

A motor behavioral evaluation method for children with developmental disorders during music therapy sessions: A pilot study.

Zu Soh¹, Ryo Migita², Kayoko Takahashi³, Koji Shimatani⁴, Hideaki Hayashi¹, Yuichi Kurita¹, Toshio Tsuji¹

¹Institution of Engineering, Hiroshima University, Japan.

²Graduate school of Engineering, Hiroshima University, Japan

³Orange Studio, Japan

⁴Department of Physical Therapy, Prefectural University of Hiroshima, Japan.

Abstract

Background: Although music therapy has long been recognized as an effective treatment for children with developmental disorders, evaluation of their motor behavior during therapy sessions has always depended on subjective and qualitative methods. Additionally, music therapists may face difficulties in conveying opinions based on observations conducted in therapy sessions to parents due to a wide disparity in the characteristics of children's behavior in different circumstances.

Objective: This pilot study was conducted to examine a computer-aided evaluation method for music therapy involving the use of video cameras and several sensors. The system is used to evaluate gross motor function and response to instructions from a therapist.

Methods: The experiments performed included a hand bell-playing task and several non-musical tasks, such as preparation of the hand bells and returning the bells to the storage box. The evaluation indices were (1) the strength of wrist-jerk movements, (2) the time of response to instructions for musical performance from the therapist, and (3) the time taken to perform non-musical tasks. Work was performed to clarify the correlation between the results of evaluation with the proposed method and those of an inquiry-based approach called the Achenbach System of Empirically Based Assessment (ASEBA), which is a standard screening method for developmental disorders.

Results: The results from the proposed system were more consistent with ASEBA results collected from therapists than with those collected from parents.

Conclusion: This indicates that the method can be used as a tool for conveying therapists' opinions to parents using the evaluated indices as objective evidence.

Keywords: Computer aided assessment, Developmental disorder, Music therapy.

Accepted July 08, 2016

Introduction

According to a survey conducted by Centers for Disease Control and Prevention, which is a national public health institute of the United States, the total number of children with developmental disabilities is increasing on a global basis [1]. To improve and support the development of these children's abilities, early diagnosis of disabilities and an effective means of early intervention are desirable [2].

Early-diagnosis methods have been extensively studied in recent years. One example is the Achenbach System of Empirically Based Assessment (ASEBA), which is a tool for developmental disorder assessment that can be used

both for adults and for children [3]. ASEBA is intended to provide comprehensive evaluation of psychosocial adaptation and maladaptive functioning based on 100 questions regarding the subject's behavior. In ASEBA, parents and teachers assess the child's behavior by answering these questions, and the answers are converted into multiple assessment scales (such as an introversion scale and a social ability scale) in order to identify the characteristics and problematic behaviors of the child. Based on the answers, then, a doctor diagnoses whether the child with a developmental disorder. The questions refer to the child's behavior over the preceding six months, which allows monitoring of development over time. ASEBA is today a standard diagnostic method for developmental

disorders because the scores are standardized for different populations and different cultures [3].

If a child is diagnosed with a developmental disorder, early intervention is desirable for the prevention of secondary disabilities and for support in forming interpersonal relationships [4,5]. Several early-intervention methods are generally employed, such as cognitive behavioral therapy (which improves self-awareness through interactive interviews), physical therapy (which supports development of motor function through posture and gait training) and music therapy (which trains children in both sociality and motor function through the playing of musical instruments and singing).

This paper focuses on music therapy in consideration of three points that are important for children with developmental disorders [6,7]: (1) Since music therapy enables communication with children through music without verbal language, intervention can be started at a very early age. (2) It has been reported that, based on the elicitation of rhythmic motility, music therapy can reduce tension and anxiety and facilitate physical exercise as well as self-expression. (3) Music therapy group sessions are effective for cultivating consideration and the ability to get along with others as capacities from which children derive self-esteem and a strong sense of identity. However, to the best of the authors' knowledge, no objective and quantitative method for evaluating the behavior of children with developmental disorders during the music therapy sessions has ever been proposed, and evaluation targets have therefore depended on the experience and subjective opinion of the therapist. More importantly, children's behavior generally varies widely in different circumstances, causing discrepancies in evaluation between parents and therapists. As a result, therapists face difficulties and the risk of miscommunication when delivering an opinion on a child to the parents without supportive or objective evidence.

As a first step toward solving this problem, the authors conducted a pilot study on a computer-aided motor behavior evaluation method for specific activities in music therapy sessions. Examination of motor behavior is a favorable starting point for future development of an evaluation system for music therapy because today's rapid progress in the field of image analysis is expected to lead to the capacity for wholly automatic evaluation in the near future. Two children diagnosed with Autism Spectrum Disorder (ASD) and one with Typical Development (TD) participated to the study. A music therapist instructed the children to hit hand bells, and indices including hand jerk and the lag time between the therapist's instruction and the ringing of each bell were evaluated using simple image analysis and auditory analysis methods. Although these indices may not cover the whole scope of music therapy, they can be considered to reflect part of children's overall behavior in such sessions. Accordingly, evaluation was performed to determine the correlation between these indices and the opinions of therapists and parents based on

comparison with ASEBA scores. ASEBA is advantageous because it allows evaluation by parents and teachers alike based on the same protocol. This enables clarification of discrepancies in evaluation results among evaluators, which can be caused by differences in circumstances among children. The system also supports testing to determine the feasibility of this computer-aided evaluation method for the conveyance of therapists' opinions to parents.

This paper is organized as follows: Section 2 gives an outline of music therapy, related work, Section 3 describes the proposed method and Sections 4 and 5 discuss the experiment and its results. Finally, Section 6 concludes the paper.

Music Therapy

Starting in the late 1960s, the effects of music therapy were demonstrated and proved through various experiments, stochastic analyses and measurement technologies [8]. Studies were actively conducted on pediatric patients, and focused on indicators such as respiration rate and the crying behavior of children. It was in this context that music therapy was adapted for children with developmental disorders such as autism, spectrum disorders and attention-deficit hyperactivity disorder. Recent studies have revealed that music therapy sessions improve joint attention, attention span and language development [9,10].

