game, users voted on the validity of each other's moves using green and red buttons (Meneses Rodríguez & Sánchez, 2024).

Contextual Elements

Some visualisations in user interfaces did not serve as direct instructional aids but instead established narrative or situational *context*, primarily in gamified or immersive environments. For example, background visuals such as roadside trees (Feng et al., 2022; Pesek et al., 2024; Y.-H. Wang, 2023), a forest with animals (X. Wei et al., 2023), or a 2D character appearing across different scenes (Y. Lu et al., 2021) were used to enhance game appeal and engagement.

Other digital tools recreated real-life music settings, such as a concert hall or a music classroom, to invoke a sense of presence and immersion in a performance context (Blewett & Gerhard, 2023; Das et al., 2017). In a metaverse-based tool, for instance, users explored a virtual concert hall to learn about a symphonic orchestra and its instruments (Delegido et al., 2024).

Contextual elements were also used metaphorical. Hong et al. (2016), for instance, paired Curwen-Kodaly hand signs (e.g., mi hand sign) with animated natural scenes (e.g., calm but moving ocean) to evoke both the expressive quality of tones and its corresponding hand posture.

4. Discussion

This scoping review provided a comprehensive overview of how visual information is applied in digital technologies for music education. Addressing three research questions, the findings cover the general characteristics of the included studies (Q1), the methodological features of the empirical research (Q2), and the key characteristics of the original digital tools developed for teaching and/or learning music (Q3). Below, we present the discussion structured around five key findings.

4.1. Key Findings

4.1.1. Persistent Methodological Limitations

Consistent with earlier reviews, (e.g., Caamaño Liñares et al., 2023; Kardeş, 2022), surveys and were the most common methods and many studies demonstrate limitations in methodological rigour. Analysis of the included studies reveal prevalence of convenience sampling and recurrent limited reporting on sample characteristics, procedures, and data collection methods. In addition, many studies relied on descriptive analysis, hindering interpretability and generalizability of their findings. Some studies claimed participant involvement without providing basic methodological information (e.g., number of participants, data collection methods, and data analysis). Taken together, these limitations constrain the validity, reliability and generalizability of the reported findings (Andrade, 2021; Robb et al., 2025).

4.1.2. Narrow Focus on Individual Performance Skills

Most digital tools aimed at supporting the development of performance skills, with limited attention to other music learning objectives, such as creating, responding, connecting, and conducting. While several studies introduced tools for different instruments and for singing, the piano was the most frequently targeted instrument. This dominance of the piano was also evident within the *creating* category, with half of the tools addressing piano improvisation.

Most digital tools designed to improve musical *responsiveness* focused on internal responding, i.e., awareness and recognition of sound characteristics. In contrast, tools addressing external responsiveness centred almost exclusively on synchronization during performance. This narrow interpretation is noteworthy, as responding to music can take multiple forms (Bauer, 2014). In particular, bodily movement during music listening is a fundamental mode of musical responsiveness (Abril & Gault, 2016). Our findings therefore suggest that current visual interfaces in music educational technologies address a limited range of learning goals, often neglecting dimensions such as creativity, expression, and embodied understanding. A more holistic approach would support learners across the full spectrum of educational objectives, including performing, creating, responding, and connecting, as well as broader dimensions of musical understanding (Abril, 2024; Bauer, 2014). Nijs and Behzadaval (2024) highlight how emerging technologies, such as motion tracking and real-time interaction, can promote embodied engagement and creative exploration through visual modalities. For example, tools that incorporate body movements (Ilsar & Hughes, 2020) or drawings (Sen et al., 2020) to sound illustrate how visual information can support embodied and creative learning.

4.1.3. Lack of Solid Pedagogical Foundations

Analysis of the conceptual background across the reviewed studies revealed that, among the thirteen identified categories, the most frequently recurring categories were: interactive learning, feedback in learning, multimodal learning, individual learning, game-based learning, and embodied learning. These were explicitly either framed as conceptual and pedagogical foundations or applied in the tool design. Though analytically distinct, they sometimes overlap. For example, visual and haptic feedback, multimodality, and embodied learning all involve the integration of multiple sensory modalities (e.g., visual, haptic, and auditory) yet varied how they were conceptualized and operationalized. Visual and haptic feedback typically referred to system responses, while embodied, interactive, and multimodal learning approaches emphasized active, multisensory learner involvement.

Embodied learning, in particular, extended beyond physical interaction to include visual representations. For instance, X. Xiao et al. (2014) and X. Xiao et al. (2016) employed figurative gestures to support learners understanding of musical expression. Hong et al. (2016) integrated embodied metaphors through 3D natural scenes aligned with Curwen-Kodály hand gestures. In such cases, visual information served to externalize musical embodiment.

Most digital tools prioritised technological features supporting feedback and individual learning, which can enhance learner autonomy, especially in contexts where a teacher is not present. This emphasis often came at the expense of collaborative learning, a vital dimension of music education (Cangro, 2016). Interestingly, visual interfaces for collaborative music-making are more frequently explored in research conducted outside the domain of teaching and learning, offering nevertheless valuable insights for the future development of collaborative digital tools for music education (Nijs & Behzadaval, 2024; see also Hopkins et al., 2022; Martin & Tørresen, 2017; Men et al., 2019).

Notably, some studies did not explicitly articulate a pedagogical foundation to underpin their work. Moreover, design rationales referring to concepts such as feedback, interactivity, individual learning, or multimodality, often focused on technological features rather than grounding their work in established pedagogies or other relevant theories. This limitation, echoed in previous reviews (e.g., Caamaño Liñares et al., 2023), may limit the educational effectiveness of these digital tools, as robust pedagogical grounding is essential for guiding meaningful learning and sustained experiences (Bauer, 2014).

4.1.4. Underuse of Design Principles

Few studies reported the use of formal design principles (e.g., HCI, cognitive load theory, gamification). This is noteworthy as design principles are widely recognised as fundamental in educational technology research. They guide the design and development of robust and pedagogically grounded tools (du Preez & Jacobs, 2025; Herrington & Reeves, 2011). In addition, pedagogical design principles grounded in learning theory (e.g., Galperin, UDL, cognitive load) turn digital environments into mediators of deeper learning and learner agency, not just content delivery (Engeness, 2021). This absence of explicit reporting limits the transparency and hinders the replicability of design processes in educational technology research.

