Review

How Technology Applied to Music-Therapy and Sound-Based Activities Addresses Motor and Social Skills in Autistic Children

Grazia Ragone *, Judith Good and Kate Howland

Department of Informatics, University of Sussex, Falmer BN1 9QJ, UK; j.good@sussex.ac.uk (J.G.); k.l.howland@sussex.ac.uk (K.H.)
* Correspondence: g.ragone@sussex.ac.uk

Abstract: Autism affects how people perceive and make sense of the world around them. Autism is a spectrum condition which impacts people in different ways. Also referred to as Autism Spectrum Disorder (ASD), it is characterized by challenges in the domains of social, cognitive and motor functioning, which differ in severity. Previous research suggests that music can have cognitive, psycho-social, behavioural, and motor benefits in this population. We systematically review the use of technology in Music-therapy and related sound-based activities to improve the motor and social skills of children. In May 2020 we conducted a systematic search on Music-therapy and musical activities for autistic children in research databases including Science Direct, APA PsycNet, Cochrane, IEEE and Web of Science, to collect relevant studies. We initially collected 5,179 papers of which only 27 studies were identified as suitable for the scope of this review. In the paper, we analyse and describe key characteristics of each project. We then highlight the commonalities, strengths and limitations of existing work, and identify implications for future interaction design.

Keywords: autism spectrum disorder; autism; motor skills; social skills; music-therapy technology; motion caption; interactivity; literature review; accessibility; user experience

1. Introduction

a spectrum of challenges, ranging from differences in the areas of communication and social interaction to the manifestation of repetitive interests and behaviours [1]. According to the Diagnostic and Statistical Manual of Mental Disorders (DSM-5, the handbook used by health care professionals in the US and much of the world as the authoritative guide to the diagnosis of mental conditions), Autism Spectrum Disorder (ASD) is characterized by two core features: deficits in social communication and social interaction, and restricted, repetitive patterns of behaviour, interests, or activities [1]. This first feature, deficits in social communication and social interaction, makes it more difficult for autistic people to predict the actions of others [2]. The DSM-5 manual [1] also states that motor deficits are common, especially an odd gait, clumsiness and walking on tiptoes, however, they are not considered to be a core trait of autism, and many practitioners, including doctors, do not consider motor problems to be closely associated with autism [3]. Nonetheless, research by Bhat and colleagues [4] suggested that a high percentage (87%) of the autistic population have motor difficulties, ranging from an atypical gait (gross motor skills) to difficulties with handwriting (fine motor skills). Furthermore, autistic children present motor differences such as clumsiness, abnormal motor coordination, postural insecurity, and weak performance in standardized testing of motor functioning [5]. When motor difficulties are present, they significantly constrain learning and social interaction, key aspects of a child’s growth [6]. Despite some interesting studies [7–9], research on this
topic is still not clear and under-investigated. For example, some researchers suggest that motor planning problems, or dyspraxia, may be a core deficit of autism or a sign of the nervous system differences underlying autism [10].

Some autistic children can face social barriers to physical activities which may put them at risk for a sedentary lifestyle and later obesity [11]. An autistic child with motor impairment (such as postural insecurity) may, for example, have a limited and different motor experience with respect to their counterparts in terms of running, jumping, and ball throwing [12]. Those are usually skills which are grounded in the postural control system, in the age of 7–10 years [13]. Wodka and colleagues [14] recently conducted a comparison study between autistic children, typically developing children, and children with attention deficit hyperactivity disorder, using a test of motor skills. They found that independent experts who were not aware of the children’s diagnosis rated performance on balance and catching a ball significantly lower for autistic children in comparison with the other groups. Researchers argue that autistic children put less emphasis on what they see when trying to catch a ball, and more emphasis on input from their own muscles.

Imitation is another area in which autistic children can face difficulties when copying others either through actions with objects, with gestures and body movements or with sounds or words. Imitation is related both to social and motor domains. For some authors [15,16] it is strictly connected with motor development and indirectly with the improvement of social skills. However, it is still not clear whether the numerous difficulties in social interaction that typically persist throughout adolescence and adulthood [1,17] may be affected by motor differences or vice versa. Early social differences may contribute, with several other factors, to the way children go through the milestones of imitation for instance. These discrepancies impact mental and physical development, learning, and behavioural functioning across settings and may contribute to the fact that even high functioning individuals have difficulty contributing to the workforce in adulthood [17,18], although this is highly speculative. In addition, the differences in social and motor skills seem to be related to other core indicators of autism such as language as well as to motor and synchronization differences in autistic individuals [19,20]. Given the evidence supporting the interplay between the motor and social domains [21], further investigations are essential to examine the potential value of motor skills to improve social skills in autistic children.

Music can have a powerful impact on autistic people because it functions as a vehicle of interaction, social and emotional development without the necessary involvement of verbal communication [22]. The experience of music is cross-modal and involves several senses at the same time [23]. Research shows that children with autism usually display musical skills in line with those of typically developing children [14]. Music-therapy has been shown to be helpful for autistic children to improve their communicative, motor and social skills as well as contributing to increasing their social adaptation skills and to furthering the quality of parent-child relationships [24]. Moreover, music is predictable in its rhythmic structure and Music-therapy can adapt to children’s rhythms and emotional needs in a flexible way [25]. Hence, Music-therapy can provide a pleasant and non-frightening means to stimulate social interactive skills [24,26]. Music-therapy is an approved approach by the World Health Organisation (WHO) for autistic children [27], and in the last decade research evidence supporting its effectiveness has grown [28]. Music based interventions are being utilized for autistic children to support the development of their social communication and motor skills [29–32]. In particular, listening and synchronising to rhythm, which provides repetition and structure in music, has displayed great potential in improving the social-communication and motor activities of autistic people [31].

