

Development of a System Combining a New Musical Instrument and Kinect: Application to Music Therapy for Children with Autism Spectrum Disorders

Tomoko Ichinose, Naomi Takehara, Kakuko Matsumoto, Tomomi Aoki, Toko Yoshizato, Ryuhei Okuno, Shinichi Watabe, Katsumi Sato, Tsutomu Masuko, Kenzo Akazawa
Tohoku University, Mukogawa Women's, Setsunan University, Graduate School of Educational Informatics, Social Welfare Organization Kibounoie

Abstract

This study describes a novel system that links an electronic instrument called Cyber Musical Instrument with Score (Cymis) and a game device called Kinect to provide music therapy for children with autism spectrum disorders (ASD). The system was developed to facilitate independent and active participation of children with ASD in music activities and teach them to integrate visual and audio sensory, motor, and physical awareness. The system combining Cymis and Kinect has been applied to both typically developing children and those with ASD, demonstrating that it can be used appropriately by either group. Preliminary studies indicate that the opportunity to “play” a familiar song by making desired movements can motivate children with ASD or similar cognitive symptoms to improve on-task behavior and collaborate effectively with a partner while the accompanying video images can be either motivating or distracting. The data obtained from these trials can be used for further empirical research and practical application of the system in music therapy for children with ASD.

1. Introduction

Music therapy is used with people of all ages who have a wide range of diseases and disabilities to exploit the beneficial physiological, psychological, and social functions of music. Music therapy is also used with children who have autism spectrum disorders (ASD) in clinical and educational settings.

In Japan, the implementation of the School Education Law in 2007 enhanced the provision of “special support education,” or education targeted at serving students with special needs[1]. These students may have difficulties in social interaction, interpersonal skills, cognitive ability, language, and other areas. The need to provide individual support

for these students has increased and its importance has gradually become more acknowledged in recent years.

Music therapy provides support tailored to each individual’s age, developmental stage, and the characteristics of his/her disability; it thus has the potential to address various special needs effectively. Music can also be linked with modern forms of information and communication technology (ICT), which can be used to enhance the customized support provided to a particular student depending on his or her diagnosis or degree of disability [2].

Typically, music therapy sessions involve activities such as performing with musical instruments, singing, moving to music, and performing ensembles to help each participant achieve specific goals related to his or her needs. The goals of music therapy are unique for each individual, and thus the activities and methods employed in music therapy sessions are linked to that individual’s goals.

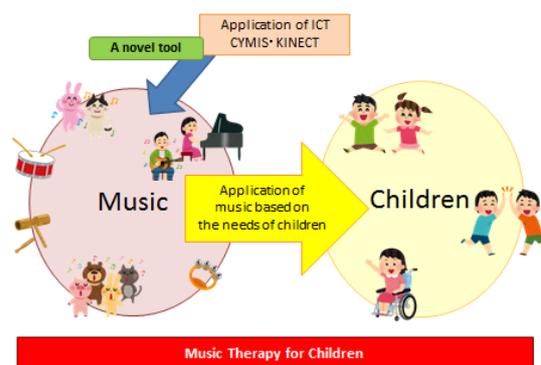


Figure 1. The rationale for combining Cymis and Kinect in music therapy for children

ASD is a spectrum disorder in the sense that it encompasses a wide range of symptoms, skills, and degrees of developmental delay [3]. Therefore, the goals of music therapy for this population also vary, but they commonly include improvisational

communication through music, which can facilitate self-expression and cognitive development, promote self-efficacy and confidence, and improve the appropriateness of sensory processing, motor coordination, and emotional release. The methods of utilizing music also vary and can include the use of previously composed music or improvisational music and sounds.

With regard to the use of ICT in music therapy, Whithead-Pleaux et al. [4], Magee et al. [5, 6], and Hahna et al. [7] reported on its application in medical rehabilitation and cognitive rehabilitation with a wide range of populations and in numerous clinical areas. Krout [8] described the use of music applications with individuals who have ASD; children with this diagnosis frequently enjoy performing on the drums, piano, or other instruments. A tool for music performance that utilizes ICT could thus have great potential as a complement to traditional music therapy (see Figure 1).

We have developed a new electronic musical instrument called Cyber Musical Instrument with Score (Cymis). Using Cymis, people who have not learned how to read a musical score can nevertheless “perform” music by making suitable movements, which are perceived by an interface and converted into sound. Even persons with neurological, motor, and cognitive impairments can enjoy playing music in this way; moreover, the system fosters learning as participants can improve their playing with practice, producing a familiar song in continuous, recognizable fashion rather than a fragmentary collection of notes.

In Japan, Cymis has been utilized by 16 organizations, including hospitals, special support schools, welfare facilities for disabled persons, and research organizations. The positive effects of using Cymis for rehabilitation purposes have been described [9]. The music therapy research laboratory at Mukogawa Women’s University, which provides music therapy services for children with special needs, has also used Cymis.