4.1.5. Expanding Technological Possibilities

While screen-based tools appear to still dominate, an increasing number of studies explore immersive technologies such as AR, VR, and MR. Computers continue to play a central role as display devices in both screen-based and AR applications. However, we

https://doi.org/10.3390/educsci16020309

found that AR technologies use a wide range of display devices, including head-mounted displays (HMDs), mobile devices, projected or installed lights, and computers. In particular, the increased implementation of HMDs reflects a trend towards hands-free, movement-based interaction, allowing learners to engage with music in more physically immersive ways. These technologies open up promising avenues for integrating visual and musical elements within embodied learning environments and for designing collaborative learning experiences that extend beyond the affordances of traditional screen-based approaches (Nijs & Behzadaval, 2024).

4.1.6. Limited Application of Visual Information

The analysis of visual information in music educational technologies revealed recurring patterns concerning types and functions of music. Colours play a key role in instruction and feedback, particularly for pitch recognition and performance accuracy. Modern technologies extend these uses through augmenting real-world environments (e.g., piano keys) and real-time feedback. Similarly, plots and spectrograms, enabled by computational analysis, were predominantly used as feedback to support technical precision. Arguably, this prevailing focus on feedback and accuracy reinforces a behaviourist approach to music learning and teaching (Rabbani et al., 2023). Moreover, it often comes at the expense of creativity, musical expression, and musical responsiveness. Recent work (e.g., Nijs, 2018; X. Xiao et al., 2014; Yang et al., 2024) suggests that integrating visuals with movement and music can support a constructivist approach and promote embodied and creative learning. Such perspectives highlight the importance of balancing the corrective and expressive functions of visual information in music education technologies, guiding future designs toward fostering creativity, responsiveness, and deeper musical understanding.

Our analysis identified two primary modes of applying visual information: static and dynamic. Most digital tools used dynamic visuals to represent time-based musical features, particularly for synchronization. Some, however, extended this use to include expressiveness, posture, and interactive engagement, often grounded in bodily movement. Such examples point to the educational potential of dynamic visuals, especially when tied to body movement, to move beyond their role as mere synchronization aids and instead function as mediators that connect perceptual, motor, and cognitive dimensions of music learning. Indeed, according to Fortuna (2017), when embodiment, visual representation, and sound are combined in a coherent learning environment, they enhance musical understanding, expression, and engagement.

4.2. Limitations

While this scoping review provides a comprehensive analysis of the application of visual information in digital music education technologies, several limitations should be acknowledged. First, only articles published in English were included, which may have excluded relevant research reported in other languages. Second, the review focuses on the literature published between 2014 to 2024, meaning that earlier studies that could offer valuable information were not considered.

In addition, while this review systematically categorised types of visualisations and their functions, it did not evaluate the effectiveness of different visualisation strategies in relation to music learning outcomes. Furthermore, the classification of visual information was based primarily on images of the digital tools and the limited description provided by the authors. As a result, some visual features may not have been documented in the original publications and were therefore not considered in the classification process.

https://doi.org/10.3390/educsci16020309

5. Conclusions and Recommendations for Future Research

This scoping review examined how visual information is applied in digital technologies for music education. Based on sixty-three publications published between 2014 and 2024, it mapped current practices and identified key limitations in how visual information is conceptualised, designed and pedagogically embedded. The findings indicate a growing interest in, predominantly, interactive and real-time visual feedback systems, with visualisations such as colours, shapes, animations, plots and waveforms, and figurative representations to enhance learners' understanding and use of pitch, rhythm, dynamics, and form. At the same time, the review highlights persistent gaps between technological capabilities and pedagogical implementation.

Collectively, the findings point to several priorities for advancing the field. First, the persistent methodological limitations (key finding 1) underscore an urgent need for substantially improved methodological rigour, including transparent sampling strategies, systematic measurement, and appropriate analytical approaches to strengthen the evidence base.

Second, the narrow focus on individual performance skills (key finding 2) and the weak pedagogical grounding (key finding 3) suggest that many digital tools prioritise performance monitoring over holistic musical development. Addressing this requires a shift from technology-driven innovation toward pedagogy-led design that support creativity, expression, collaborative music making, and embodied responsiveness. Such pedagogy-led design should be informed by established theoretical frameworks from musicology, and music psychology and pedagogy, as well as historically grounded educational and artistic explorations such as Tobin's colour system, Boal-Palheiros' musicograms, or Truslit's motion curves. In addition, a pedagogy-driven approach to the design of digital technologies needs to consider different educational levels (e.g., beginner vs. advanced) and contexts (e.g., formal, informal).

Third, the limited and inconsistent use of formal design principles (key finding 4) highlights the need for more systematic application of frameworks from human computer interaction (HCI), cognitive load theory, and gamification frameworks with explicit documentation of design decisions and their pedagogical.

Fourth, existing digital overemphasise feedback for accuracy at the expense of creativity and expression (key finding 6), using visual information within a more behaviourist approach rather than as a catalyst for constructivist, embodied learning. This finding reveals a critical research gap in understanding how different types of visual information and interaction modes affect the cognitive, emotional, and physical dimensions of music learning. Robust empirical investigation is needed to develop evidence-based design principles that balance monitoring with expressive or creative functions and complement existing design frameworks (key finding 4). In addition, an effective integration of visual digital tools into teaching and learning requires appropriate professional development and training of educators, including knowledge acquisition on the appropriate use of different types of visualisation and interaction modes.

Finally, the expanding technological possibilities (key finding 5), particularly immersive technologies such as AR, VR, and MR, offer promising opportunities that remain underexplored. Both software solutions (e.g., computational analysis) and hardware devices (e.g., HMDs, motion sensors) create new possibilities for embodied and collaborative learning experiences. Future research could include comparative studies investigating optimal conditions for learning across diverse technological platforms, moving beyond screen-based digital tools to explore how immersive environments can support dynamic visualisations connected to bodily movement and collaborative music making.

https://doi.org/10.3390/educsci16020309

In sum, while digitised visual information holds significant potential to enrich music education, realising this potential requires a shift from technology-driven explorations to a pedagogy-driven innovation, embracing a holistic approach to music learning, fostering creative, embodied, and collaborative music educational approaches alongside individual technical skill acquisition.

Author Contributions: Conceptualization, B.B. and L.N.; methodology, B.B., L.S.M. and L.N.; software, B.B. and L.S.M.; validation, B.B. and L.N.; formal analysis, B.B., L.S.M. and L.N.; investigation, B.B.; resources, B.B. and L.N.; data curation, B.B.; writing—original draft preparation, B.B.; writing—review and editing, B.B. and L.N.; visualization, B.B.; supervision, L.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available on request.