The purpose of this systematic literature review is to review the use of technology in Music-therapy and related sound-based practice to improve the motor and social skills of autistic children and to identify and critique definitions of social and motor skills used to guide future practice and research. We also included systems which cannot be considered
part of Music-therapy; but which use auditory and musical feedback with the aim of creating an interactive environment beneficial for autistic people. To date, we identified several literature reviews exploring the use of music with autism [29,32–34]. The previous reviews investigated intervention of traditional Music-therapy following the original definition of Music-therapy as ‘the clinical and evidence-based use of music interventions to accomplish individualized goals within a therapeutic relationship’ [35].

This review shows how digital technology has been contributing to the change and evolution of interventions where music and sounds are still used as a channel of communication and expression, although not always matching the original and traditional description of Music-therapy. Hence, for the current review, we also considered interventions where music was used as a way to interact with children, outside of a formal music therapeutic structure [36–38].

In the next section, we will briefly report some neurological explanation of how music and sounds affect the brain of autistic individuals. Also, we will introduce Music-therapy with some historical notes and a partial account of its practice as there are many theoretical/practical approaches that we are unable to cover in this context. Then, we make a link with the technology which has definitely contributed to the change of Music-therapy, offering a wider scope for different forms of music making.

2. Background

In this section, we describe Music-therapy approaches and the challenge of demonstrating its efficiency, which motivates the use of recording and analysis technology. We then discuss the challenges faced by some autistic children, which motivates the development of new digital instruments. In the last decade, empirical evidence in the medical literature, in addition to qualitative and experiential stances, increasingly support the case for the impact of music on autistic children [22]. Music allows people to communicate thoughts, emotions, feelings and comprehend those subtle distinctive human experiences that cannot easily be communicated through words [22]. Music is composed by melody, harmony, tempo, dynamics, timbre, form, and rhythm [39]. Rhythm has been shown to be very useful when adopted as timekeeper in the therapeutic application of music for motor rehabilitation goals or auditory-motor synchronization [40–42]. Rhythm plays a crucial role in learning, development, and performance, as timing of movement is important in many motor control and cognitive functions [43,44]. Rhythm supports the regulation of movement patterns in time, space, and force dimensions to facilitate movement execution [45].

Music-therapy may have the potential to improve the lives of autistic people by giving them new means to express, communicate and discover new resources in a world in which they too often find themselves labelled, excluded and marginalized. Music-therapy is used in several therapeutic contexts with autistic people because it facilitates non-verbal communication and interaction as well as for the music itself as being similar to language [46]. Some investigations report case studies of individuals who exhibited a preference to express themselves in songs rather than verbalized words [47,48]. In support of the studies suggesting the heightened responsiveness of autistic children to musical and language-like sounds rather than verbal utterances [49], some parents have reported autistic children who will not speak but can sing instead [47]. This could be explained by differences in neural activation between speech and song perception in autistic children [49,50]. There is some evidence that the predisposition of songs to speech stems from the fact that music and language develop in different areas of the brain. Usually, in the typically developing (TD) brain of a non-musician, language use and its interpretation occurs in the left hemisphere while the interpretation of music occurs in the right hemisphere. Studies with brain images through electroencephalogram (EEG) reported that autistic individuals show decreased activity in the left hemisphere but increased in the right hemisphere [51,52]. This may explain a few reports from parents that their child can sing but not speak [47].
Spoken language has several variables shaping the prosody, which interplay with each other such as pitch, loudness, tempo, and rhythm in the speech which express different meanings. Autistic people find prosody difficult to hear, understand and reproduce [53]. It is widely reported that autistic people often exhibit music processing skills equal or superior to their typically developing counterparts [54]. Within the autism community, there are a few musical prodigies [55,56]. Although it is rare, musical savantism represents another example of enhanced auditory processing in autism. Savant individuals often show a virtuosistic capacity to encode and remember passages, for example they can play back or repeat whole musical sections, impeccably reproducing complex melodic, harmonic, and rhythmic pieces, after listening only once [57]. More generally, there is good evidence of intact music processing abilities in autistic people [58]. Research shows that autistic children enjoy listening to and making music just as much, if not more than, their typically developing peers [22,58].

The Music-therapy literature describes two broad categories of music interventions: active and receptive. The active approach typically uses musical instruments, which the client (child or adult) and the therapist play, alternating synchronous and asynchronous patterns [59]. The music generated by the dyad or the group develops as a communication between the interacting parts. In contrast, the receptive approach involves listening to live or recorded music to enhance mood, or lessen the stress, pain or anxiety felt by the client. The active approach presents opportunities for non-verbal embodied communication, and in that way fosters the development of the client’s agency and independence, primarily by improving their body interaction and social skills [60]. One of the most recognised active methods of Music-therapy is Musical Improvisation, which is a complex cognitive method that merges creativity, goal-directed action, sensory monitoring, and social communication [61]. Moreover, the music therapeutic active approach usually uses acoustic melodic and rhythmic instruments as well as singing [62]. Music-therapy has been often considered as part of occupational therapy interventions. There are several methodologies employed in Music-therapy, some with solid bases demonstrating their efficacy [24,62]. Music-therapy uses music and rhythm to improve sensorimotor, synchronization and communication skills. Music-therapy also supports children in dealing with anger, aggression, and frustration in several ways [63]. For this purpose, several techniques have been developed to help increase self-esteem and challenge self-aggression. Gold and colleagues [32] describe Music-therapy as a scientific and systematic intervention where the therapist endorses wellbeing through musical experiences. The term Music-therapy appeared for the first time in 1789 in a Columbian publication titled ‘Music Physically Considered’ and the first record of a systematic experiment in Music-therapy dates from the 1800s [64].