What can we expect when we use this electronic musical instrument with children who have ASD? One reasonable answer is that Cymis should facilitate the integration of sensory and motor systems. Schneider et al. [10] reported that stroke patients improved their motor functions after playing electronic drums as a means of training in motor skills. This result suggests that playing percussion instruments with audio feedback positively impacts the brain’s neurological system. Furthermore, we believe that playing electronic instruments could help children with ASD to develop their capacity for independent and active participation. Finding creative, motivating ways is necessary to facilitate the integration of visual and audio sensory, motor, and physical awareness within this population. According to Wilson [11], embodied cognitive

theory claims that cognitive processes are rooted in the body’s interaction with the environment and that sensory and motor functions are crucial for successful interaction with the world.

Individuals with ASD tend to experience severe difficulty with executive functions, including motor control. A study reported that the development of motor control was delayed among individuals with ASD [12]. Piek et al. [13] found a strong correlation between problems in motor coordination and attention.

As for the relation between music and movement, Thaut [14] introduced a series of research-based techniques referred to as neurologic music therapy. According to this research, rhythmic drivers optimize the motor control of arm movement, and rhythmic auditory stimulation is effective in enabling proper walking; thus, music can play a valuable role in sensorimotor rehabilitation.

The Kinect game device, developed by Microsoft, is considered to be effective in promoting visual and motor integration. Research on educational and therapeutic applications of Kinect has been growing in recent years. For example, Garsotto et al. [15, 16] conducted an empirical study demonstrating the effective application of Kinect in addressing the attention skills of ASD children with low to moderate cognitive deficit, low to medium sensorimotor dysfunction, and difficulties in motor autonomy. Bartoli and Lassi [17] also used Kinect to address attention problems in children and adults with ASD. Boutiska [18] described the use of Kinect in both traditional and special classrooms with students who have mobility problems and specifically proposed the game’s applicability to children with autism. Casa et al. [19] used Kinect as an augmented-reality system for children with ASD, aiming at facilitating self-awareness, body schema and postures, communication, and imitation. A music engine was used in the exercise performed in this study based on the fact that music therapy proved to be effective.

From the perspective of motor control, therefore, there are ample indications that Kinect has potential educational benefit for children with ASD.

When Cymis, as an instrument that offers audio feedback, is combined with Kinect’s contributions to visual feedback and motor and physical awareness, it can be applied as a method of music therapy. In combination, Cymis and Kinect offer not only audio feedback in response to the individual’s own movements but also visual feedback as the individual sees his or her own motions on a monitor. This immediate and accurate visual feedback using ICT was previously impossible due to technological reasons and thus did not exist in conventional music therapy.

This study describes the computer engineering perspective underlying the development of the

Cymis and Kinect electronic musical instrument system as well as the development and testing of a novel music therapy method utilizing sensorimotor integration from a cognitive scientific perspective. Our initial research involved developing the system, ensuring that it could be appropriately used with children (including both typically developing children and those with ASD), and obtaining data for further empirical research and practical application for children with ASD.

2. Outline of Cymis

Although not all the functions of Cymis are used in this study, a brief description of this instrument is provided to indicate its capacities and the possibilities for further study and development.

2.1. The concept

Generally, to play a musical instrument, one must be able to read a musical score and employ proper techniques to manipulate the instrument to produce sounds that match the pitches, rhythms, and various other information indicated in the score. It is not an easy or instantaneous process. Considerable knowledge is required to understand a musical score, and acquiring the physical techniques necessary to perform a musical instrument involves extensive practice.

In contrast to these imposing obstacles, Cymis eliminates the technical barriers associated with playing a musical instrument. Its development was designed to achieve three conceptual goals, permitting anyone to (1) perform music easily, even a relatively advanced musical piece, (2) enjoy performing music, and (3) improve by practicing, thus experiencing the feeling of achievement in performance.

2.2. The structure of Cymis

In Cymis, preprogrammed music score information necessary for a musical performance is stored in a personal computer (PC). The instrument 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 (see Figures 2 and 3). Using these interfaces, people of any age and even individuals with neurological and motor impairments, such as those with cerebral palsy, can enjoy performing music. For instance, an individual who suffers from impairments that affect the feet and hands can perform through his or her breathing.

The touch panel is the main interface. The musical score of a selected piece is displayed on the touch panel's monitor. The performer plays the selected music by pointing to notes on the monitor. When the performer points to a note head displayed on the touch panel, the pitch associated with that note is sounded.

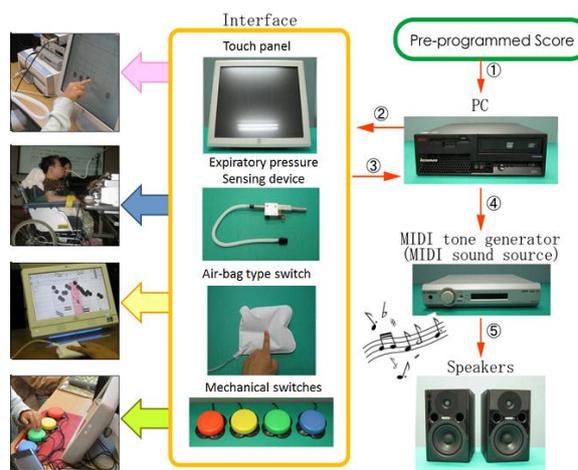


Figure 2. Structure of the Cymis system

The signal of the identified x-y coordinate is sent to the PC, where it is processed, and the sound production signals are then sent to the MIDI sound source, which emits the sound from the speakers. In this manner, a piece of music can be played. The process of understanding and reading a music score and the generation of sounds, which is necessary for a conventional musical instrumental performance, is replaced by Cymis.