Engineered approaches to music therapy have also been proposed. For example, Oshima et al. [11] proposed a system that plays the music accompany with subjects clapping. Kurizuka et al. proposed a mutual adaptive system in which the therapist assists the walking motion of the patient by playing music with an optimal rhythm to improve the smoothness of walking movement [12]. However, the purpose of these systems is to restore motor function to elderly or dementia-stricken patients rather than to support the improvement and development of communication skills and sociality in children with developmental disorders.

In practical music therapy for children with developmental disorders, the therapist sets session targets in line with individual characteristics identified from behaviors observed in a previous session. After the session, evaluation is performed and is generally video-recorded to afford a deeper understanding of the child. The targets commonly defined in private music therapy facilities include using musical instruments properly, performing cooperative actions, singing while filling in missing words, using left-arm/right-arm/both-arm approaches to musical instruments and objects, and understanding cause-and-effect relationships. The total number of targets can exceed 100. To evaluate how well these targets are met, the therapist is required to capture the characteristics of both fine and gross motor functions as well as to assess response to instructions during the session [13-17]. However, this evaluation method is subjective and qualitative. Toward the establishment of an objective, quantitative evaluation method and the future development of a music therapy evaluation system, the next section outlines the technique examined in the pilot study.

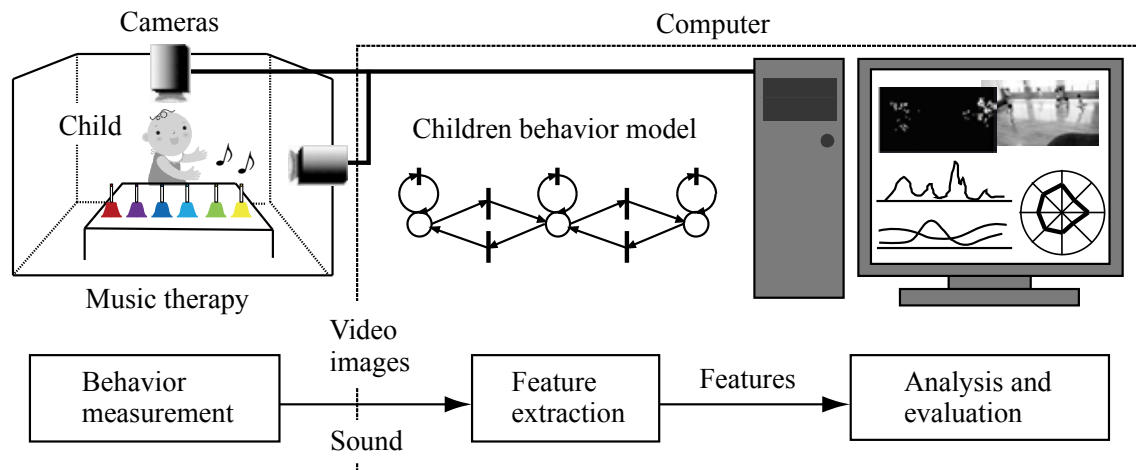


Figure 1. Overview of the proposed system

Computer-Aided Music Therapy Evaluation Method

Figure 1 gives an overview of the proposed method for evaluating the behavior of children during music therapy. The approach consists of a signal-measurement process, a feature-extraction process and a behavior-evaluation process, which together are used to evaluate (a) motor function and (b) response to instructions. The trajectories of the wrist during the playing of hand bells are analyzed to evaluate gross motor function, and the time taken for response to musical instructions as well as for task completion are also evaluated.

The proposed method is intended to test the feasibility of the computer-aided approach for quantitative evaluation representing the opinions of therapists. Once the approach is verified, modern technologies such as Kinect can be employed to establish a fully automated evaluation system. The following sections outline each process of the proposed method.

Target Task and Evaluation Target

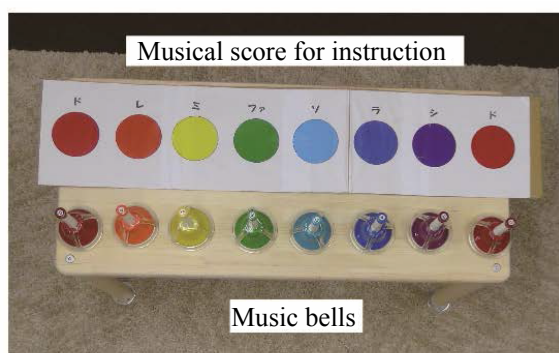
The target tasks and the evaluation target were configured to represent a compromise between achieving the aims of music therapy and facilitating computer-aided evaluation. The main target task was to play the hand bells following the instructions of the therapist. In addition, three types of non-musical tasks were also carried out, which are

preparing the hand bells, changing hand bells, and returning the hand bells to the storage box.

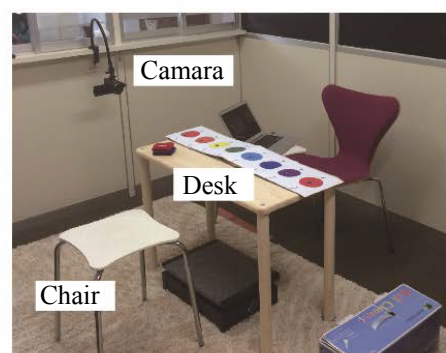
Although the goals of music therapy are not as simple as having the patient hit a bell in line with the therapist's instructions, this task allows partial evaluation of the music therapy targets described in Section 2. By way of example, gross motor function can be evaluated by tracking the wrist of a child reaching out to the bells, as such function in children with developmental disorders may differ from that in children with typical development. In addition, indices such as the lag time between instruction and hitting a bell and the number of failed attempts may be influenced by engagement with the task, and execution time for non-musical tasks may depend on children's levels of compliance. A key objective of this pilot study was to determine whether these evaluation targets reflect ASEBA scores given by therapists and parents.