There is an ongoing demand for clinical accountability in assessing the effectiveness of Music-therapy interventions for autistic children, outcomes reviewed in this paper provide evidence for music as a powerful tool for intervention [22]. There are several effective methods adopted in Music-therapy where rhythm is central to facilitating positive change in motor output in autistic children. One of the most known and recognised is Neurological Music Therapy (NMT) which includes twenty standardized techniques that have been adopted successfully to treat some aspects of neurodevelopmental disorders like autism [65]. The twenty techniques are divided into three areas of specialization: sensorimotor, speech and language, and cognitive training [66]. Three sensorimotor techniques specifically target motor skills: Rhythmic Auditory Stimulation (RAS) [67], Patterned Sensory Enhancement (PSE) [68,69], and Therapeutic Instrumental Music Performance (TIMP) [70]. The efficacy of music and sounds, with the component of rhythm, to improve motor skills in autistic children has been documented by several studies [66–69]. There are also studies [42,65,70,71] that support the claim that rhythm can engage brain’s area for movement. Nevertheless, there is a need for further research in this area, and to date, there has only been one randomised control trial on a musical intervention addressing motor abilities in autism [71].
Although some acoustic instruments, such as drums, offer an accessible way to allow someone with limited skills to interact, other instruments, such as the guitar, can sometimes be too complicated for those with cognitive or physical deficits, limiting their scope for expression [62,72,73]. The use of digital technology has made it possible to overcome these constraints, with the design of ‘easy to use’ multisensory musical surfaces and environments which enable natural interaction for persons with all kinds of special needs, including severely autistic people, for whom motor development may be highly compromised [74]. Digital technology has the potential to make activities such as music-making more accessible [28,75]. Many autistic people have been drawn towards digital devices for play, learning and communication [76]. Technologies enhance opportunities for autistic individuals to improve their lives, support their learning, facilitate communication, foster recording, and support data collection during interventions [76]. Furthermore, in the last decades, the use of technologies has improved the assessment and diagnosis of autism [77] by quantifying differences and outcomes throughout or between sessions. The attractiveness of most technology stems from its predictability, and the fact it does not demand social convention be followed or require language skills [74]. Therefore, therapists may use it to enhance dynamic engagement and make it configurable for each client as well as use it to monitor and record the changes happening during the intervention while they focus on the relationship with their client [31]. This suggests that studying how technology can be used in Music-therapy is an important avenue of research. Although there are various systems available, there is a shortage of studies in the Music-therapy literature which investigate the application of digital technologies to address social and motor skills in children. This review aims to highlight the commonalities and the strengths of the selected applications to maximise the benefits of future interaction design.

2.1. Music-Therapy and Technology

There is a clear distinction between approaches that use technology to study and reflect on therapy sessions and those that use technology actively within sessions to enhance the therapy. However, some practitioners and researchers use technology in both of these ways. In fact, to meet the challenges of gathering evidence about effectiveness while carrying out interventions, therapists started integrating technologies into their practices, using software such as Amstar [77] and Emtek [78,79] to assess the quality of their methods or for video recording and analysing the Music-therapy improvisation sessions at a later stage, thus allowing therapists to focus on interaction during the session [80].

Music and auditory environments implemented using virtual reality [81–83] (Table 1), software applications [84–89] (Table 2), machine learning [90,91] and robotics [92–94] (Table 3) have highlighted the benefits of combining multi-sensory stimuli to boost expression through sounds and music. These new forms of intervention also include therapeutic goals such as improved motor control, social-engagement, body awareness and creativity [95,96] which is this review’s focus.

Based on the assumption that digital music instruments might support therapeutic goals, research in Human-Computer Interaction (HCI) has investigated the coupling of movement with sound generation. This includes the appropriation of machine learning techniques which has enabled more precise recognition of movements, supporting interaction based on the full body. On the direct interaction side, HCI has investigated multimodal technology, robots, etc. to enhance the synchronization, the interaction of people, and in the particular case of autism to promote independence, expand communication, and increase social interactions.

The rest of the paper is structured as follows: in Section 3 we describe the method used to select the studies for this review explaining the criteria for exclusion and selection, then in Section 4 we analyse each of the relevant selected systems, aiming to understand best practice by studying the main characteristics of these studies and their goals and participants. In Section 5 we discuss our interpretation of the results and, finally, Section 6 draws conclusions with guidelines to design and develop future systems.
3. Materials and Methods

3.1. Study Selection

We structured our process following recommended practices for systematic reviews: search, study selection, data extraction, data analysis, and interpretation [97]. As part of the selection process, we studied the main characteristics of the studies and their participants and excluded those which did not meet our criteria. We first identified our research question which was to find all the existing technological systems, with a potential of being adopted in a Music-therapy context, addressing motor and/or social skills in autistic children. In the selected studies we included only one study adult participants. In our search we also included studies which were not classified as Music-therapy but where the use of sounds or music was a relevant part of the investigation.

We conducted a research strategy that considered the terms and keywords that characterize the research question, included in the titles and keywords. The databases used were Scopus search across different publishers (Web of Science, IEEE, Science Direct and All Cochrane).

To increase search sensitivity and ensure satisfactory search retrieval, we used text words, synonyms, keywords and spelling variations, which were combined using Boolean operators. The following search terms and sequences were used: “autistic disorder” OR “autism spectrum disorder” OR “autism” OR “disorder, autistic” OR “Autism condition” OR “condition, autistic” OR “ASD” AND “music therapy” OR “sound therapy” AND “motor skills” AND “social skills published between 1980 and 2021.

We performed two rounds of study selection: (1) the first round included papers based on their keywords and abstracts, and (2) the second round included papers based on whether they described an appropriate system related to our query.

3.1.1. First round of Selection

First, titles, keywords, and abstracts of found articles were examined to exclude papers which did not include autistic children (or the diagnosis was not specified), and papers not including a technological system or not reporting full details of studies. In the first exclusion process, including the initial search terms, the first author read the abstracts of the 5179 papers to select studies for potential inclusion in the review. The initial selection included 4646 papers from PsychNet, 190 from Web of Science, 12 from Cochrane trials, 18 from Science Direct and 93 from IEEE. Also, 220 papers resulted from a wider web search. Papers were excluded either because they were out of scope, i.e. they included systems or technology for activities unrelated to Music-therapy and related interventions (n = 1724) or were posters or workshop papers (n = 3421) (Figure 1).