In the currently developed Cymis and Kinect system, Kinect functions like a switch. Each time the switch is turned on, one beat of the piece is played. The score data, including not only the melody but also the accompaniment, and the tempo vary depending on the previous tempo performed. This means that the piece of music can be performed at an even tempo by repeating an operation smoothly. (Importantly, Kinect also permits the participant to “play” the music by making prescribed notions rather than by pointing to notes, as explained below.)

Alternatively, the piece of music can be played not only one beat at a time but with several measures of music played each time the switch is turned on. The number of measures played as a result of a single action can be adjusted between two and eight. With regard to Cymis, each note is played by turning the switch on. The software program for processing signals from Kinect using this method has not yet been completely developed.

2.3. The importance of improvement when playing a musical instrument

The experience of “improvement” is an essential part of continuous enjoyment when one plays a musical instrument. Even an individual with severe physical impairment can enjoy playing Cymis by starting with a one-switch interface at first and then advancing to two switches or a touch panel, thus enjoying a more complicated piece of music and a more challenging performance. This process generates a positive cycle such that the more

enjoyment the performer feels due to his or her improvement, the more he or she becomes motivated to keep practicing to improve further.

In this study utilizing Kinect, information on people's movement is used as a signal to control Cymis, so that the movements themselves become pivotal factors in the performance. Consequently, it is necessary to examine the appropriate extent of movement needed for an individual to play a piece of music and change the level of complexity and difficulty required as the individual's skill improves.

3. The Cymis and Kinect system

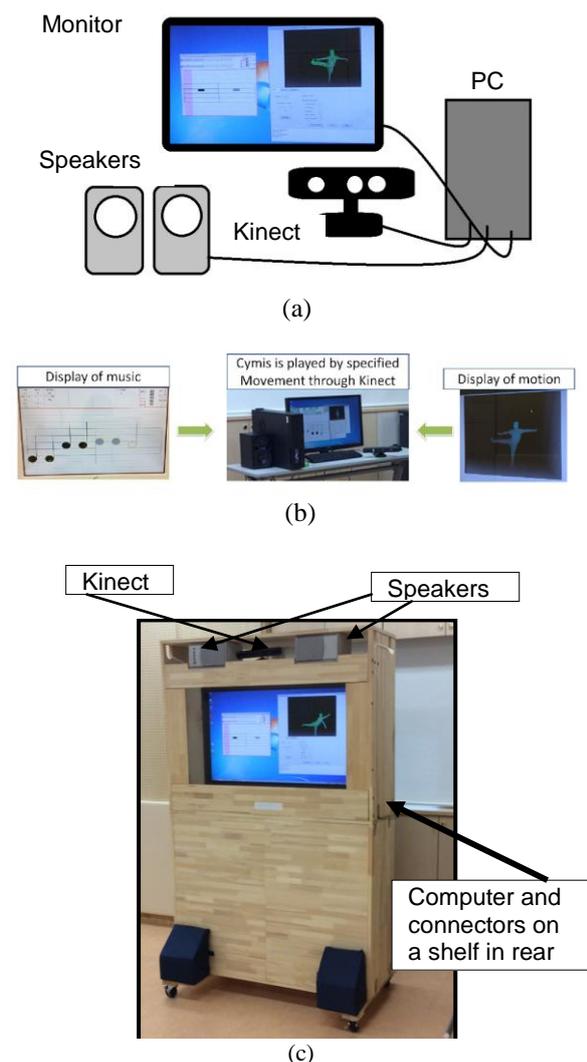


Figure 3: The Cymis and Kinect System

3.1 Kinect's features

Kinect is a simplified motion capture device created by Microsoft for Xbox 360. It comprises an RGB camera, a depth sensor, and an array of four microphones. It recognizes the skeletons of objects without any physical restriction. The device used in this study can recognize the skeletons of not more

than two persons and capture 20 points of location of the joints (using three-dimensional coordinates: X, Y, Z) for one person. The horizontal angle of view is 53 degrees, the vertical angle of view is 43 degrees, and the possible distance for acceptable measurement is between 0.8 and 4 meters.

3.2. The Cymis and Kinect device system

The Cymis and Kinect system is shown in Figure 3(a). It comprises Kinect, a monitor, speakers, and a PC. The coordinate data obtained through Kinect are processed, and the specified motions are recognized by the PC. After this, the recognized motions are sent to Cymis as signals, and the image data obtained through Kinect are processed by the PC and displayed on the monitor. Figure 3(b) shows an example of a performance created by dancing. Figure 3(c) shows how individuals with ASD can use the device safely. The structure of the device prevents individuals with ASD from touching the monitor, Kinect, or the speakers directly.