Conflicts of Interest: The authors report there are no competing interests to declare.

Abbreviations

The following abbreviations are used in this manuscript:

AR	Augmented Reality
VR	Virtual Reality
MR	Mixed Reality
Q	Question
WoS	Web of Science
PRISMA-ScR	Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
HMD	Head Mounted Display

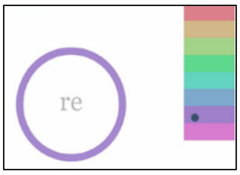
Appendix A

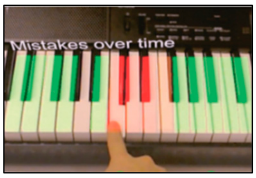
Category	Keywords in Thematic Analysis	No.	Publications
Interactive learning	interactive software, interactive visualization, interaction techniques, interactive experience, interactive multimedia, interactive systems, interactive environment, interactive visual music education software, interactive visual music pedagogy, interactive music education tool, interactive lessons, interaction techniques, interaction modes, tangible interaction, interactive music tangibles, interactive online application, interactive learning environment, interactive multimedia, interactive instructional design, interactive tutorials, interactive video tutorials, interactive games, interactive content, interactive design, interactive digital musical instrument, interactive user feedback, interactivity, interactive methods, interactivity space, interactive visual feedback, musical interactions, interactive learning, interactive music learning materials, interactive computing, interactive piano education, interactive elements, interactive theory area, interactive learning experiences, interactive modules, interactive space, interact	29	Schönfeld et al. (2014); Perdana (2014); Franceschini et al. (2014); Baldassarri et al. (2016); Martins et al. (2015); Ling (2016); Das et al. (2017); Fletcher et al. (2019); Klamka et al. (2019); Nasrifan and Saidon (2019); Pedersen et al. (2020); Guclu et al. (2021); B. Wang et al. (2021); Y. Lu et al. (2021); Corintha et al. (2021); Feng et al. (2022); Chin and Xia (2022); Avanzini et al. (2022); K. Lee (2022); Blewett and Gerhard (2023); C.-L. Ho et al. (2023); Y.-H. Wang (2023); X. Wei et al. (2023); Amm et al. (2024); Safian et al. (2024); B. T. Wei et al. (2024); Stachurska et al. (2024); Deja et al. (2024); Delegido et al. (2024)

Category	Keywords in Thematic Analysis	No.	Publications
Feedback in learning	visual feedback, immediate visual feedback, real-time visual feedback, real time feedback, feedback in learning, instant feedback, haptic feedback, audio-visual-haptic feedback	27	Oestermeier et al. (2015); Huang and Chu (2016); Pérez-Gil et al. (2016); Kumar et al. (2016); Fernandez et al. (2016); Johnson and Tzanetakis (2017); Trujano et al. (2018); Cheng (2018); Del Rio-Guerra et al. (2019); Micheloni et al. (2019); Johnson et al. (2020); C. Xiao (2020); Rigby et al. (2020); Blanco et al. (2021); B. Wang et al. (2021); Y. Lu et al. (2021); Molero et al. (2021); Corintha et al. (2021); Chin and Xia (2022); K. Lee (2022); López-Calatayud and Tejada (2023); Pinkl and Cohen (2023); Blewett and Gerhard (2023); Pinkl et al. (2024); Amm et al. (2024); Pesek et al. (2024); B. T. Wei et al. (2024)
Multimodal learning	multi-sensory and multidimensional music experiences, multi-modal perceptions, multi-modal body-syntonic interactive systems, multimodality, multiple modalities in learning, multi-sensory experience, multimodal immersive experience, multimodal integration, auditory-motor theory, multi-modal environments, multimodal interaction, Visual-Aural-Kinesthetic (VAK) learning style model	18	Perdana (2014); X. Xiao et al. (2014); Keebler et al. (2014); Martins et al. (2015); X. Xiao et al. (2016); Baldassarri et al. (2016); Ling (2016); Trujano et al. (2018); Alvarez-Molina et al. (2018); Fletcher et al. (2019); Johnson et al. (2020); Pedersen et al. (2020); Chin and Xia (2022); C.-L. Ho et al. (2023); Pinkl et al. (2024); Pesek et al. (2024); O'Toole et al. (2024); B. T. Wei et al. (2024)
Individual learning	self-development, self-regulated learning, self-directed learning, individualized sight-singing learning, self-assessment, self-evaluation, self-listening, self-analysis, self-learning, self-directed learning, self-practice, self-efficacy, self-teaching, self-study	14	Dascălu et al. (2014); T. K.-L. Ho et al. (2015); X. Xiao et al. (2016); Huang and Chu (2016); Pérez-Gil et al. (2016); Cheng (2018); Del Rio-Guerra et al. (2019); Nasrifan and Saidon (2019); Pedersen et al. (2020); Rigby et al. (2020); López-Calatayud and Tejada (2023); Amm et al. (2024); Pesek et al. (2024); Stachurska et al. (2024)