![Figure 1. Prisma Process Followed for search strategy query.](image-url)
3.1.2. Second Round of Selection

We reviewed the abstracts of all the studies where technology was used in Music-therapy and related sound-based interventions for autistic children, and we excluded studies where the focus of the intervention was not motor or social skills and not autism. This includes studies with active (where music was played using instruments or sung) or receptive (where participants listened to live or recorded music) Music-therapy as well as related interventions where sounds were used to interact with autistic children. The abstracts were then examined, and the final result highlighted 33 studies where music was used with autistic children, (or in one case, with young adults), to address motor and/or social skills in their treatment. We then identified and removed further duplicates which left a total of 27 studies meeting all inclusion criteria. In total, 5152 publications were rejected because they did not meet the eligibility criteria. There was no restriction based on language, though, the main language found in the literature was English.

4. Results

4.1. Main Areas of Classification

We list the main areas in which the studies selected for this review can be categorised as (a) Virtual Environments, (b) Interactive Multimodal Systems, (c) Robots. These categories, as explained in the following sub-sections, often overlap in terms of technology, and some systems in each category make use of machine learning techniques, as discussed below.

Virtual environments (VE), including virtual reality, augmented reality [98], games and simulations, make possible the learning and expression of human movement in augmented environments. For various authors [99,100], the use of Virtual Reality (VR) offers several advantages over other technologies for interventions aiming to improve social interaction in autistic children, because they provide an attractive interface that promotes a high level of motivation [101,102]. One explanation for high engagement of children with VR finds support in the concept of aesthetic resonance, which defines a condition where individuals are so immersed in the virtual reality (with an audio-visual control system) that the physical obstacles faced in generating the movement are forgotten [101]. Virtual environments provide unique opportunities for autistic children to learn and practice skills in a controlled replicable setting. However, not all the virtual environments use the same technology, hence, the level of immersion differs across the settings [102].

Interactive multimodal systems (IMS) usually process two or more combined user input modes such as movement, touch, speech, visual and learning in a coordinated manner with multimedia system output which can provide tactile, auditory, and visual feedback and give a greater sense of agency and autonomy to clients in the therapeutic environment. Recently, neuroimaging studies [103] have demonstrated that musical activity activates a multimodal network of brain areas such as hearing, movement, emotion, memory expanding such effects to other nonmusical domains through structural and functional changes in the brain [104]. The use of multimodal systems (with the real time interplay of sound and image) may widen the efficacy of sessions where music already gives a great contribution [87]. A study by Ringland and colleagues [105] showed that the combinations of multimodal interactions (multi-sensorial and multi-interaction design) with physical objects enhance sensory integration. In recent years, multimodal systems, involving audio-visual and tangible user interfaces, have been increasing popularity in a Music-therapy context [106].

Robots have been reported to be successful with autistic children as they inhabit a space somewhere between a toy and a human [107]. Robots have been employed in therapeutic settings to develop social communication and motor skills in autistic children [108]. Their use proved to be effective with autistic children as they are predictable and quite simple compared to the complexities of a human being [109]. Hence, robots have
been utilized particularly to address interventions on shared attention and as an intermediary to stimulate social interactions [92,110]. Starting in 1998, and running for more than 10 years, the “Aurora Project” used robots to help autistic children. Starting with ‘Labo-1’ researchers on the Aurora project created several sorts of robot to interact with the children: ‘Pekee’, an upgraded version of Labo-1; ‘Robota’, a humanoid doll robot which has an infrared sensor and some rotatable joints; ‘KASPAR’ (Kinesics and Synchronization in Personal Assistant Robotics), a humanoid robot with more than 11 moveable joints, tactile sensors, and rotatable eyeballs. KASPAR was designed to boost elementary social interaction skills in autistic children, adopting strategies of turn-taking and imitation games [111].

Machine learning techniques such as (i) gesture recognition and (ii) regression can enable the interpretation of movements and make virtual reality and Human Robot Interaction (HRI) even more attractive and engaging [112]. The main goal of gesture recognition is to create a system which can recognize specific human gestures and use them to convey information or for device control, whilst regression approaches can be used to map between motion and sounds in novel and creative ways [113–115].

4.2. Analysis and Interpretation

Our aim is not to derive best practices for all types of evaluation from this corpus but to summarize and describe the key system features and reported findings of each of them. Of the 27 selected studies, 12 of them were designed to be integrated in the Music-therapy practice while the other 16 studies, do not strictly fit in music therapeutic practice. We divided the studies into three main areas according to system type: 1. Music-Interaction and Virtual Environments – VE, 2. Interactive Multimodal Systems, 3. Music Interaction with Robots (Human Robot Interaction – HRI). We now describe the systems we found, going through the analysis and description of each one. In the following tables we list a brief overview of the studies selected, grouped in each of the three main categories created for this review. We report for each study 1) technological features; 2) Users interaction; 3) Outcomes; and 4) Aims.

4.2.1. Music-Interaction with Virtual Environments

Virtual and mixed reality environments (VE) use computer systems to create immersive, interactive environments. A virtual environment is a networked common operating space where users can interact with both the computing environment and the space of other users. VR environments can be experienced by multiple players simultaneously, through headsets which include stereoscopic displays, stereo sound, and head motion tracking sensors.

Table 1. Systems adopting Virtual Environments (VE).

<table>
<thead>
<tr>
<th>Systems Overview</th>
<th>Technological Features</th>
<th>User Interaction</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cymis and Kinect [115]</td>
<td>The system comprises a monitor display, a PC, a MIDI sound source, speakers, and several interfaces, including a touch panel, a switch, and an expiratory pressure sensing device.</td>
<td>Used in Music-therapy. Users touch the main interface, (the touch panel), pointing to the notes. Kinect permits to “play” the music.</td>
<td>4 music therapy sessions lasting 15 minutes. N= 2 children, one autistic, and the second not yet formally diagnosed. Improved the sensory integration, motor coordination, and emotional release.</td>
</tr>
<tr>
<td>Pictogram Room [82,83]</td>
<td>It uses Kinect and an Open Scene Graph (OSG) shader that facilitates combining the objects and the user’s movements and integrating them with the virtual world.</td>
<td>The child works through educational activities and games, controlling a virtual avatar with their movements learning self-awareness, body schema, postures, communication, specific gestures and imitation.</td>
<td>Study 1: Improvement of self-recognition and body schema (participants N= 20 TD and 5 autistic). Study 2: analyzed its effectiveness in improving the joint attention of participants (N=6 autistic children).</td>
</tr>
</tbody>
</table>

Reproduced with permission
from the authors Peris, J.S and Fernandez, M.