3.3. Development of a program for motion recognition

We applied the free software called FFAST to the program when it was recognizing one person as a target. For two persons as targets, we developed an original program called "High five." In this program, when the two participants' hands touch each other (more precisely, when the distance between the hands becomes less than approximately 15 cm), the signal of the character "Z" is the output. It is assumed that this program will be played by a child and an adult (perhaps the child's mother). There are four possible combinations of hands (since each player has two hands); for all four combinations, the program calculates the direct distance and judges whether it is less than 15 cm. Because individuals' posture can vary widely, the coordinate values of the hand could not always be measured accurately, and measurement errors of the order of several centimeters often occurred. We detected and examined these errors using a 3D position measurement device called OPTOTRACK. Due to these shortcomings of Kinect, we used a program that makes judgments whether the hands were "touching" based on the coordinate values of a hand and a wrist. Consequently, more reliable judgments have been obtained.

In the High-five program, an image video appears on the upper left side of the monitor, and the motion sensed by Kinect is shown on the lower left side of the monitor. The Cymis music score appears on the right side of the monitor (see Figure 4).

4. Application to Music Therapy

After we confirmed that a typical adult could play music using the Cymis and Kinect system in the laboratory level, we tested the device with typically developing children and those with ASD to see how it functioned and assessed its potential application for clinical or educational purposes. This study was approved by the ethical committee of Mukogawa Women's University, and informed consent regarding participation in the research project was obtained from the children's parents. All sessions were recorded by several video cameras for subsequent analysis. The sessions were conducted with the involvement of a music therapist, creating a relatively natural educational setting; verbal prompts and praise were used to guide and reinforce behavior.

4.1. Use of Cymis and Kinect (with FFAST) by typically developing children

We tested the device with typically developing children of different ages and heights. They performed "Twinkle, Twinkle, Little Star." A three-year-old boy, 95 cm in height, moved his right arm up to his head and down to his thigh. The music was played on Cymis in response to his motion (Figure 4a). Also, a seven-year-old girl, 125 cm in height, performed music on the Cymis system by walking in place (Figure 4b). The children naturally understood the correlation between their motion and the musical performance as cause and effect.

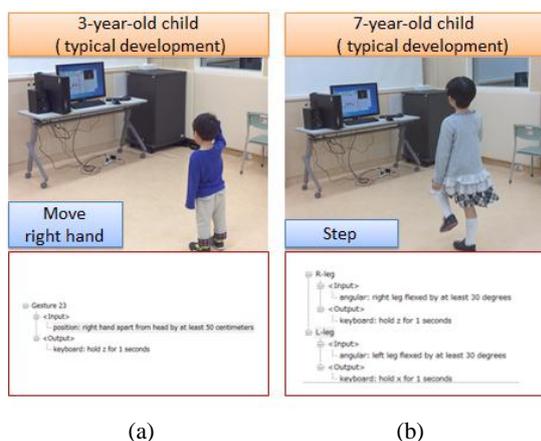


Figure 4. Typically developing children playing Cymis and Kinect (FFAST).

4.2 Use of Cymis and Kinect (High five) by typically developing children

Cymis and Kinect (High five) was used with adult-child pairs. The four typically developing children who participated included a four-year-old

boy, a six-year-old boy, a seven-year-old boy, and an eight-year-old girl. Music with which children are generally familiar, such as "Rudolph the Red-Nosed Reindeer" or a popular Japanese anime song, was selected.

Even the seven- and eight-year-old children needed some amount of practice with their adult partner to attain a performance level at which the resulting piece sounded like music and could be recognized as a familiar piece being performed.

Generally, the children recognized which skeleton reflected which player and were then able to retain that knowledge, since the two skeletons had different colors. They watched the display while giving a high five, but they did not seem to exhibit a special or intense interest in the video display of their own image.

The four-year-old boy thoroughly enjoyed giving a high five to an adult (his father) and commented, "Clapping was fun." He did not seem to recognize that he had played "music." The typically developing children exhibited enjoyment of the sounds created and the skeleton movement reflected by their own motion, continued playing while watching the monitor, and understood when they had played the music all the way to the finish. The children improved as they repeated playing with their adult partners.

4.3. Use of Cymis and Kinect (FFAST and High five) by children with ASD

The system was also used in sessions at the music therapy research laboratory of the Mukogawa Women's University School of Music. On average, 19 children, most of whom are enrolled in special support schools or classes, receive music therapy services there twice a month. Music therapists certified by the Japanese Music Therapy Association (JMTA) work with these children. As the university's program is JMTA-certified for educating prospective music therapists, the laboratory simultaneously functions as a clinical, educational, and research setting.

Two children whom the music therapists expected would obtain benefits from playing Cymis and Kinect were selected, and the system was used with each child. Child A, aged eight, was already diagnosed as having ASD; child B, aged six, was not yet formally diagnosed but had exhibited special cognitive-related needs common among children with ASD and was thus considered to be appropriate for this study.