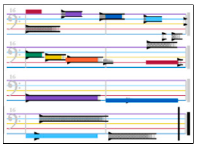
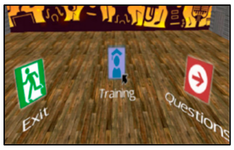
Category	Keywords in Thematic Analysis	No.	Publications
Game-based learning	game-based learning, Learning Mechanics-Game Mechanics, gamification, Game-Based Music Education, gamification techniques, gamification approach, Educational Videogames, gamified learning	11	Paule-Ruiz et al. (2017); Alvarez-Molina et al. (2018); Micheloni et al. (2019); Haferkamp and Dengel (2021); Molero et al. (2021); Feng et al. (2022); Avanzini et al. (2022); Y.-H. Wang (2023); Meneses Rodríguez and Sánchez (2024); Amm et al. (2024); Pesek et al. (2024)
Embodied learning	embodiment, embodied music cognition, embodied metaphors, embodied musical expression, embodied cognition, Virtual Co-embodiment, the Inspect, Embody, Invent design framework	10	X. Xiao et al. (2014); Keebler et al. (2014); X. Xiao et al. (2016); Hong et al. (2016); Alvarez-Molina et al. (2018); Corintha et al. (2021); Avanzini et al. (2022); Pinkl and Cohen (2023); Pinkl et al. (2024); B. T. Wei et al. (2024)
Immersive learning	immersive learning, immersive teaching, immersive experience, Immersive Education, immersive interactive virtual environment	9	Johnson et al. (2020); Haferkamp and Dengel (2021); Y. Lu et al. (2021); Blewett and Gerhard (2023); Safian et al. (2024); Delegido et al. (2024); Pesek et al. (2024); O'Toole et al. (2024); Stachurska et al. (2024)
Tangible interfaces in learning	tangible interaction, Tangible Interfaces, Tangible User Interfaces	5	Oestermeier et al. (2015); Baldassarri et al. (2016); Klamka et al. (2019); Micheloni et al. (2019); Y. Lu et al. (2021)
Music learning theories	Dalcroze Eurhythmics, Gordon's theory, Curwen-Kodaly (Kodaly method), Gordon's music learning theory	4	X. Xiao et al. (2016); Pérez-Gil et al. (2016); Hong et al. (2016); Y.-H. Wang (2023)
Computer assisted learning	computer-mediated solutions, computer-assisted music instrument tutoring (CAMIT)	4	Dascălu et al. (2014); Johnson and Tzanetakis (2017); Johnson et al. (2020); K. Lee (2022)
Collaborative learning	collaborative learning, collaborative music, collaborative music education	4	Franceschini et al. (2014); Oestermeier et al. (2015); Baldassarri et al. (2016); Meneses Rodríguez and Sánchez (2024)
Visual learning	visual learning, interactive visual music pedagogy	3	T. K.-L. Ho et al. (2015); Ling (2016); Amm et al. (2024)
Multimedia in learning	multimedia animation, interactive multimedia, multimedia learning	3	L. Liu et al. (2016); Nasrifan and Saidon (2019); D. Lu (2024)

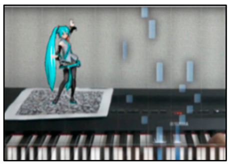
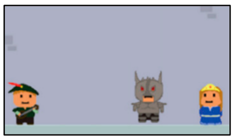
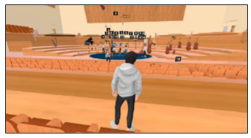
Appendix B

<table>
<tr><th></th><th>Instruction</th><th>Feedback</th><th>Interactive Elements</th><th>Contextual Elements</th></tr>
<tr><td>Colours</td><td>

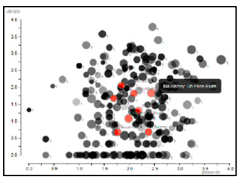
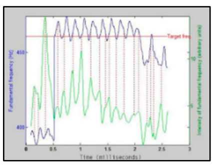
Perdana (2014)</td><td>

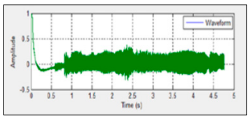
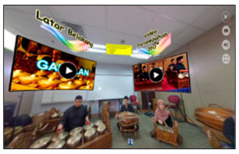
Trujano et al. (2018)</td><td>

Meneses Rodríguez and Sánchez (2024)</td><td>—</td></tr>
<tr><td>Shapes</td><td>
D. Lu (2024)</td><td>
Chin and Xia (2022)</td><td>

Pedersen et al. (2020)</td><td>—</td></tr>
<tr><td>2D and 3D images</td><td>
X. Xiao et al. (2014)</td><td>
Fernandez et al. (2016)</td><td>
Alvarez-Molina et al. (2018)</td><td>
Delegido et al. (2024)</td></tr>
<tr><td>Text</td><td>

Y.-H. Wang (2023)</td><td>

Alvarez-Molina et al. (2018)</td><td>
Martins et al. (2015)</td><td>—</td></tr>
<tr><td>Traditional notation</td><td>
Pérez-Gil et al. (2016)</td><td>—</td><td>—</td><td>—</td></tr>
<tr><td>Plots</td><td>
Schönfeld et al. (2014)</td><td>
T. K.-L. Ho et al. (2015)</td><td>—</td><td>—</td></tr>
</table>

	Instruction	Feedback	Interactive Elements	Contextual Elements
Sound waveforms	 M. Ng and Schutz (2017)	 Kumar et al. (2016)	 B. Wang et al. (2021)	
Videos	 Safian et al. (2024)	—	—	—

References

Abril, C. R. (2024). A holistic approach to music education. In *The sage handbook of school music education* (pp. 295–324). Torrossa.

Abril, C. R., & Gault, B. M. (2016). *Teaching general music: Approaches, issues, and viewpoints*. Oxford University Press.

Aksoy, Y. (2023). Seeing sounds: The effect of computer-based visual feedback on intonation in violin education. *International Journal of Education and Literacy Studies*, *11*(2), 2–12. [CrossRef]

Alvarez-Molina, K. A., Jöllenbeck, A., & Malaka, R. (2018). Listen and play: Auditory-motor interaction in a bard's tale video game. In *Immersive learning research network: 4th international conference, iLRN 2018, Missoula, MT, USA, June 24–29, 2018, Proceedings 4* (pp. 198–210). Springer.

Amm, V., Chandran, K., Engeln, L., & McGinity, M. (2024). Mixed reality strategies for piano education. *Frontiers in Virtual Reality*, *5*, 1397154. [CrossRef]

Andrade, C. (2021). The inconvenient truth about convenience and purposive samples. *Indian Journal of Psychological Medicine*, *43*(1), 86–88. [CrossRef] [PubMed]

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, *8*(1), 19–32. [CrossRef]

Avanzini, F., Baratè, A., Cottini, M., Ludovico, L. A., & Mandanici, M. (2022). Learning tonal harmony through augmented reality: Bridging the gap between music embodiment and digital experiences. In *International conference on Computer-human interaction research and applications* (pp. 54–73). Springer.

Baldassarri, S., Marco, J., Bonillo, C., Cerezo, E., & Beltrán, J. R. (2016). ImmertableApp: Interactive and tangible learning music environment. In *Human-computer interaction. Novel user experiences: 18th international conference, HCI international 2016, Toronto, ON, Canada, July 17–22, 2016. Proceedings, Part III 18* (pp. 365–376). Springer.