**BendableSound [73]**

Reproduced with permission from Elsevier and Copyright Clearance Center

It uses a Kinect sensor, speakers, and an ultra-short throw projector located behind spandex fabric. 3D background displaying an animation of a nebula with translucent space-based elements, like stars and planets, appearing on top of the fabric. BendableSound endorses two open-ended activities and one goal-oriented task.

**Art Activity ([84], pp. 1148-1151)**

Reproduced under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

The prototype has been built for a desktop platform and extended to a web platform. It consists of 3 activities. 3D avatar simulates the movements made by the child, playing a musical instrument.

**The Magic Room [116]**

Reproduced with permission from Garzotto, F., and Gelsomini, M.

Two projectors, one video camera, one ambient sound system, one Kinect, various smart objects smart applications. Philips Luminous Carpet. Pressure sensors, multi-axis accelerometers, a gyroscope, a magnetometer, and various actuators for light, sound, and vibration.

**Interactive Multimodal Systems**

Interactive Multimodal Systems (IMS) use sophisticated pattern recognition techniques to support human-computer interaction using multiple interaction modes, including motion capture, voice, touch or wearable devices. Usually, IMS do not use traditional desktop and GUI interfaces.

<table>
<thead>
<tr>
<th>Table 2. Interactive Multimodal Systems.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Systems Overview</strong></td>
</tr>
<tr>
<td><strong>Soundbeam [85,86]</strong></td>
</tr>
<tr>
<td><strong>MidiGrid [117]</strong></td>
</tr>
<tr>
<td><strong>Multimodal Technol. Interact. 2021, 5, 11</strong></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td><strong>Reproduced with permission from Elsevier and Copyright Clearance Center</strong></td>
</tr>
<tr>
<td><strong>CymaSense</strong> [87]</td>
</tr>
<tr>
<td><strong>Reproduced with permission from McManus, M.</strong></td>
</tr>
<tr>
<td><strong>Reproduced with permission from Royal Society</strong> [88]</td>
</tr>
<tr>
<td><strong>Trocas</strong> [119]</td>
</tr>
<tr>
<td><strong>Reproduced with permission from Silva, M.L.</strong></td>
</tr>
<tr>
<td><strong>Reproduced with permission from Jorda, S., Villafuerte, L., Markova, M.</strong></td>
</tr>
<tr>
<td><strong>Reproduced with permission from Kiefer, C.</strong></td>
</tr>
<tr>
<td><strong>SensoryPaint</strong> [105]</td>
</tr>
<tr>
<td><strong>Reproduced with permission from Ringland, K.</strong></td>
</tr>
<tr>
<td><strong>It is a multimodal system that allows users to paint on a large display using physical objects, body-based interactions, and interactive audio. Rubber balls held by the user are detected by the system and act as paintbrushes of various sizes, textures, and colours, allowing for the painting of lines.</strong></td>
</tr>
</tbody>
</table>
Scoog [122]

Reproduced with permission from Rinta, T.

It is an innovative musical instrument that can be used by everyone. It is tactile and easy to hold, with no awkward hand positions, and a comfortable soft cube shape. The process of trial and error required to get a song or sequence of notes correct develops resilience, and the satisfaction of achievement when songs are successfully played.

OSMoSIS [123]

Reproduced with permission from Ragone, G.

The system required the development of software that processes the data from a Microsoft Kinect camera. OSMoSIS captures the body movements of one or more individuals and transforms them in sounds, in real time. The software is built with open-source JavaScript web technologies.

4.2.3. Music-Interaction with Robot (Human Robot Interaction—HRI)

The application of robots in autism research can be grouped into 4 typologies: (i) Cooperative human-robot dance, (ii) Imitation of human dance motions, (iii) Synchronization for music and (iv) Creation of robotic choreography [87].

**Table 3. Interactive systems with robot.**

<table>
<thead>
<tr>
<th>Systems Overview</th>
<th>Technological Features</th>
<th>User Interaction</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keepon [92]</td>
<td>It has a number of sensors to express attention, body movement, excitement and fear (by vibrating). Keepon, behaves (in terms of movement and sound) in synchrony with perceived rhythm. Keepon can perceive rhythms in a diversity of modalities and can synchronize recurring dance-like movement to these rhythms. The body has four degrees of freedom: nodding, turning, rocking and bobbing which produce two qualitative actions: Attentive and Emotive. Music begins to play and Keepon starts dancing as soon as the child holds and moves the toy, which is connected to an accelerometer. Keepon is monitored by an operator, and after two songs have been played for a period of 4 minutes, the child is free to play for indefinite time, with rhythmic synchrony to the movement of the toy. The outcomes of the study with N = 25 TD children and 30 pervasive developmental disorder (PDD) children (included autism) showed a successful elicitation of social behaviour from autistic children as well as enhancing their interaction.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Darwin-OP Robot [93]</td>
<td>It is a humanoid robot designed with open architecture, so it can be easily changed in its framework and program. Darwin robot plays the role of a facilitator. The authors tried to make the robot act as a humanoid through remote control, which has still limited DARwIn-OP was built to teach and play music, through imitation games, to autistic children. The aim of the study was to teach social or communication skills, through music, to autistic children. The study was conducted with N = 2 severe autistic children, showing the improvement of communication, social skills and motor coordination while playing the instrument.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
NAO Robot, aka ‘Nima’ [94]

Reproduced with permission from Taheri, A.

NAO is a humanoid robot, with 25 degrees of freedom which gives it a morphology that is very similar to that of a human which teaches how to play music. A configurable user-friendly GUI (Graphical User Interface) and some rhythm patterns have been developed to run different modes of the music scenarios, controlled either manually by the operators or automatically in real-time.