High five was used with both children, and FFAST was used with one of them. The sessions lasted approximately 15 minutes.

The program settings were as follows. When the High-five program was used, one touch generated the sound of one note, whereas with the FFAST

program, if a child moved his arms and legs or his whole body freely, two measures of the piece of music were played. If he adopted a posed position without moving, the music stopped.

Only High five was used with child A (see Figure 6). The song “Twinkle, Twinkle, Little Star,” which was familiar to him, was played. The system was set up in such a way that the music would play if the child and a therapist gave high fives to each other.

Child A showed a strong interest in visual feedback from the image video and skeleton and stared at the monitor while standing close to it. At the beginning of the session, two adults (therapists) modeled the process for him, and after he watched them, one of the therapists touched his hand. The child gradually learned to initiate touching the therapist’s hand, and thus they gave each other high fives by touching each other’s hand. Through practice, they successfully performed “Twinkle, Twinkle, Little Star” as a whole piece of music. The child understood that he had played the music of “Twinkle, Twinkle, Little Star” based on comments by his mother, who observed the performance.

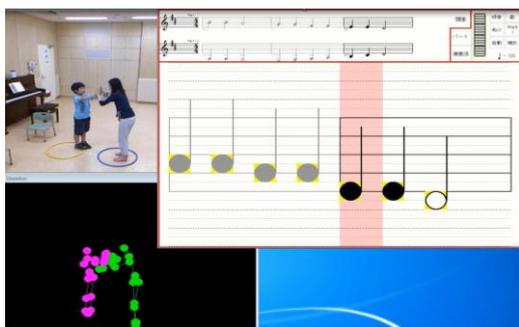


Figure 5. A child with ASD playing Cymis and Kinect (High-five)

Child B played both the FAAST and High-five programs. In the first session, he played Cymis and Kinect (High five) with his mother and father by clapping his hands together with each of them. The music selected by a therapist was a simple song with which the boy was familiar.

The child showed more interest in the images than in the music and sounds. He stared at his own image video and skeleton and enjoyed making various poses and motions in front of the system. It seemed difficult for him to make the specific motions required by High five to play in such a way that the sounds could be considered to be a coherent piece of music. In fact, even his parents did not recognize the song, as the mother indicated after the session because the notes were generated so fragmentarily.

In the second session, FAAST was used with the child alone, without his parents participating. The system was set such that if he made a gross motion with his arms and legs or moved his body freely

forward or backward, two measures from “Twinkle, Twinkle, Little Star” were played.

The child remembered the first session in which he had seen his image video on the monitor and to see it again by saying, “Show me the mirror.” This request suggests that he was strongly fascinated by the visual image. He moved around the system, back and forth, right to left, lying down on the floor, and making specific poses. Gradually, he noticed that if he made motions freely in a specific area in front of the system, sounds would be generated.

4.4. Further Application of Cymis and Kinect (High five)

This preliminary study showed that both children were strongly fascinated by the pictures of skeletons in the High-five program. We also found that when the children moved fast, the system could not fully detect the motions and errors could occur, thus reducing the consistency and immediacy of the feedback provided. Therefore, we further explored possible music therapy applications.

The current implementation of High five is limited to either walking in place, clapping, or touching hands. Therefore, in search of additional applications that would be aligned with children’s needs and interests, the “Wizard of Oz method” was applied. This term was coined by Kelly in 1984 to describe a testing or interactive design methodology in which an experimenter, called the “wizard” in a laboratory setting, manipulates the interface [20]. For example, a person may give input to an interface and then recognizes the reactions as coming from the interface, whereas in actuality, the experimenter is using a keyboard to control the interface. This method offered us additional ways to develop and test the user interface.

In this study, a music therapist functioned as the wizard, manipulating the keyboard and giving commands to Cymis from an adjacent room while watching through a one-way mirror (Figure 6). In this procedure, the wizard pressed a key to cause Cymis to make sounds when the child performed a specific motion.

We wanted to encourage the children to maintain a pose without moving to facilitate motor control and recognize that if they held a pose, music would be played. In addition, we set the goal of having the child perform and recognize a complete piece of music.

Based on these goals, we presented three new tasks to children A and B: (1) striking a victory pose with the arms, (2) stretching the arms straight to each side and then moving them straight above the head, and (3) in High five, touching the therapist’s hands above the head, then moving apart, moving around, and touching each other’s hands in various spots (near the floor, up in the air, etc.). For task 3, the

time (in seconds) required to complete the task was evaluated by video review.