Bauer, W. (2014). *Music learning today: Digital pedagogy for creating performing and responding to music* (1st ed.). Oxford University Press. [CrossRef]

Blanco, A. D., Tassani, S., & Ramirez, R. (2021). Real-time sound and motion feedback for violin bow technique learning: A controlled, randomized trial. *Frontiers in Psychology*, *12*, 648479. [CrossRef] [PubMed]

Blewett, D. J., & Gerhard, D. (2023). Brass haptics: Comparing virtual and physical trumpets in extended realities. *Arts*, *12*(4), 145. [CrossRef]

Boal-Palheiros, G., & Wuytack, J. (2006, August 22–26). *Effects of the 'musicogram' on children's musical perception and learning*. 9th International Conference on Music Perception and Cognition (pp. 1264–1271), Bologna, Italy.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. [CrossRef]

Caamaño Liñares, T., Rodríguez Rodríguez, J., Castro Rodríguez, M., & Marín Suelves, D. (2023). Digital didactic resources and music: Mapping the last decade of research. *Music Education Research*, *25*(4), 351–366. [CrossRef]

Cage, J. (1969). *Notations*. Something Else Press.

Cangro, R. (2016). Student collaboration and standards-based music learning: A literature review. *Update: Applications of Research in Music Education*, *34*(3), 63–68. [CrossRef]

Cheng, M. (2018). Introducing motion-capturing technology into the music practice room as a feedback tool for working towards the precision of rubato. *Journal of Music, Technology & Education*, *11*(2), 149–170.

Chin, D., & Xia, G. (2022). A computer-aided multimodal music learning system with curriculum: A pilot study. In *NIME 2022*. The University of Auckland.

Corintha, I., Outeiro, L., Dias, R., & Bernardes, G. (2021). *AM-I-BLUES: An interactive digital music instrument for guiding novice pianist in the improvisation of jazz melodies. Advances in design, music and arts* (Vol. 9, pp. 689–698). Springer.

Das, S., Glickman, S., Hsiao, F. Y., & Lee, B. (2017, May 15–18). *Music everywhere–augmented reality piano improvisation learning system*. International Conference on New Interfaces for Musical Expression (NIME) (pp. 511–512), Copenhagen, Denmark.

Dascălu, M.-I., Coman, M., Postelnicu, R., & Nichifor, C. (2014). Learning to play a musical instrument in adulthood–Challenges and computer-mediated solutions. *Procedia-Social and Behavioral Sciences*, *142*, 311–317. [CrossRef]

Deja, J. A., Štor, S., Pucihar, I., Pucihar, K. Č., & Kljun, M. (2024). Teach me how to ImproVISe: Co-designing an augmented piano training system for improvisation. *arXiv*, arXiv:2402.02999.

Delegido, J. M. A., Almagro, B. G., & López, M. R. (2024). Metaverse as a higher education tool: Notes from musical curricula at early childhood education and primary school education bachelor' s degrees. *Harmonia: Journal of Arts Research and Education*, *24*(2), 376–385. [CrossRef]

Del Rio-Guerra, M. S., Martin-Gutierrez, J., Lopez-Chao, V. A., Flores Parra, R., & Ramirez Sosa, M. A. (2019). AR graphic representation of musical notes for self-learning on guitar. *Applied Sciences*, *9*(21), 4527. [CrossRef]

Dunbar, L. (2017). Enhancing content understanding through animated listening map development. *General Music Today*, *30*(2), 25–28. [CrossRef]

du Preez, I., & Jacobs, L. (2025). Pedagogical and visual design principles for online learning material: Insights from quality guiding documents in diverse educational contexts. *Interactive Learning Environments*, 1–17. [CrossRef]

Engeness, I. (2021). Developing teachers' digital identity: Towards the pedagogic design principles of digital environments to enhance students' learning in the 21st century. *European Journal of Teacher Education*, *41*(1), 96–114. [CrossRef]

Feng, Y., Li, S., & Wei, Y. (2022). Design of a game-based learning environment for basic rhythmic patterns in music. In *Third international conference on computer science and communication technology (ICCSCT 2022)* (Vol. 12506, pp. 1343–1348). SPIE.

Fernandez, C. A. T., Paliyawan, P., Yin, C. C., & Thawonmas, R. (2016). Piano learning application with feedback provided by an AR virtual character. In *2016 IEEE 5th global conference on consumer electronics* (pp. 1–2). IEEE.

Fletcher, C., Hulusic, V., & Amelidis, P. (2019). Virtual reality ear training system: A study on spatialised audio in interval recognition. In *2019 11th international conference on virtual worlds and games for serious applications (VS-Games)* (pp. 1–4). IEEE.

Fortuna, S. (2017). Embodiment, sound and visualization: A multimodal perspective in music education. *Zbornik Radova Akademije Umetnosti*, *5*(5), 120–131. [CrossRef]

Franceschini, A., Laney, R., & Dobbyn, C. (2014, September 14–20). *Learning musical contour on a tabletop*. Joint ICMC/SMC 2014 Conference, Athens, Greece.

Fürst, D., Miller, M., Keim, D. A., Bonnici, A., Schäfer, H., & El-Assady, M. (2020). Augmenting sheet music with rhythmic fingerprints. In *2020 IEEE 5th workshop on visualization for the digital humanities (VIS4DH)* (pp. 14–23). IEEE.

Gorbunova, I., & Hiner, H. (2019). Music computer technologies and interactive systems of education in digital age school. In *International conference communicative strategies of information society (CSIS 2018)* (pp. 124–128). Atlantis Press.

Grahn, J. A. (2012). See what I hear? Beat perception in auditory and visual rhythms. *Experimental Brain Research*, *220*, 51–61. [CrossRef] [PubMed]

Guclu, H., Kocer, S., & Dundar, O. (2021). Application of augmented reality in music education. *The Eurasia Proceedings of Science Technology Engineering and Mathematics*, *14*, 45–56. [CrossRef]

Haferkamp, M., & Dengel, A. (2021, May 17–June 10). *Learning organology in virtual environments: A pilot study for virtual music education*. 2021 7th International Conference of the Immersive Learning Research Network (iLRN) (pp. 1–5), Eureka, CA, USA.

Hamond, L. F., Welch, G., & Himonides, E. (2019). The pedagogical use of visual feedback for enhancing dynamics in higher education piano learning and performance. *Opus*, *25*(3), 581–601. [CrossRef]

Herrington, J., & Reeves, T. C. (2011). Using design principles to improve pedagogical practice and promote student engagement. In *Proceedings ASCILITE 2011-changing demands, changing directions*. Murdoch University.

Hills, C. S. (2011). *Graphic notation as means of musical gesture: Examining percussion works by john cage, Morton Feldman, and Vinko Globokar* [Doctoral dissertation, University of Kansas].