The system was mainly designed to teach rhythmic patterns to autistic children as well as notes, playing the xylophone. The developers attached mallets to the robot's hands. The choice of the instruments of drum and xylophone comes by the fact that they are both easy to play.

The results from the preliminary exploratory study with a small mixed group of high and low functioning N=4 autistic children. It showed positive outcomes in improving cognitive, social as well as motor skills (imitation).

Xylotism [81]

Reproduced with permission from Taheri, A.

It is associated with the NAO robot. Both were designed in a virtual environment called Xylotism to teach music to autistic children. The name derives from the combination of Xylophone and Autism. The application is interactive and provides feedback to the child by means of verbal and non-verbal reactions.

It is an interactive game to improve learning and teach music to autistic children. It supports the improvement of the skill of playing rhythms in autistic children. The Xylotism, as an app, is versatile and can be used as an auxiliary tool specially for using at home.

The study was conducted with N=20 autistic children, conducted. The results showed an improvement of the engagement, reinforce the imitation and hand-eye coordination.

ROMO (Romotive) [103]

Reproduced with permission from Javed, H. and Park, C.H.

The robots ROMO and Mini communicate through gestures and speech, expressing a variety of emotions. The robot can also display some physical pattern as representative of fatigue, or dizziness. For gestures, Robotis-Mini is designed to express same sets of emotional and social contexts as Romo, through bodily gestures and movements chosen based on the potential responses autistic children may exhibit as reaction to stimuli.

The system can provide some measures of the child’s emotional states detecting motion analysis and vocal reactions. “Five Senses,” is an interactive session realized with ROMO based on the knowledge that autistic individuals often display several sensory issues associated with the five senses and vestibular/proprrioception.

The study was conducted with N=13 TD and N=9 autistic children. The study reported improvement in social engagement, and enhancement of gaze and smile behaviors as well as imitation and self-initiated interaction.

Social Virtual Reality Robots (V2Rs) [124]

Reproduced with permission from Taheri, A.

The system is a combination of virtual reality, social robots, and music training. In the virtual room are present 1) Two virtual humanoid robots (Nima and Sina) and 2) Virtual Xylophone/Drum. The virtual xylophone has eight colored bars which have the same sound as a standard xylophone. There is also a virtual drum in the room which sounds like a real drum. The virtual musical instruments are placed in front of the V2Rs.

The robots interact, providing audio and social feedback to the child during different scenarios to improve their performance. The application uses the participant’s visual attention and tracks different objects. The robots are programmed to play/teach music notes/rhythms with a xylophone and a drum in the virtual environment and assess their behaviours.

The study was conducted with a group of N=14 autistic and 21 TD children. Results showed the potential use of the designed V2R music-based game as autism screening. It also provides analysis of imitation accuracy/agility during the sessions for behavioural assessment.

Rero Robot [125]

Reproduced with permission from Yusof, H. M.

It is created with several parts such as controller, head, cube servo, cube joint, u-joint, long u-joint, castor, slim wheel, foot plate, claw, infrared sensor, touch sensor, extension cables, connector and spacer. It is also mobile, speech enabled, controllable, programmable, and attractive.

Five modules of interaction have been built and adopted from special education teachers and therapists. Several authors developed their own prototypes or with the aim to enhance learning and communication. The modules have been organized to help children to be capable to emulate actions, follow instructions, name objects.

Investigations on its use with a group of N=9 autistic children, display good outcomes in terms of social and motor interaction and communication.
The devices have bidirectional communication with tablet. The app works only on iOS devices and can collect data on the users’ physical skills such as reaction time, sequencing skills, logical reasoning, and motor skills of speed of movement, force and tremor. The collected data is then anonymized. It comprises smart hardware devices, an application connected to the devices featuring a variety of activities designed by therapists and teachers.

Cozmo is an interactive robotic character that plays video game character animation in a toy rather than on the screen. The facilitator often modeled actions with the Cozmo units. The system combines art, music, and technology. Cozmo provides a data analytics platform that may be used for the assessment of user progress. Cozmo’s random activities like singing and giving players fist bumps caused positive affect as well.

A group of N=4 autistic children who took part to the research attended 5 sessions, all led by the same facilitator, who let the children lead by imitating their sounds and physical actions. It is useful to develop social skills (turn taking, sequencing, storytelling, vocalization) and facilitate fine and gross motor skills.

Waseda (WAS-5) [91]

WAS-5 is a humanoid robot which performs music with an ordinary saxophone and can be set in 2 modes: autonomous or it can be played in real-time in a Wizard-of-Oz mode. Music can be played in two different modes: a) Pre-loaded files (autonomous) or b) Remote input (real time activation). Gaze detection is based on OpenCV with its machine learning library to detect the gaze of the subject on WAS-5.

Cozmo’s random activities like singing and giving players fist bumps caused positive affect as well.

This research, designed for autism, has done a preliminary experiment with a mock program of the training program with TD people with positive effects. Further experiments with autistic participants will be done with hardware upgrade.

V2R Robot [126]

The study was conducted at the Social and Cognitive Robotics Lab at SUT using an HTC VIVE virtual reality headset in the presence of the child, one human teacher, and an operator in a Wizard of Oz style control.

V2R-assisted music-based intervention consists in a virtual reality classroom, designed for musical rehabilitation with two virtual humanoid robots, Nima and Sina, designed to teach/play virtual musical instruments (a xylophone and a drum).

This research, designed for autism, has done a preliminary experiment with a mock program of the training program with TD people with positive effects. Further experiments with autistic participants will be done with hardware upgrade.