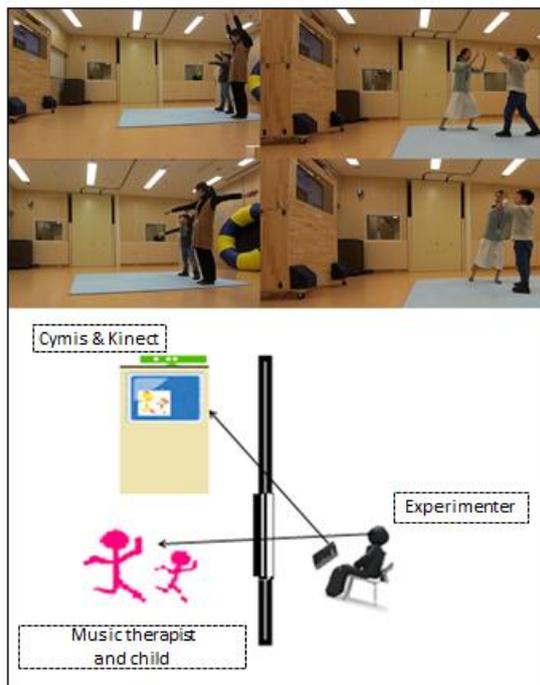


Figure 6. Playing Cymis and Kinect using the “Wizard of Oz” method

Child A participated in two sessions using the Wizard of Oz method. In task 1, he understood the rules easily and completed the task right away. In task 2, the child was to perform “Twinkle, Twinkle, Little Star” by moving his arms to his right and left, posing for two measures of the song, and then raising his arms straight up and posing there for two measures. If he repeated this sequence four times, the whole song, consisting of 14 measures including an introduction, was played. The boy performed this task successfully on his own after the therapist modeled it.

The child exhibited enjoying the achievement of performing the song and recognized that when he posed, the music played. By doing this repeatedly and hearing the music play, he could grasp the cause-effect relationship. Also, he learned to pose, stop, and then listen, which is important for both motor control and attention.

The third task required the child to perform the same song with a modified, moving set of high fives. Child A touched both of the therapist’s hands and posed there for two measures. Once they separated, the music stopped, and it then began again when they touched hands in a different location.

At the second session, Child A repeated task 3 four times successfully without distraction, with the result that the time spent on the task each time

remained relatively stable (Table 1). At the beginning and the end of the song, he watched the monitor to see his own image and jumped up and down, expressing a feeling of enjoyment. He did not gaze at the monitor, however, while engaged in the moving high five. The child did not need much prompting or encouragement from the therapist but engaged willingly and independently in the performance.

Table 1. Child A’s time to perform “Twinkle, Twinkle, Little Star” by moving high five

| | 1st | 2nd | 3rd | 4th |
|-------------------------------|-----------|-----------|-----------|-----------|
| Performing time (sec.) | 38 | 28 | 36 | 25 |

Child B participated in four sessions using this method. The songs “Twinkle, Twinkle, Little Star,” “Froggy Song,” and “Rudolph the Red-Nosed Reindeer” were used.

In task 1, “Twinkle, Twinkle” was used. The child was encouraged to strike a victory pose, which the therapist found to be one of his favorite poses. If he did so, two measures of music were performed.

At the second session, task 3, the modified, moving high five was introduced. Child B seemed to gradually become bored and lose interest. This may have been because of the repetition of the same song or because the song was a bit too long for his attention span. Therefore, at the third session the music was changed to “Froggy Song,” which was a familiar, shorter song that he could complete easily. The boy successfully performed “Froggy Song.”

At the fourth session, “Rudolph the Red-Nosed Reindeer” with modified, moving high five was used. At the beginning, Child B was distracted during the process of performing and needed the therapist’s prompts and encouragements. He then engaged in moving high five while singing along with the therapist. He repeated this four times. The first time, he could not stay on task; the second time, the music was heard in fragments; finally, the fourth time, the music was performed very smoothly (Table 2). At the end of the session, he correctly named the song title when asked by the therapist.

Child B, almost every time he made a high five with the therapist, looked at the monitor to see the image or skeletons. He also commented quite often about the colors of the skeleton, such as by saying, “I am red and you are green.”

Table 2. Child B’s time to perform “Rudolph the Red-Nosed Reindeer” by moving high five

| | 1st | 2nd | 3rd | 4th |
|-------------------------------|-----------|----------|-----------|-----------|
| Performing time (sec.) | 20 | 3 | 14 | 39 |
| | | | | 19 |
| | | | | 4 |
| | | | | 50 |

4.5. Discussion of applications of Cymis and Kinect

We discovered that both typically developing children and those with ASD were generally interested in the visual feedback from the image video and skeleton. They recognized the music and showed awareness of the beginning of and any interruption in the music. It was important for the music to be familiar to the children using the system. We confirmed that the system could be safely used with children.

The two children with ASD were particularly interested in the image video and skeleton. This might be because individuals with ASD tend to be visually oriented. If this strong interest in visual feedback is exploited appropriately in teaching children with ASD, Cymis and Kinect might contribute to improving their body image or body recognition.

Child A understood the task when he was face to face with the therapist and they gave each other high fives; thus, he accomplished the task of performing a piece of music by collaborating with the therapist. His completion of this task might have been easier because he had recently been gaining interest in imitating other people's motions through motion songs and in playing musical instruments such as bells.

Child A was engaged in all three tasks in the second study (i.e., striking a victory pose, moving the arms to each side and above the head, and modified moving high five). At first, the therapist provided prompting and modeling, but quickly the child learned to perform more autonomously. This child, as long as his safety is ensured, could be taught to play the system on his own. From the perspective of musical performance and improvement, it would be beneficial for him to try to expand his repertoire to longer, more complicated songs (but still ones with which he is familiar). For child A, both visual and audio feedback seemed effective and he recognized the construction of the music and exhibited a sense of completion when the song ended.