Ho, C.-L., Lin, T.-G., & Chang, C.-R. (2023). Interactive multi-sensory and volumetric content integration for music education applications. *Multimedia Tools and Applications*, *82*(4), 4847–4862. [CrossRef]

Ho, T. K.-L., Lin, H., Chen, C.-K., & Tsai, J.-L. (2015). Development of a computer-based visualised quantitative learning system for playing violin vibrato. *British Journal of Educational Technology*, *46*(1), 71–81. [CrossRef]

Hong, M. H., Hicks, W. S., & Horn, M. S. (2016). MiLa: An audiovisual instrument for learning the curwen hand signs. In *2016 CHI conference extended abstracts on human factors in computing systems* (pp. 1691–1697). Association for Computing Machinery.

Hope, C. (2017). Electronic scores for music: The possibilities of animated notation. *Computer Music Journal*, *41*(3), 21–35. [CrossRef]

Hopkins, T., Weng, S. C. C., Vanukuru, R., Wenzel, E. A., Banic, A., Gross, M. D., & Do, E. Y.-L. (2022). AR drum circle: Real-time collaborative drumming in AR. *Frontiers in Virtual Reality*, *3*, 847284. [CrossRef]

Hu, P., Gao, B., & Li, K. (2024). Learning interpreting in virtual reality: A scoping review. *Interactive Learning Environments*, *33*(1), 347–366. [CrossRef]

Huang, Y. T., & Chu, C. N. (2016). Visualized comparison as a correctness indicator for music sight-singing learning interface evaluation—A pitch recognition technology study. In *Frontier computing: Theory, technologies and applications* (pp. 959–964). Springer.

Ilsar, A., & Hughes, M. (2020). A new audio-visual gestural instrument for unlocking creativity. In *Extended abstracts of the 2020 CHI conference on human factors in computing systems* (pp. 1–4). Association for Computing Machinery.

Johnson, D., Damian, D., & Tzanetakis, G. (2020). Evaluating the effectiveness of mixed reality music instrument learning with the theremin. *Virtual Reality*, *24*(2), 303–317. [CrossRef]

Johnson, D., & Tzanetakis, G. (2017, May 15–18). *Vrmin: Using mixed reality to augment the theremin for musical tutoring*. International Conference on New Interfaces for Musical Expression (pp. 151–156), Copenhagen, Denmark.

Kardeş, B. (2022). Investigation of publications on the use of technology in music education. *Journal of Educational Technology and Online Learning*, *5*(4), 795–807. [CrossRef]

Keebler, J. R., Wiltshire, T. J., Smith, D. C., Fiore, S. M., & Bedwell, J. S. (2014). Shifting the paradigm of music instruction: Implications of embodiment stemming from an augmented reality guitar learning system. *Frontiers in Psychology*, *5*, 471. [CrossRef] [PubMed]

King, A., Himonides, E., Ruthmann, A., & Ruthmann, S. A. (2017). *The Routledge companion to music, technology, and education*. Routledge.

Kivijärvi, S. (2019). Applicability of an applied music notation system: A case study of Figurenotes. *International Journal of Music Education*, *37*(4), 654–666. [CrossRef]

Klamka, K., Wojnar, J., & Dachselt, R. (2019). Scaledial: A novel tangible device for teaching musical scales & triads. In *Extended abstracts of the 2019 CHI conference on human factors in computing systems* (pp. 1–4). Association for Computing Machinery.

Kumar, A., Chandra, M., & Agarwal, S. (2016). GUI implementation of real time feedback system to learn singing. In *2016 IEEE international conference on recent trends in electronics, information & communication technology (RTEICT)* (pp. 1051–1054). IEEE.

Lam, C. K. (2024). Technology-enhanced creativity in K-12 music education: A scoping review. *International Journal of Music Education*, *42*(4), 691–703. [CrossRef]

Lee, K. (2022). Designing an intelligent learning system for practicing the Oboe embouchure. In *Adjunct proceedings of the 2022 ACM international joint conference on pervasive and ubiquitous computing and the 2022 ACM international symposium on wearable computers* (pp. 278–283). Association for Computing Machinery.

Lee, P. N. (2013). Self-invented notation systems created by young children. *Music Education Research*, *15*(4), 392–405. [CrossRef]

León-Garrido, A., Barroso-Osuna, J. M., & Llorente-Cejudo, C. (2022). Conceptual cartography for the systematic study of music education based on ICT or EdTech. *Societies*, *12*(5), 136. [CrossRef]

Li, R., Kamaruzaman, F. M., & Nasri, N. M. (2025). The research trends of music general education: A bibliometric review. *Edelweiss Applied Science and Technology*, *9*(8), 427–439. [CrossRef]

Lima, H. B., Santos, C. G. D., & Meiguins, B. S. (2021). A survey of music visualization techniques. *ACM Computing Surveys (CSUR)*, *54*(7), 1–29. [CrossRef]

Lin, F., & Yunus, M. M. (2024). A systematic review of metaverse-based learning in music education. *International Journal of Advanced Research in Future Ready Learning and Education*, *37*(1), 41–58. [CrossRef]

Ling, Y. (2016). Interactive visual music education software for children. In *20th international academic mindtrek conference* (pp. 453–456). Association for Computing Machinery.

Liu, C., Hwang, G.-J., Tu, Y., Yin, Y., & Wang, Y. (2023). Research advancement and foci of mobile technology-supported music education: A systematic review and social network analysis on 2008–2019 academic publications. *Interactive Learning Environments*, *31*(7), 4535–4554. [CrossRef]

Liu, L., Pang, Y., & Hu, Z. (2016). Application of spectrogram analysis in traditional vocal music teaching and multimedia animation vocal music teaching. *International Journal of Emerging Technologies in Learning*, *11*(11), 64. [CrossRef]

López-Calatayud, F., & Tejada, J. (2023). Self-efficacy in instrumental practice with intonation assessment software: A multiple case study with novice violin and viola students. *Journal of Music, Technology & Education*, *16*(1–2), 57–76.

Lu, D. (2024). The impact of augmented reality on motor skills in piano education: Improving the left-hand technique with AR piano. *International Journal of Human–Computer Interaction*, *41*(13), 8398–8408. [CrossRef]

Lu, Y., Huang, T., Liu, J., & Gong, J. (2021). Design of children's entertainment and education products based on AR technology. In *International conference on human-computer interaction* (pp. 292–301). Springer.

Ludovico, L. A., & Mangione, G. R. (2014). *Self-regulation competence in music education.*. ERIC.