Looking at all the 27 applications selected for this review, 12 of them have been integrated in Music-therapy contexts. The remaining 15 have the potential of being integrated in Music-therapy practices. We can characterize some common goals and practices used across the technology applications (number of applicable papers shown in brackets):

1) enhance motor or/and social skills (26)
2) gain the child’s attention (17)
3) promote imitation, interpersonal synchrony (15)
4) teach musical rhythm and patterns, ICT skills (12)
5) monitor/record the child’s responses (10)
6) give simultaneous feedback to body movements (18)
7) screen for autism (4)
8) behaviour management (4)

From the current review, we can summarise such technological features as follows (number of applicable papers shown in brackets):
- simple and intuitive tangible screen design (3)
- sensors (accelerometer, gyroscope, and magnetometer) (9)
- motion capture/Augmented reality (3)
- Kinect sensor, speakers, projector located behind or embedded in fabric (6)
- 3D avatar, which simulates the movements made by the child (5)
integration of the user’s movements with the virtual world (3)
- aesthetic resonance for the user’s full immersion in the augmented (3)
- interactive tools, toys, shapes (3)
- humanoid robots (4).

A couple of the systems considered in this review allow people with limited movement to engage in free improvisation thanks to specific sensors [85, 117]. Applied on the head or neck which generate sounds according to their movements. The simultaneous feedback generated by motion capture, for instance, makes children more aware of their bodies and facilitates their responses to interact with the researcher or music therapist [123]. Also, simultaneous musical feedback generated by body movements may trigger curiosity in new body movements’ exploration as well as offer a non-intimidating way of interaction. Virtual environments contribute to overcoming cognitive and perceptual differences. In addition, VEs provide autistic children with new opportunities to generalise experience to the real world [73,82,84,116].

Furthermore, in most of the examined cases, the activities are proposed to children like music lessons, so rooms, where studies take place, are not typically set up as they are in cases of therapeutic interventions [82,121]. The adoption of new tools may eventually lead to the reorganisation of the therapeutic space. For example, in the case of technological devices such as physical tools with embedded sensors, it is essential to consider the position of these devices in the room. In the case where no instruments are needed, it may be important to consider the reaction children may have in an empty therapeutic room. The light, their position and their intensity need to be well planned in order to avoid negative experiences for autistic individuals. Everything needs to be well planned and structured and being placed and presented in the least obtrusive way for children.

We report that generic benefits for social or motor skills (or both) emerge in all the studies’ findings. When we refer to social skills, we consider the development of joint attention, social reciprocity and body or non-verbal communication, which music and sounds generation trigger in most cases [60,94]. The review highlights the contribution of technology in attracting children’s attention and motivating their interaction either with facilitators, with their peers or with robots. Almost all studies included in this review targeted social skills, as verbal and non-verbal interaction, through gestures and body language. The development of gross motor skills like walking, running, balancing, and motor-sequence praxis, or the enhancement of fine motor skills such as synchronization of hands and fingers with eyes for example, as well as small movements of wrists or fingers, are not always addressed by the systems we examined. Technology expands the range of opportunities the music therapist may offer to the child to express themselves and interact with others. The collection of data seeks to present more accurate results in comparison with traditional Music-therapy settings. What is changing in Music-therapy, other than the musical tools used, are also the typology and source of data. In fact, the selected studies use sensors [82, 92,115-117] which feel pressure, distance, proximity, and direction which transform movement to music or sounds or offer identifiable visual and tactile characters that generates real-time 3D graphics [73,84,87]. As already specified, not all of the selected studies have been carried out in a Music-therapy setting. For instance, in one of them [88], the aim was the detection of anxiety. Using wearable sensors on a child’s fingertips and chest, authors tracked physiological input such as pulse breathing, temperature and sweat. These inputs were correlated to emotions and associated to an algorithm that interprets that data as sound, generating an electronic, musical output (not danceable). The theoretical background in support of most of the selected studies stem from multimodal embodied interaction, theories of embodied cognition and sensory integration, as well as synaesthesia and cross modality literature [23]. We included one of the studies [91] as a work in progress, although it did not present results with autistic individuals but with typically developing participants. However, the system was initially designed for autism and presents features which can be taken into consideration for future design.
Well-designed interfaces can remove the intimidation factor that comes with traditional musical instruments, such as guitar or saxophone, and their requirement for prior technical abilities. However, the choice of sounds needs to be planned and, ideally, tailored for each individual. Autistic people often report auditory processing issues which may trigger conditions like anxiety and confusion in social situations, so the therapist needs to carefully consider all the factors taking part in the therapeutic setting.

Previous studies [127,128] suggest that in therapeutic sessions with autistic individuals, human-robot interaction can result in better outcomes in comparison to the human-human interaction. Social robots used to teach fundamentals of music to autistic children in therapy sessions showed a promising effect on involving children to interact and learn [81,94,129,130]. Taheri and colleagues [94] report a positive correlation between the intervention with social robots, fine motor imitation, communication skills and even the autism severity of participants. Sounds generated by body movements or by interactive, virtual environments enhance the child’s proprio perception and self-confidence. The interaction with the robot, which mimics and moves in synchrony with the child [123], offers opportunities to improve motor skills, timing and body coordination. The settings are usually intended to be fun and attractive. For example, robots, therapists or facilitators might teach children how to play drums, or rhythmic patterns as well as notes in virtual reality, or users might be invited to move freely or guided pretending to play instruments to promote social interaction. The technologies generally aim to simplify the environment according to the level of input stimuli that is appropriate for the child and most of them try to tailor to each individual’s need. The applications seek to be attractive and stimulate the child’s creativity. Augmented [83,128], virtual environments [131] or robots [90,132] aim to encourage the children’ self-awareness, body schema and postures, communication, and imitation through a framework of activities where technology such as body tracking helps. Imitation, one of the main targets in the interventions, is mainly achieved through an avatar or humanoid robot which, as an intermediate, tries to mirror the child and elicit their active interest in responding to the gestures and activities proposed. Only one of the studies uses a human facilitator to achieve the same objective, i.e. the facilitation of children’s imitation and body schema [83,123].

In addition, we noted the need to create affordable, accessible, and useful systems for human-computer/robot interaction, resulting in easy to use, user-centred systems. The design of such systems required the interdisciplinary collaboration of several specialties such as computer science, cognitive science, music and occupational therapists, researchers and human-factors engineering. The collaboration of such disciplines made possible the development of systems useful primarily to expand, facilitate the autistic user’s needs and for the related figures such as therapists, scientists, teachers, and families.