Child B was weak in coordinated movements such as activity requiring fine motor skills, jumping, or throwing a ball. Cymis and Kinect could be used to facilitate his awareness of body parts and to improve attention and control, thus helping him to learn various movements. Child B's visual attention was quite strong, and he commented on the colors of the skeletons and watched them constantly while performing on the system. He needed frequent encouragement from the therapist and enjoyed interacting with the therapist.

Child B improved his performance in the course of his interaction with the system and the therapist. Especially in his fourth session of the second study,

it became clear that the system could effectively facilitate his concentration, on-task behavior, and enjoyment.

With regard to how to introduce the system to children, it was most appropriate to facilitate their recognition of how the music should sound through slow rather than fast movement. This is because Kinect does not recognize motions if they are made too swiftly, with the result that Cymis cannot play the music appropriately. For the children with ASD who participated in this study, it was considered beneficial to establish the programs so that they could play Cymis and Kinect through making collaborative motions (such as motion imitation or motions in turn with another person), thus facilitating mutual communication skills.

With regard to the further application of the Wizard of Oz method, in which an experimenter manipulated Cymis instead of letting the input from Kinect function directly as the switch, several potential therapeutic and educational benefits were observed. The High-five program is limited to three motions, but the Wizard of Oz method permits use of a wider variety of motions depending on the child's situation, ability level, needs, and interests, while the attraction of the video images and skeletons retains interest.

Kinect can program the motions, but it is not feasible to change the motion program quickly when warranted by the process of interactions in an actual therapeutic or educational setting. The Wizard of Oz method permits more flexibility in changing or modifying the tasks on the spot, in response to the child's progress or interest. In this sense, the Wizard of Oz approach may be useful in FAAST applications as well as in High five.

In this empirical study, intervention by therapists was necessary. However, Cymis and Kinect can be used flexibly, either for interaction between an objective and the system or for three-way interaction involving an objective, the system, and the therapist, depending on the needs and goals of a particular music therapy session.

Cymis and Kinect as presented here enabled the performers to obtain both visual and audio feedback on their own motions. Visual feedback such as the image video and skeleton can provide an effective incentive or a form of engagement while the music and sounds offer effective audio feedback to facilitate awareness and cognition and provide cues for controlling motions. Establishing visual, audio, and motor tasks based on an appropriate assessment of the educational or therapeutic needs and characteristics of each child are necessary. With suitable individualization, the system could be applied effectively in future empirical studies or in actual therapeutic practice.

5. Conclusion

There are some limitations in our study, as our use of Cymis and Kinect remains in its early stages. This study primarily reports on the system's development, followed by preliminary, descriptive case studies involving two children who have ASD or similar symptoms. This study is significant, however, in several regards.

First, we have established the effectiveness of a novel system that permits people to perform musical pieces on Cymis as their motions are recognized by Kinect. As a preliminary study, we applied this system to four typically developing children, who generally enjoyed playing and receiving feedback from Cymis through both audio feedback and visual images. This feedback helped the children to understand cause and effect, and they improved in playing the music as they practiced.

Cymis and Kinect was also applied with two children who have ASD or similar cognitive deficits. It was found that children with ASD were more interested in visual feedback and showed awareness of the beginning of or any interruption in the music. They were thus also able to recognize cause and effect as well as to use the system safely.

As a result of these preliminary studies, we conclude that it is possible to gain both clinical and educational benefits using Cymis and Kinect with children who have ASD.

6. Future Perspectives

In this exploratory case study, we asked the participating children to complete several levels and types of tasks. We plan to include a wider range of tasks in the future.

In this study, one therapist and one child participated in each session. However, children A and B have been receiving music therapy together in their regular, traditional sessions separate from this study. We intend to attempt to use Cymis and Kinect with the two children as a pair, doing high fives and then moving high fives rather than pairing each child with a therapist. This experiment might enable us to further explore the system's potential not only to produce positive cognitive effects but also to facilitate constructive interaction between children.

As a next step, we hope to establish a series of motion tasks, from easy to more complex, that require executive functioning and the integration of visual, audio, and motor coordination. That is, we wish to identify visual, audio, and motor tasks based on an appropriate assessment of the individual child's clinical and educational needs, characteristics, learning level, and interests, thus making it possible to effectively apply the system in children with ASD in various settings, such as music

therapy sessions, special support schools, or other educational environments.

From the perspective of further development of the program, it would also be desirable to utilize a greater variety of sounds, such as those made by animals or a mother's voice. Another possible innovation would be to substitute other characters that are popular with children for the images of skeletons using computer graphics. In addition, a wider monitor or screen and speakers that enable the child to feel more surrounded by the sound could enhance children's engagement and enjoyment.

7. Acknowledgments

This study was supported by JSPS KAKENHI (Grants-in-Aid for Scientific Research (S)) Grant Number 26381347. T. Kinukawa of Setsunan University wrote the computer program for Cymis and Kinect High five.