Signal-Measurement Process

This process involves the use of two video cameras. One of these is installed on the ceiling and the other is fixed beside the table so that the behavior of the child can be observed from both lateral and overhead viewpoints. The cameras record images at f_0 [Hz] and sound at a frequency of f [Hz]. An instruction board with circles of eight colors corresponding to the colors of the hand bells is used, and the therapist gives instructions by pointing at



(a) Arrangement of the music bells



(a) Arrangement of the experimental equipment

Figure 2. Layout of the experimental environment

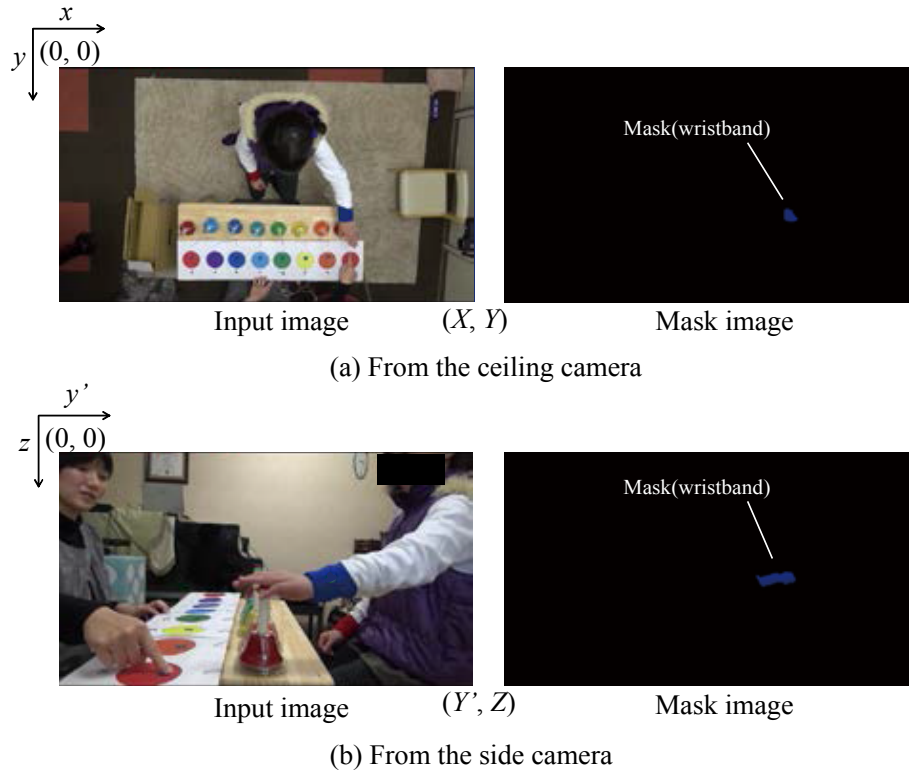


Figure 3. HSV images obtained from input images

color circles of the kind shown in Figure 2 (a). A touch sensor is attached at the center of each color circle to capture the instruction time t_i^{teach} and record it to the computer. Here i represents the scale of the hand bell, where $i = 1, 2, \dots, I$.

Feature-Extraction Process

Image processing

To evaluate the gross motor function of the arms from the recorded video (Figure 3), HSV components of a wristband attached to the child are extracted and the wrist's movements are tracked in each frame. First, as shown in Figure 3, each frame of the video is converted into both a brightness component image and a mask image, which are generated by extracting the area that has HSV values close to those of the wristband. Noise is then reduced by performing expansion and reduction processes on the mask image, and the contour with the maximum area is extracted and identified as the area of the wristband. The equations shown below give the center of gravity G as the wrist position of the child in frame l , where l is the total number of frames to be analyzed. Here the center of gravity of the z axial direction was determined using images from the video camera installed on the lateral wall.

$$G_x = \frac{1}{M^l} \sum_{x=0}^X \sum_{y=0}^Y x g^l(x, y) \quad (1)$$

$$G_y = \frac{1}{M^l} \sum_{x=0}^X \sum_{y=0}^Y y g^l(x, y) \quad (2)$$

$$G_z = \frac{1}{M^l} \sum_{y=0}^{Y'} \sum_{z=0}^Z z g^l(y', z) \quad (3)$$

$g^l(x, y)$ in Equations (1) and (2) and $g^l(y', z)$ in Equation (3) are pixel values in frame number l of the ceiling camera and lateral camera, respectively. When the pixel value (the value of the pixel extracted using the mask image) is 1, $g^l(x, y) = g^l(y', z) = 1$; otherwise, $g^l(x, y) = g^l(y', z) = 0$. The coordinates of the x -axis of the images taken from the ceiling camera are numbered as $x = 1, 2, \dots, X$ [pixels], and the coordinates of the y -axis taken from the lateral camera are numbered as $y = 1, 2, \dots, Y'$ [pixels], M^l being the total number of pixels in the mask area. When there is no area with HSV components close to the wristband, the mask image from the previous frame is used.

Frequency analysis

To measure the time it takes for the subject to ring each bell, the power spectrum of the audio signal $S(t)$ is computed using short-time Fourier analysis with window width ω and overlap ν . The total power $P_i(t)$ in the frequency band of $f_i - df$ to $f_i + df$ [Hz] corresponding to the i -th scale of the hand bell is then calculated. The time it takes for $P_i(t)$ to exceed the threshold of θ [%] above the power of sound in the environment is defined as the response time t_i^{child} ; that is, the time it takes for the child to ring the i -th hand bell.

Behavior Evaluation Process

The behavior evaluation process involves calculation of evaluation indices based on features determined from the feature extraction process.

Root mean square of jerk

Gross motor function during the playing of hand bells is evaluated using hand trajectories, with each hand movement between two bells being considered a reaching movement. Recent research has shown significant differences in reaching movement between children with typical development and those with developmental disorders [18-20]. For example, Mari et al. measured the reaching-to-grasp movement of children with autism spectrum disorder and compared the results to those of children with typical development. It was found that the two types of children differed in trajectory planning as well as in execution process [20]. It has also been reported that children with typical development are more sensitive to biological motions characterized by small jerks [21]. However, to the authors' knowledge, no studies have employed reaching-motion models to evaluate children with developmental disorders. There are three major models for describing reaching movement: (1) the minimum-jerk model, which is based on the assumption that humans naturally select the smoothest trajectory connecting the start and end points [22]; (2) the minimum torque-change model, which introduces joint dynamics to the minimum-jerk model and is based on the assumption that humans select the motion trajectory that minimizes variation in joint torque [23]; and (3) the minimum variance end-point error model, which is based on the assumption that humans select the trajectory that minimizes the effect of biological noise generated by muscle and neuronal activity [24].