Ma, Y., & Wang, C. (2025). Empowering music education with technology: A bibliometric perspective. *Palgrave Communications*, *12*(1), 1–14. [CrossRef]

Martin, C. P., & Tørresen, J. (2017). Microjam: An app for sharing tiny touch-screen performances. In *NIME* (pp. 495–496). Aalborg University.

Martins, V. F., Gomes, L., & de Paiva Guimarães, M. (2015). Challenges and possibilities of use of augmented reality in education: Case study in music education. In *International conference on computational science and its applications* (pp. 223–233). Springer.

Men, L., Bryan-Kinns, N., & Bryce, L. (2019). Designing spaces to support collaborative creativity in shared virtual environments. *PeerJ Computer Science*, *5*, e229. [CrossRef] [PubMed]

Meneses Rodríguez, C. P., & Sánchez, J. (2024). Implementation and usability evaluation of an online videogame for learning musical harmony. In *International conference on human-computer interaction* (pp. 82–93). Springer.

Merchán Sánchez-Jara, J. F., González Gutiérrez, S., Cruz Rodríguez, J., & Syroyid Syroyid, B. (2024). Artificial intelligence-assisted music education: A critical synthesis of challenges and opportunities. *Education Sciences*, *14*(11), 1171. [CrossRef]

Micheloni, E., Tramarin, M., Rodà, A., & Chiaravalli, F. (2019). Playing to play: A piano-based user interface for music education video-games. *Multimedia Tools and Applications*, *78*, 13713–13730. [CrossRef]

Molero, D., Schez-Sobrino, S., Vallejo, D., Glez-Morcillo, C., & Albusac, J. (2021). A novel approach to learning music and piano based on mixed reality and gamification. *Multimedia Tools and Applications*, *80*(1), 165–186. [CrossRef]

Muresan, R. E., & Pop, A. S. (2017). Visual feedback as a modern method of vocal training. *Information and Communication Technology in Musical Field*, *8*(2), 81–90.

Nasrifan, M. N. H., & Saidon, Z. L. H. (2019). Designing online interactive application of learning music theory in blended learning mode. *International Journal of Recent Technology and Engineering*, *7*(6), 226–233.

Nettheim, N., & Becking, G. (1996). How musical rhythm reveals human attitudes: Gustav Becking's theory. *International Review of the Aesthetics and Sociology of Music*, *27*(2), 101–122. [CrossRef]

Ng, H. H. (2019). Project stomp: Using graphic notation for effective multimodal collaborative learning. *General Music Today*, *33*(3), 5–13. [CrossRef]

Ng, M., & Schutz, M. (2017). Seeing sound: A tool for teaching music perception principles. *Canadian Acoustics*, *45*(3), 104–105.

Nijs, L. (2018). Dalcroze meets technology: Integrating music, movement and visuals with the music paint machine. *Music Education Research*, *20*(2), 163–183. [CrossRef]

Nijs, L., & Behzadaval, B. (2024). Laying the foundation for augmented reality in music education. *IEEE Access*, *12*, 100628–100645. [CrossRef]

Nijs, L., & Leman, M. (2013, July 17–20). *Interactive technologies in the instrumental music classroom: A case study with the music paint machine*. 6th Conference of the European Network of Music Educators and Researchers of Young Children (pp. 117–133), The Hague, The Netherlands.

Novotná, E., & Ašenbrenerová, I. (2021). Visualization of music and its application in the process of education. *Gramotnost, Pregramotnost a Vzdělávání*, *4*(3), 25–34.

Oestermeier, U., Mock, P., Edelmann, J., & Gerjets, P. (2015). LEGO music: Learning composition with bricks. In *14th international conference on interaction design and children* (pp. 283–286). Association for Computing Machinery.

O'Toole, P., Mancini, M., Pitt, I., & Maye, L. (2024). Chromaesthetica: A virtual reality app for learning musical theory using sound-colour cross-modal associations. In *International conference on mobile and ubiquitous multimedia* (pp. 513–515). Association for Computing Machinery.

Pace, I. (2007). Graphic notation. In *Lecture*. University of London Institutional Repository.

Paule-Ruiz, M., Álvarez-García, V., Pérez-Pérez, J. R., Álvarez-Sierra, M., & Trespalacios-Menéndez, F. (2017). Music learning in preschool with mobile devices. *Behaviour & Information Technology*, *36*(1), 95–111.

Pedersen, K., Hulusic, V., Amelidis, P., & Slattery, T. (2020). Spatialized audio in a custom-built OpenGL-based ear training virtual environment. *IEEE Computer Graphics and Applications*, *40*(5), 67–81. [CrossRef]

Perdana, I. (2014). Teaching elementary school students new method of music performance with Leap Motion. In *2014 international conference on virtual systems & multimedia (VSMM)* (pp. 273–277). IEEE.

Pesek, M., Hirci, N., Žnideršič, K., & Marolt, M. (2024). Enhancing music rhythmic perception and performance with a VR game. *Virtual Reality*, *28*(2), 118. [CrossRef]

Pérez-Gil, M., Tejada, J., Morant, R., & De Martos, A. P.-G. (2016). Cantus: Construction and evaluation of a software solution for real-time vocal music training and musical intonation assessment. *Journal of Music, Technology & Education*, *9*(2), 125–144.

Pinkl, J., & Cohen, M. (2023). VR drumming pedagogy: Action observation, virtual co-embodiment, and development of drumming "halvatar". *Electronics*, *12*(17), 3708. [CrossRef]

Pinkl, J., Villegas, J., & Cohen, M. (2024). Multimodal drumming education tool in mixed reality. *Multimodal Technologies and Interaction*, *8*(8), 70. [CrossRef]

Rabbani, L. M., Alarabi, K. S., Alsalhi, N. R., & Sa'deh Al Marashdi, H. (2023). Behaviourist vs. constructivist views of feedback: Provisions and implications. *Arts Educa*, *36*, 197–206.

Rigby, L., Wünsche, B. C., & Shaw, A. (2020). piARno-an augmented reality piano tutor. In *32nd Australian conference on human-computer interaction* (pp. 481–491). Association for Computing Machinery.