5. Discussion

This review highlights some of the challenges encountered in research efforts towards the development of technologies for autism interventions: small numbers of study participants, nongeneralizable outcomes, technology considerations, privacy, legal and ethical issues, and assessment procedures which are not shared or not defined. This review, through analysis of the various studies examined, highlights how music and sounds can be deployed in various ways and be used in therapeutic context for developmental and learning conditions. Digital technology has contributed to changing the way music therapists deliver their therapeutic goals as well as the tools used as a channel for expression and communication. In real scenarios, it is rare for technology to be completely absent in Music-therapy contexts. Sometimes traditional musical instruments are still used along technology while in some other cases, sounds are generated by invisible devices and software [123]. Most of the studies selected were researchers whose scope was to expand the range of approaches to using music in therapeutic contexts to enhance the interaction and communication of autistic children. The positive outcomes identified therefore need further investigation if they are to become part of specific music therapeutic approaches.
The aim of this review was to search for and examine published studies focusing on motor and social skills in autistic children using technology enhanced music-based interventions. The results of this review revealed that there are three main areas in which technologies have been developed: (i) Musical activities with Virtual Environments; (ii) Musical activities with specific applications; (iii) Musical activities with Robots. We have also included papers where music therapy was not specifically mentioned, even though sound played a central role in improving children’s interaction and communication. Since the consistent number of studies which show how autistic children enjoy music and have heightened musical skills [47,58], socially embedded musical context that facilitate eye contact, turn taking, imitation and interpersonal synchrony, it is advisable to design and develop tools which enhance these opportunities as well as the wellbeing of children using those tools.

In some of the studies examined, authors propose new electronic musical instruments with an intuitive design, which let users to be creative and explore the sounds’ world through a tangible tabletop or screen [106,117,119]. Considering the studies that report on design approaches, authors generally seek to adopt a user-centred design methodology. However, there are no specific shareable methodologies or frameworks for future designs proposed. In most of the studies examined, the people conducting the research are researchers with an interest in music but with a different academic background and who do not follow specific methodologies of Music-therapy. However, these studies represent an important step in furthering the progress of methodologies where music induces beneficial outcomes in the treatment of specific conditions.

Future work should build on the efforts and findings coming from these studies to design system with evidence-based games and exercises which can be easy to transfer to teachers and parents and be performed at school and at home. Well-designed interfaces may allow people who experience limited ways of expressing themselves to lose the awareness of them using technology and free their emotional and creative expression. Qualities considered to be essential for effective design were system adaptability and user-centred, as well as user participation, iterative prototyping, and interdisciplinary development teams. The robots’ design resulting efficient to attract children’s attention have humanoid [91,93,94,124,125] resemblance or attractive colours [90,94,124,125] looking more like interactive toys [92,94] providing audio and social feedback to the child during different scenarios to improve their performance. Some systems use the participant’s visual attention and tracks different objects as well as providing analysis of imitation accuracy/agility [85,124] during the sessions for behavioural assessment of the children [90,124].

Also, consideration of the role of synchrony, which has been shown to be relevant in autism, has so far been limited in this area, so there are no relevant findings yet to suggest promising design features tailored for autistic users. Instead, we can confirm the value of technologies where sound and visual stimuli react to gestures and footsteps of children enhancing their sense of synchrony. Synchrony is not directly addressed in most of the cases; however, the simultaneous reaction of technology is nevertheless a valuable technique.

From this review we confirm a scarcity of robust, quantitative, longitudinal, studies in this field, and that the type of data collected generates concerns when evaluating suitability or efficacy of the technologies used to address users’ involvement.

6. Conclusions

This review of studies contributes evidence that the use of technology in Music-therapy interventions has potential to improve the wellbeing of individuals, although it is clear that further research is required. Empirical evidence shows that there have been many attempts to address the core challenges of the condition, including motor and social interaction of autistic individuals. Overall, we try to draw general guidelines which can

---

Multimodal Technol. Interact. 2021, 5, 11
combine the most pertinent aspects of each study examined, which can be considered for a future unique methodological framework.

Where the goal is to facilitate exploration and acquisition of social skills and motor skills by young autistic children, it is crucial to adapt the environment to the individual child, avoiding many distractions as well as over-specialization of the interaction to what may be interesting of some children and not all. We have seen as the association of movement and music often works but we need to pay attention to the sounds and their features adopted in each project. In designing technology for autism, we may consider an inclusive, interdisciplinary dialogue which can lead to the generalization of learning. Therefore, it is necessary to consider the quality of attractiveness, of simultaneous features which respond in real time to children’s input and be careful to calibrate the stimuli in order to avoid overstimulating children which can reverse instead the good benefit any intervention might have. Technologies are able to perform more functions and help both clients and therapists in stages of therapy and musical interaction. However, as a tradeoff, teachers or researchers need to learn how to use them correctly and adequately with autistic children.

Music-therapy may have the potential to improve the quality of lives of autistic children and their families, and the relationships between parents and their children. However, we still do not have sufficient studies to develop a comprehensive framework of Music-therapy, with the added value of technology, as a practice. Today, it may be more challenging to develop a unified framework as the technologies deployed to produce and process music and sounds become ever more varied and diverse. An important finding is that no specific unpleasant effects or destructive outcomes were observed by any of the studies included in this review, and all the interventions mentioned in this review was shown to be well accepted by almost all participants. To advance this research, studies employing larger samples and longitudinal designs are needed to further validate these outcomes and to explore whether the effects of these music-based interventions are long lasting.

Author Contributions: Conceptualization, G.R. and K.H.; methodology, G.R.; investigation, G.R.; writing—original draft preparation, G.R.; writing—review and editing, G.R. K.H. J.G.; visualization, G.R.; supervision, K.H. J.G. 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: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: We would like to express our gratitude to Emeline Brulé and James Jackson for their valuable feedback on the initial draft of the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References