As this study focused on the smoothness of reaching movement, the minimum-jerk model was employed to evaluate the motion of children. This model predicts the trajectory of reaching motion by minimizing the following cost function when the movement duration t_f is determined [22]:

$$C_j = \frac{1}{2} \int_0^{t_f} \left(\frac{d^3x}{dt^3} \right)^2 + \left(\frac{d^3y}{dt^3} \right)^2 + \left(\frac{d^3z}{dt^3} \right)^2 dt. \quad (4)$$

The analytical solution can be derived by applying the variation method to Equation (4). As a result, the trajectory of x -axis $x_{\text{sim}}(t)$ can be expressed by the fifth-order function

$$x_{\text{sim}}(t) = x_f (-6\tau^5 + 15\tau^4 - 10\tau^3) \quad (5)$$

where, $\tau = t/t_f$, t is time, and x_f is the end point of the reaching movement. Velocity and acceleration can be assumed to be 0 at both the starting point and the end point. The trajectories on the y - and z -axes can be derived in the same manner. To reduce the noise component generated by image processing, a simple moving-average K -order method is employed. By differentiating Equation (5), velocity can be calculated, and the waveform of velocity in the ideal reaching movement takes on a bell shape. Using the mean sum square error, this bell-shaped velocity was compared with the velocity actually measured in children. In addition, in order to evaluate the smoothness of reaching movement, the effective value of jerk J_{rms} is calculated

because less jerk indicates smoother motion. The effective jerk value is then calculated based on the root mean square (RMS) over time as in the following equation:

$$J_{\text{rms}} = \sqrt{\frac{1}{N} \sum_{i=1}^N j(t)^2}, \quad (6)$$

where N is the total number of samples collected during a tune.

Number of failures and lag time

To evaluate the child's response to the therapist, calculation was performed to determine the number of failures and the lag time R_i between the therapist's instruction and the ringing of each bell. Failure here is defined as the child's tapping of a different bell from the one the therapist indicated. Lag time R_i is calculated by subtracting time t_i^{child} (when the bell is rung) from time t_i^{child} (when the therapist gives the instruction).

Task execution time

The time required to complete non-musical tasks was also evaluated. The behavioral transit caused in a child by the therapist's instruction can be described using an infant behavior model previously proposed by other authors based on the Petri-nets theory [25].

Figure 4 shows a schematic diagram of the model that describes the states of the child and the therapist's instructions. In Figure 4, places P , P' and I respectively represent the state of a child who is on task, the state of a child who is off task, and the instruction from the therapist. T is the transition between states in each task. The current behavioral state of the child is represented by the token of a solid black circle, and the current instruction is represented by the token of a solid gray circle. When the therapist gives an instruction (for example, "Put away the hand bells"), the solid gray circle moves to the corresponding place I . This enables the child's token to make a transit to the instructed place P . If the child does not comply with the instruction, the token either does not make a transition or makes a transition to an off-task place. In this manner, the Petri-net-based model can visually describe the behavior of the child. This means that the therapist can employ the model in evaluating on-/off-task states by using it to track transition times between places. However, as determining whether the child is on or off task requires subjective evaluation, it is difficult to clearly separate the behavioral states of the child. Accordingly, the time $T^q = T_2 - T_1$ required for task completion was simply calculated, where q represents the task number, T_1 represents the time when the therapist gives an instruction, and T_2 represents the time when the child completes the task. As T^q increases along with off-task time, this index serves as an indicator for the evaluation of off-task behavior.

Comprehensive evaluation score

The proposed method outputs a comprehensive evaluation score through the following procedure. First, the indices measured from a child subject were compared with those

collected from a typical development group using t -tests. The comprehensive evaluation score was then defined as the average of t -values where significant differences were found. As t -values express the distance between two groups in t -distribution, the comprehensive evaluation score describes the difference from the typical development group. This definition was derived from Achenbach System of Empirically Based Assessment (ASEBA) in which behavior indicators are calculated from the t -score compared to the typical development group [3]. In the same manner, ASEBA scores were also averaged over all indices for comparability to the defined scores.

Experiments

To verify the accuracy of the proposed method and indices, the behavior of children in a music therapy facility was monitored. The measurement and evaluation results for two children diagnosed with autism spectrum disorder (ASD) were compared with those for a child with typical development (TD). In addition, evaluation results from the proposed method were compared with ASEBA scores collected from the parents and the therapist.

The experimental environment and method configuration are described in this section.

Subjects and Therapist

Three children (Sub. A: 9 years old/female/TD; Sub. B: 7 years old/male/ASD; Sub. C: 7 years old/male/ASD) and a music therapist participated in the experiments. Monitoring and analysis were carried out for each child. The monitoring was conducted in a private music therapy facility under parent/therapist supervision in accordance with the Declaration of Helsinki to ensure that sufficient care was taken with the children and to prevent their exposure to excessive risks and burdens. The ethics committee of the music therapy facility approved the monitoring and analysis protocols. The parents provided written informed consent for their children's participation in the experiments.

The music therapist who participated in the experiments is certified with the Japanese Music Therapy Association and the Certification Board for Music Therapists.