Robb, S. L., Story, K. M., Harman, E., Burns, D. S., Bradt, J., Edwards, E., Golden, T. L., Gold, C., Iversen, J. R., Habibi, A., Johnson, J. K., Lense, M., Perkins, S. M., & Springs, S. (2025). Reporting guidelines for music-based interventions checklist: Explanation and elaboration guide. *Frontiers in Psychology*, *16*, 1552659. [CrossRef]

Roels, J. M., & Van Petegem, P. (2014). The integration of visual expression in music education for children. *British Journal of Music Education*, *31*(3), 297–317. [CrossRef]

Safian, A. R., Saidon, Z. L., Nasrifan, M. N., Djahwasi, H. R., Sampurno, M. B. T., Mazlan, C. A. N., & Koning, S. I. (2024). Design and development of e-MARZ: A practical guide to implementing cost-effective VR experiences for music education. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, *40*(1), 129–145. [CrossRef]

Sanzone, D. (2018). *Teaching the colourstrings violinist: Student development through and beyond colourstrings* [Ph.D. thesis, University of Tasmania].

Sánchez-Marroquí, J., & Vicente-Nicolás, G. (2024). Bibliometric study of scientific production in music education from 2000 to 2022. *International Journal of Music Education*, *43*(3), 372–388. [CrossRef]

Schönfeld, F., Berndt, A., Hähnel, T., Pfleiderer, M., & Groh, R. (2014). Vocalmetrics: An interactive software for visualization and classification of music. In *9th audio mostly: A conference on interaction with sound* (pp. 1–8). Association for Computing Machinery.

Sen, S., Tahiroğlu, K., & Lohmann, J. (2020). Sounding brush: A tablet based musical instrument for drawing and mark making. In *International conference on new interfaces for musical expression* (pp. 331–336). Birmingham City University.

Sethares, W. A., & Bañuelos, D. (2007). *Rhythm and transforms* (Vol. 1). Springer.

Sköld, M. (2019, October 14–18). *Visual representation of musical rhythm in relation to music technology interfaces—An overview*. 14th International Symposium on Computer Music Multidisciplinary Research (p. 725), Marseille, France.

Soysal, F. (2012). Teaching music theory in an effective and amusing way: An approach to teach music theory by painting. *Review of Higher Education & Self-Learning*, *5*(14), 36.

Spence, C., & Di Stefano, N. (2025). Augmenting art crossmodally: Possibilities and pitfalls. *Frontiers in Psychology*, *16*, 1605110. [CrossRef] [PubMed]

Stachurska, W., Witoszek-Kubicka, A., & Igras-Cybulska, M. (2024). Harmosphere VR: Enhancing harmonic learning in music schools through virtual reality. *CSEDU*, *1*, 765–771.

Strayer, H. R. (2013). From neumes to notes: The evolution of music notation. *Musical Offerings*, *4*(1), 1. [CrossRef]

Şen, E. (2021). An analysis on music maps with their symbolic and graphic representations. *Egitim ve Bilim*, *46*(207), 439–463. [CrossRef]

Tan, K., & Lim, C. (2019). Malaysian Music Augmented Reality (MMAR): Development of traditional musical instruments using augmented reality. *International Journal of Innovative Technology and Exploring Engineering*, *8*(4S), 340–345.

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., . . . Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, *169*(7), 467–473. [CrossRef]

Trujano, F., Khan, M., & Maes, P. (2018). ARPiano efficient music learning using augmented reality. In *Innovative technologies and learning: First international conference, ICITL 2018, Portoroz, Slovenia, August 27–30, 2018, Proceedings 1* (pp. 3–17). Springer.

Verschaffel, L., Reybrouck, M., Janssens, M., & Van Dooren, W. (2009). Using graphical notations to assess children's experiencing of simple and complex musical fragments. *Psychology of Music*, *38*(3), 259–284. [CrossRef]

Wang, B., Yang, M. Y., & Grossman, T. (2021). Soloist: Generating mixed-initiative tutorials from existing guitar instructional videos through audio processing. In *2021 CHI conference on human factors in computing systems* (pp. 1–14). Association for Computing Machinery.

Wang, T., & Yu, X. (2024). Practical analysis of multi-modal teaching behavior in elementary school music singing game teaching. *Arts Educa*, *41*, 87–102.

Wang, Y.-H. (2023). Exploring the effects of using various designs of game-based materials on music learning. *Interactive Learning Environments*, *31*(5), 2650–2664. [CrossRef]

Weatherly, K. I. C. H., Wright, B., & Lee, E. Y. (2024). *Digital game–based learning in music education: A systematic review between 2011 and 2023*. Research Studies in Music Education. [CrossRef]

Wei, B. T., Almeda, S. G., Tam, E., & Abrahamson, D. (2024). Demonstration of sympathetic orchestra: An interactive conducting education system for responsive, tacit skill development. In *Adjunct proceedings of the 37th annual ACM symposium on user interface software and technology* (pp. 1–3). Association for Computing Machinery.

Wei, X., Qiu, R., & Liu, R. (2023, December 4–8). *Using augmented reality to facilitate music learning for preschool children*. International Conference on Computers in Education, Matsue, Japan.

Xiao, C. (2020). MADSA: Music accuracy development through spectral analysis. In *2020 ACM southeast conference* (pp. 204–210). Association for Computing Machinery.

Xiao, X., Puentes, P., Ackermann, E., & Ishii, H. (2016). Andantino: Teaching children piano with projected animated characters. In *15th international conference on interaction design and children* (pp. 37–45). Association for Computing Machinery.

Xiao, X., Tome, B., & Ishii, H. (2014, June 30–July 3). *Andante: Walking figures on the piano keyboard to visualize musical motion*. International Conference on New Interfaces for Musical Expression (pp. 629–632), London, UK. [CrossRef]

Yang, M., Llano, M. T., & McCormack, J. (2024). Exploring real-time music-to-image systems for creative inspiration in music creation. *arXiv*, arXiv:2407.05584.

Zeng, H., He, X., & Pan, H. (2019). FunPianoAR: A novel AR application for piano learning considering paired play based on multi-marker tracking. *Journal of Physics: Conference Series*, *1229*(1), 012072. [CrossRef]

Zhang, X. (2023). Multi-sensory research of singing visualization in pre-school music education. *Interactive Learning Environments*, *31*(5), 2830–2840. [CrossRef]

Zhang, Y., Fen, B. W., Zhang, C., & Pi, S. (2024). Transforming music education through artificial intelligence: A systematic literature review on enhancing music teaching and learning. *International Journal of Interactive Mobile Technologies*, *18*(18), 76–93. [CrossRef]

Zhou, L., Seah, S. P., Zhang, M., Chen, C. W., & Zhu, K. J. (2026). Graphic notation activities in early childhood music teaching. *Educational and Humanities*, *2*(2), 1–12. [CrossRef]

https://doi.org/10.3390/educsci16020309