8. References

- [1] Ministry of Education, Culture, Sports, Science and Technology. http://www.mext.go.jp/b_menu/hakusho/nc/06072108.htm
- [2] T. Ichinose, K. Matsumoto, N. Takehara, and S. Watabe, "ONGAKU RYOHO NI OKERU TEKUNOROJII NO KATSUYO: 2000 NEN IKO NO BUNKEN REBYUU WO CHUUSIN NI" ("Application of Technology in Music Therapy: A Literature Review Published after 2000"), *Japanese Journal of Music Education Practice*, 11 (2), 2013, pp. 60-65.
- [3] National Institute of Mental Health (2015). Autism Spectrum Disorder. Retrieved from <http://www.nimh.nih.gov/>
- [4] A. M. Whitehead-Pleaux, S. L. Clark, L.E. Spall, "Indication and Counterindication for Electronic Music Technologies in a Pediatric Medical Setting", *Music and medicine*, 3, 2011, pp. 154-162.
- [5] W. L. Magee, M. Bertolami, L. Kubicek, M. LaJoie, L. Martino, A. Sankowski, J. Townsen, A. M. Whitehead-Pleaux, and J. B. Zigo, "Using Music Technology in Music Therapy with Populations across the Life Span in Medical and Educational Program", *Music and medicine*, 3(3), 2011, pp. 146-153.
- [6] W. L. Magee, "Electronic Technologies in Clinical Music Therapy: A Survey of Practice and Attitudes", *Technology and disability*, 18, 2006, pp. 139-146.
- [7] N.D. Hahna, S. Hadley, V.H. Miller, and M. Bounaventura. "Music Technology Usage in Music Therapy: A Survey of Practice." *The Arts in Psychotherapy*, 39(5), 1999, pp. 456-464.
- [8] R.E. Krout, "Engaging iPad Applications with Young People with Autism Spectrum Disorders" Magee, W.L (Ed.) *Music Technology in Therapeutic and Health Settings*. Jessica Kingsley Publisher, London and Philadelphia, pp. 181-197.

[9] K. Akazawa, T. Kawai, R. Okuno, T. Masuko, and N. Nishida, Novel Electronic Musical Instrument for Persons with Cerebral Palsy to Play and Enjoy Together, In Proceedings of 9th International Conference on Disability, Virtual Reality and Associated Technologies, 2012, pp. 1-4.

[10] S. Schneider, P.W. Schönle, E. Altenmüller, and T.F. Münte, "Using Musical Instruments to Improve Motor Skill Recovery Following a Stroke", *Journal of Neurology* 54, 2007, pp. 339-346.

[11] M.Wilson. "Six Views of Embodied Cognition", *Psychonomic Bulletin and Review*. 9(4), 2002, pp. 625-636.

[12] T.Tsutomu, H.Shinoda, and H.Ozaki. "A Review and Case Study of Executive Function and Motor Control in Autistic Spectrum Disorder", *Rissho University Sinrigaku Kenkyu Nenpo*, 2, 2011, pp. 23-32.

[13] J.P.Piek, M.J.Dyck, A.Nieman, M.Anderson, D.Hay, L.M.Smith, M.McCoy, and J.Hallmayer. "The Relationship between Motor Coordination, Executive Functioning and Attention in School Aged Children", *Archives of Clinical Neuropsychology*, 19, pp. 1063-1076.

[14] Thaut, M.H. "Rhythm, Music, and the Brain: Scientific Foundation and Clinical Application", Routledge, New York, 2005.

[15] F. Grazzot, M.Valoriani, and L. Bartoli, "Touchless Motion-Based Interaction for Therapy of Autistic Children", M. Ma et al.(eds.) *Virtual, Augmented Reality and Serious Games for Healthcare 1*, Intelligent Systems Reference Library 68, DOI: 10.1008/978-3-642-54816-1_23

[16] L.Bartoli, C.Corradi, F.Garzotto, and M.Valoriani, "Exploring Motion-based Touchless Games for Autistic Children 's Learning", *IDC'13 Proceedings of the 12th International Conference on Interaction Design and Children*, ACM, NewYork, 2013, pp. 102-111.

[17] L.Bartoli and S.Lassi. "Experimental Study of Results Obtained from the Interaction with Software Motion-Based Touchless Created for Habilitation-Rehabilitation in Users with Diagnosis of Autism Spectrum Disorders", *Procedia Manufacturing* 3, 2015, pp.5176-5183.

[18] E. Boutsika "Kinect in Education: A Proposal for Children with Autism", *Procedia Computer Science* 27 2013.pp.123-129.

[19] X.Casa, G. Herrera, I. Coma, and M. Fernández, "A Kinect-Based Augmented Reality System for Individuals with Autism Spectrum Disorders", In Proceedings of the International Conference on Computer Graphics Theory and Application (GRAPP-2012), pp.440-446.

[20] J.F.Kelly. "An Empirical Methodology for Writing User-Friendly Natural Language Computer Application" *CHI'83 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1983, pp.193-196