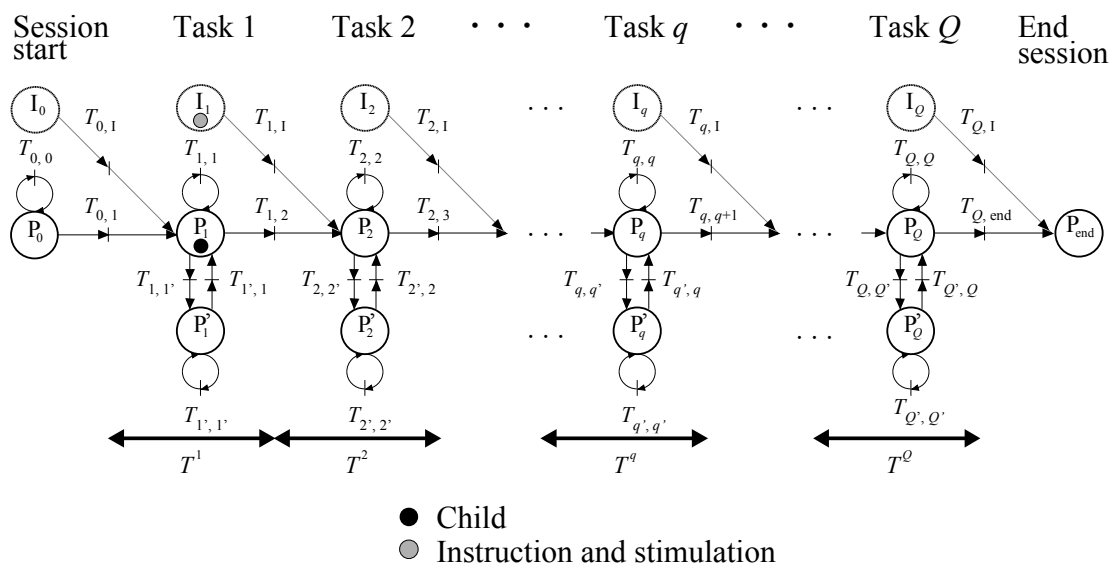


Figure 4. Petri-net model for behavioral evaluation of children

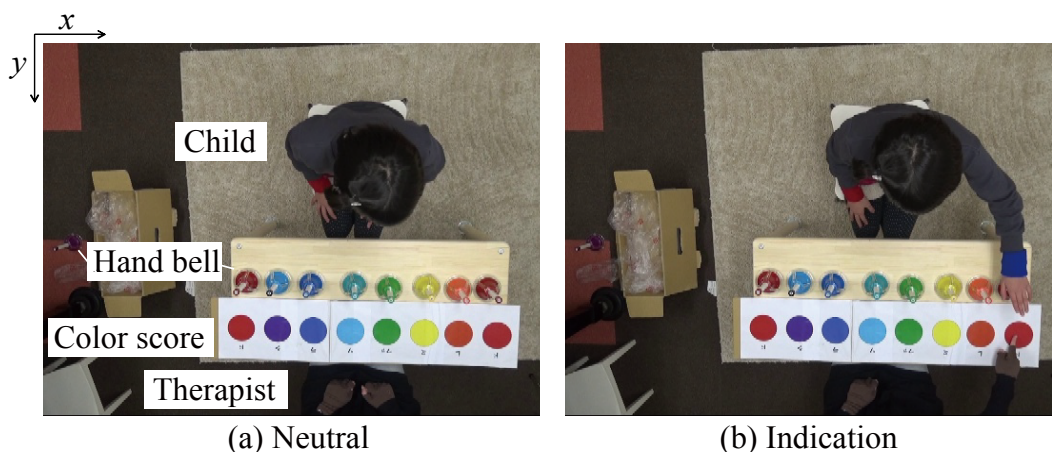


Figure 5. Images of a child during experiments

Experimental Environment

Figure 5 shows the hand bells and musical score used in the experiment as well as the experimental environment. A camera is fixed on the ceiling about 2.4 [m] from the floor (ceiling camera). The other camera is installed on a fixture attached to the wall and positioned approximately 0.4 [m] from the desk and 0.8 [m] above the floor (wall camera). The behavior of each child during the music therapy session was simultaneously recorded from these vertical and horizontal viewpoints. As shown in Figure 5, during the session the child and the therapist sat on chairs facing each other across a wooden table (width: 0.9 [m]; depth: 0.4 [m]; height: 0.6 [m]) on which the hand bells were aligned. The type of hand bell (ZEN-ON Co., Ltd., Tokyo: music bell, color-touch type) used for the experiment is a common instrument in music therapy and can be rung by hitting a button on its top.

Experimental Protocol

The repertoire used for the experiments included "Mary Had a Little Lamb" (2/4 time) and "My Grandfather's Clock" (4/4 time). These two tunes were chosen because the subjects had played them before so would not be confused during the experiments by having to play new tunes. Another reason was that the two tunes have different difficulty levels: "Mary Had a Little Lamb" is a simple tune employing only four notes, whereas "My Grandfather's Clock" is a relatively difficult tune that has a longer playing time and employs eight notes. The frequency of each note of the hand bells used in the experiments is shown in Table 1 ("Mary Had a Little Lamb") and Table 2 ("My Grandfather's Clock").

The experiments were arranged to require the subjects to alternately perform non-musical tasks and a musical

Table 1. Musical scales and corresponding frequencies: *Mary Had a Little Lamb*"

Musical notes	Frequency [Hz]
C ₄	1046.5
D ₄	1174.65
E ₄	1318.51
G ₄	1567.98

Table 2. Musical scales and corresponding frequencies: *My Grandfather's Clock*"

Musical notes	Frequency [Hz]
C ₄	1046.5
D ₄	1174.65
E ₄	1318.51
F ₄	1396.91
G ₄	1567.98
A ₄	1760
A# ₄	1864.65
C ₅	2093

task. There were a total of five sequential tasks arranged as follows: (1) prepare the hand bells, (2) play "Mary Had a Little Lamb," (3) change hand bells, (4) play "My Grandfather's Clock," and (5) return the hand bells to the storage box.

Parameter Configuration of the Proposed Method

Monitoring part

For the musical task, the therapist gave instructions to the subjects by pointing at color circles corresponding to the colors of the hand bells. A touch sensor attached to the center of the circle sensed each pointing action, and the computer received the signal from the sensor and recorded the timing of the instruction. The resolution of the camera was 720×480 [pixels], and its frame rate was $f_0 = 29$ [Hz]. The sampling frequency of the audio signal was $f = 44.1$ [kHz].

Feature extraction part

For the short-time Fourier transformation, the parameters of window width $\omega = 34.5$ [ms] and overlap width $\nu = 17.3$ [ms] were set. The frequency margin for detecting notes was set to $df = 20$ [Hz], and the detection threshold was set to $\theta = 10$ [%]. It should be noted that when the same note was continuously played or when the sound from the hand bell was not loud enough even though the subject had hit the bell, it was difficult to systematically determine the time at which it was rung. In such cases, the time was manually extracted using movie-editing software (Corel Corporation, VideoStudio Pro X4).

Behavior evaluation part

The data number for the moving average was set to $K = 5$. The end of the reaching movement x_f and the time taken for its completion t_f were manually extracted from the video footage recorded. The start time T_1 and completion time T_2 of the action task were extracted based on the instruction messages from the therapist. Movie-editing software (Corel Corporation, VideoStudio Pro X4) was used to carry out these procedures.

Calculations of jerk and lag time R_i were normalized based on beats per second (bps) for each tune to compensate for variation in instruction tempo. The root mean square of jerk was calculated for each bar of each tune and failed trials in which subjects hit incorrect hand bells were excluded from the calculation of average time lag.

In addition, the numbers of failures were manually counted by comparing the power $P_i(t)$ of the sound signals of the i -th note with the signals from the touch sensors that indicated the therapist's instructions.

Results

Evaluation of Reaching Movement

Figures 6a–6c shows examples of video images recorded while a subject was playing "Mary Had a Little Lamb."

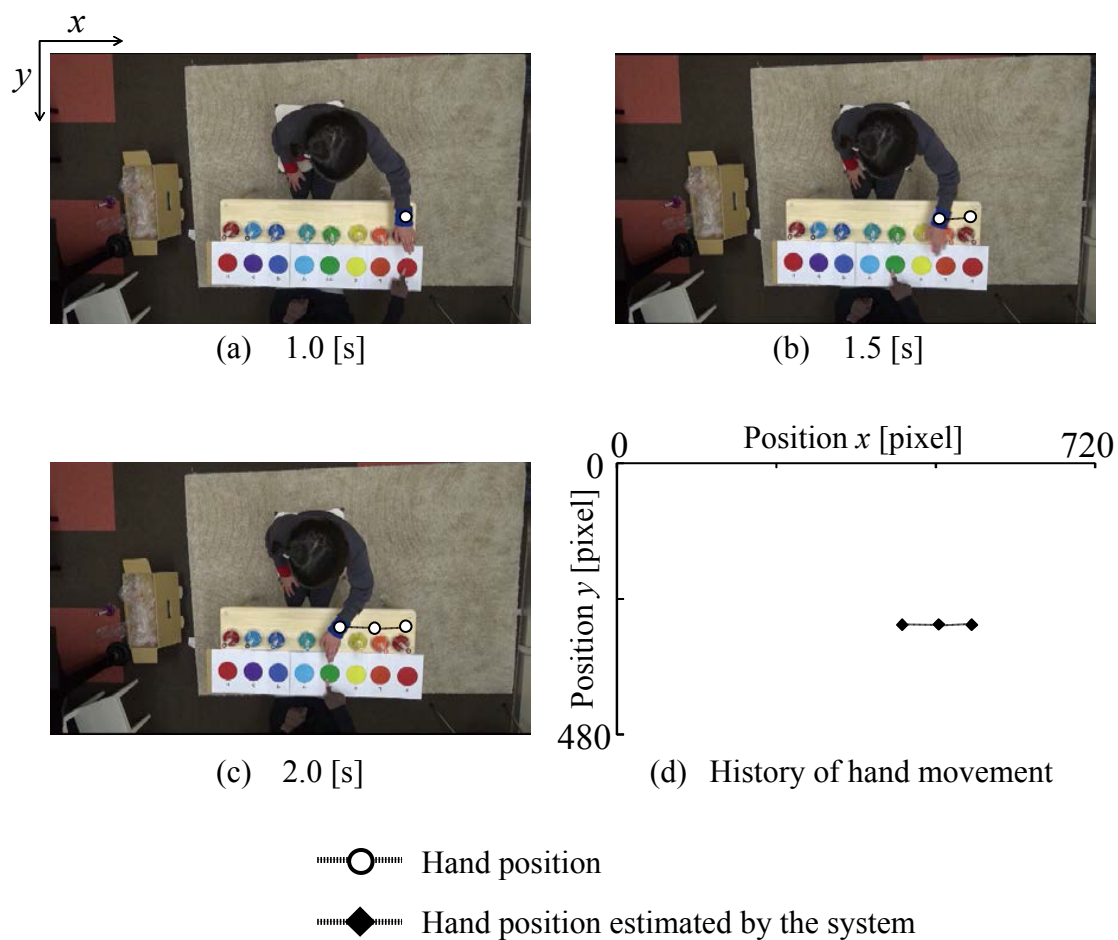


Figure 6. Example of video images and wrist movement trajectory

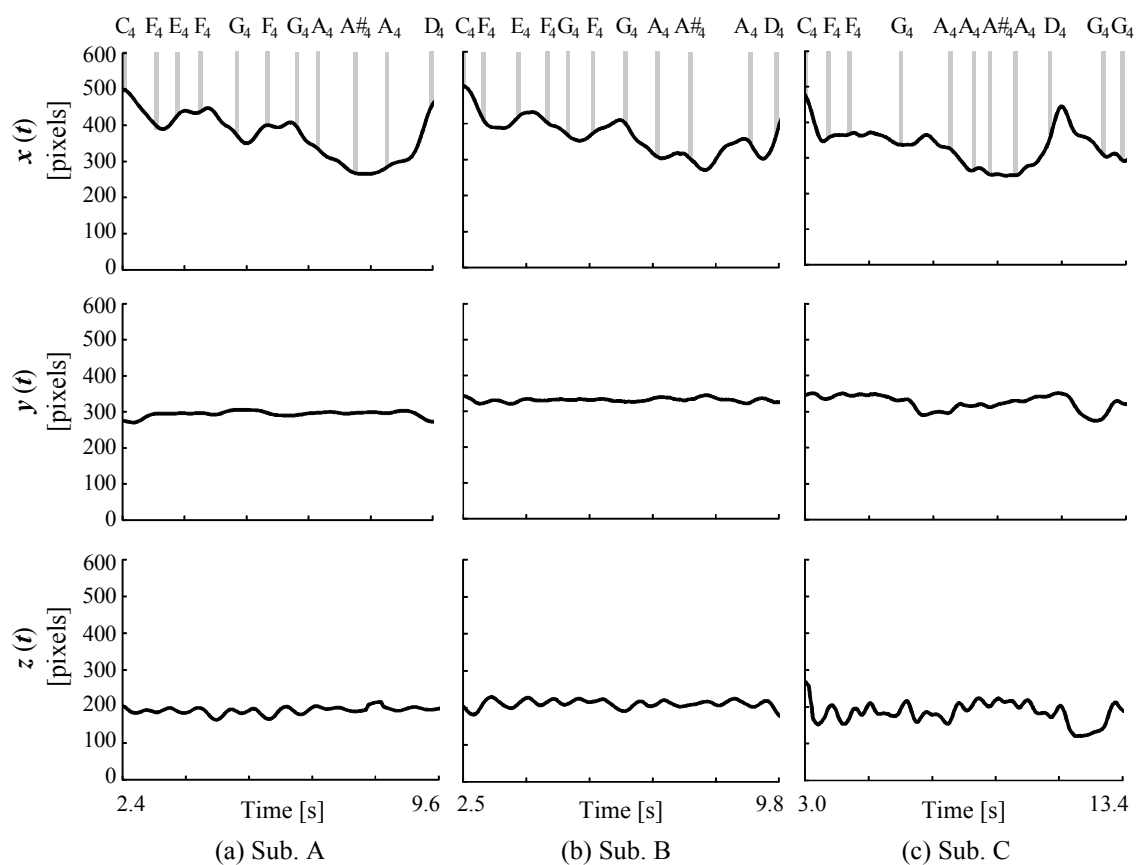


Figure 7. Wrist joint trajectories

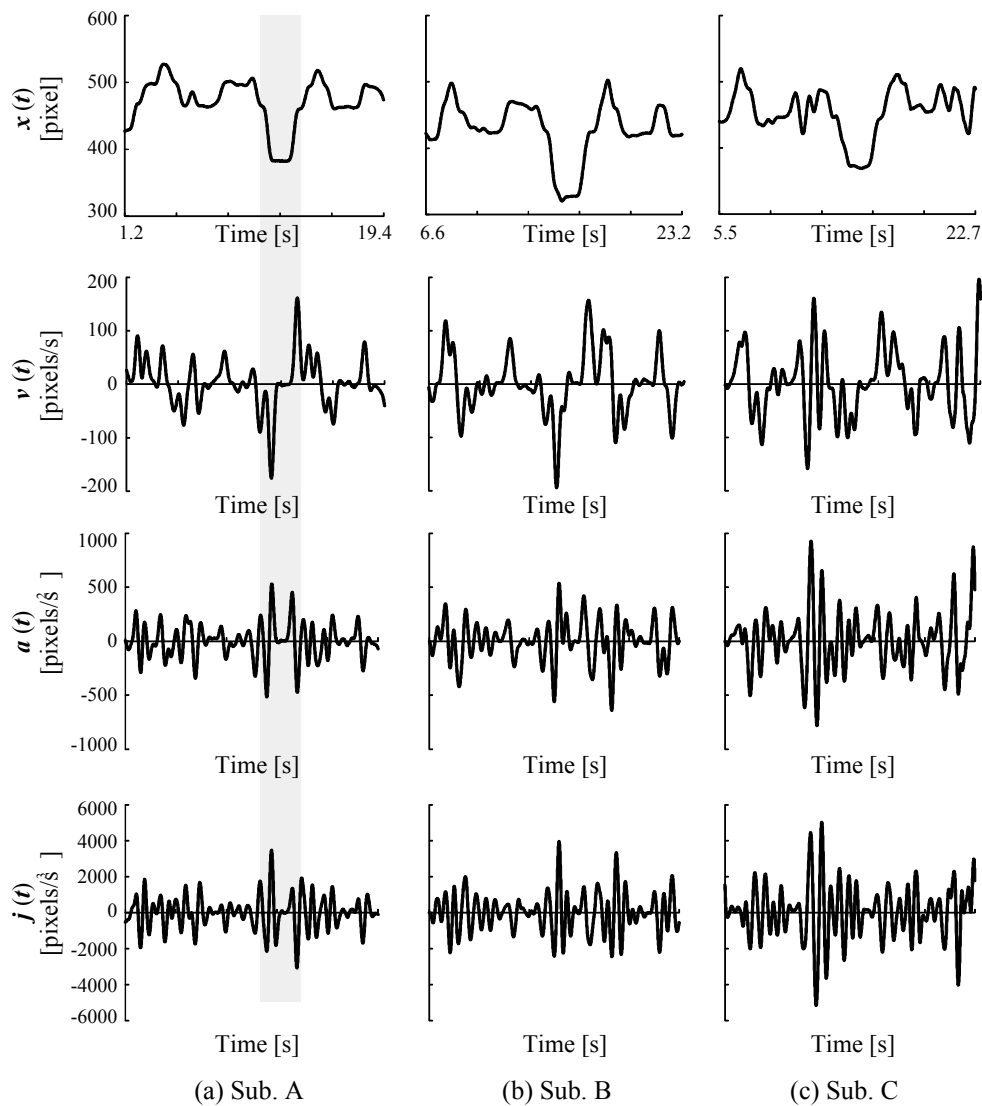


Figure 8. Wrist position measurement results

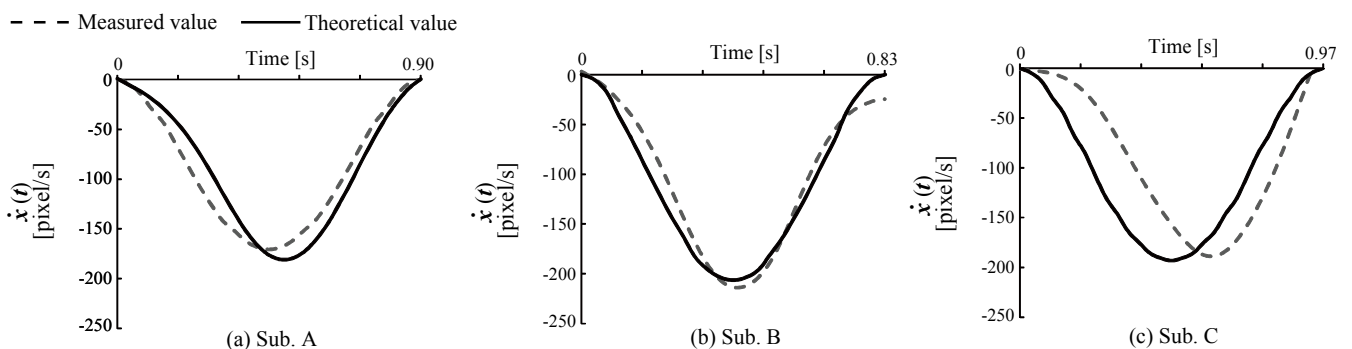


Figure 9. Example of measured wrist velocity compared to theoretical velocity calculated using the minimum-jerk model

Figure 6d shows trajectories of the wrist position extracted using the proposed method. For comparison, the position of the wristband as extracted by visual observation is also shown in Figure 6c. These figures indicate that both trajectories are very close, which confirms the tracking ability. Figure 7 shows wrist positions (x, y, z) extracted from a session. It can be seen that changes in position on the y - and z -axes were small compared to those on the x -axis; because of this, focus was placed on motion along the x -axis for the subsequent evaluation. Figure 8 shows an example of time-

series variations in position $x(t)$, velocity $v(t)$, acceleration $a(t)$ and jerk $j(t)$ among subjects playing the same tune. The shaded areas in Figure 8a show the interval that required each subject to sequentially play bells placed more than 300 [mm] apart. In this interval, position $x(t)$ and velocity $v(t)$ changed the most, which confirmed that movement of the hand had been successfully extracted. Focusing on jerk $j(t)$, the highest amplitude of Subject A was nearly 2,000 [pixels/s³], which was much lower than that for Subjects B and C, whose amplitudes were as great as 4,000–5,000 [pixels/s³].

Table 3. Number of sessions for analysis of hand-reaching movements

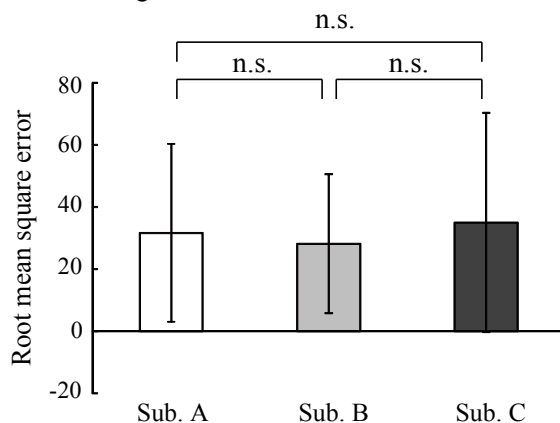
	Sub. A	Sub. B	Sub. C
"Mary had a Little lamb"	6	4	4
"My Grandfather's Clock"	12	12	8

Figure 9 shows examples of the measured velocity and the velocity calculated based on the minimum-jerk model in the interval that required each subject to sequentially play bells placed more than 300 [mm] apart. Table 3 shows the number of sessions used for analysis of reaching movement. Note that data were excluded from the calculation of jerk when (1) the hand position could not be extracted because the hands were hidden by the body, or (2) when the subject was not wearing the wristband. In the evaluation of reaching motion, additional data were excluded from Table 3 when such motion was interrupted for any of the following reasons: (1) the subject scratched his or her face during the reaching motion, (2) the subject moved his or her hand in the opposite direction from the target bell, or (3) the subject was unsure about which hand should be used to tap the bell.

The resulting numbers of reaching movements used for analysis were five for Subject A, six for Subject B and five for Subject C. Absolute discrepancies between the measured velocity and that calculated using the model were calculated for each sampling time, and the mean values and standard deviations are shown in Figure 9. This figure suggests that although the error trend of Subject C was slightly larger than that for Subjects A and B, no significant differences among the subjects were observed.

Figure 10 shows the average RMS of jerk j_{rms} . It can be seen that jerk for Subjects B and C had larger RMS values than that for Subject A. Considering the multiplicity of problems involved, stochastic comparison was performed using the Bonferroni method. As shown in Figure 11a, a significant difference with a 5 [%] significance level between Subjects A and C was detected in the simple tune

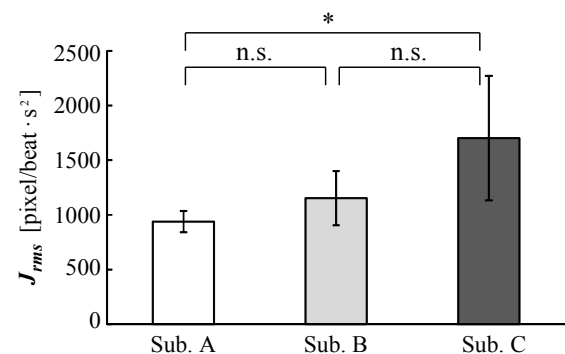
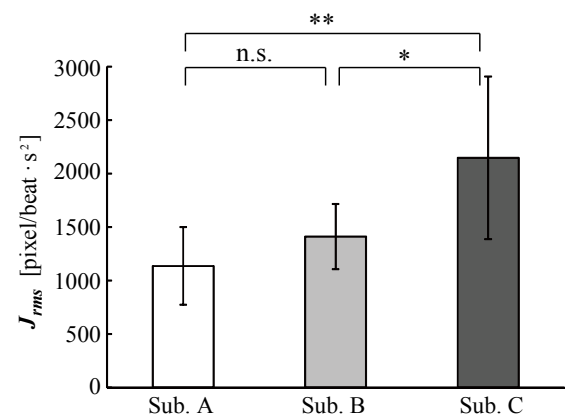
n.s. : Not significant

**Figure 10.** Average root mean square errors between measured velocities and those calculated using the minimum-jerk model

* : Significant level 5.0% ($p < 0.05$)

** : Significant level 1.0% ($p < 0.01$)

n.s. : Not significant

**(a) Mary Had a Little Lamb****(b) My Grandfather's Clock****Figure 11.** Average root mean square of measured jerk**Table 4.** Number of failures

		Sub. A	Sub. B	Sub. C
"Mary had a Little lamb"	Trail 1	0	2	2
	Trail 2	0	1	1
	Trail 3	0	0	3
"My Grandfather's Clock"	Trail 1	0	0	10
	Trail 2	3	2	5
	Trail 3	0	2	5

"Mary Had a Little Lamb", while Figure 11b indicates a significant difference with a 1 [%] significance level between Subjects A and C and a significant difference with a 5 [%] significance level between Subjects B and C in the difficult tune "My Grandfather's Clock". These results suggest that the hand motion of Subject C became more awkward with greater tune difficulty. This may have been caused by the subject's loss of concentration in the latter half of the tunes.

Evaluation of Response to Instructions

Table 4 shows the number of failures for each subject in each session. The data suggest that Subjects B and C tended to make more mistakes than Subject A did, which is consistent with video observations. While Subject A waited for and then confirmed instructions from the therapist before tapping the bell, Subjects B and C tended

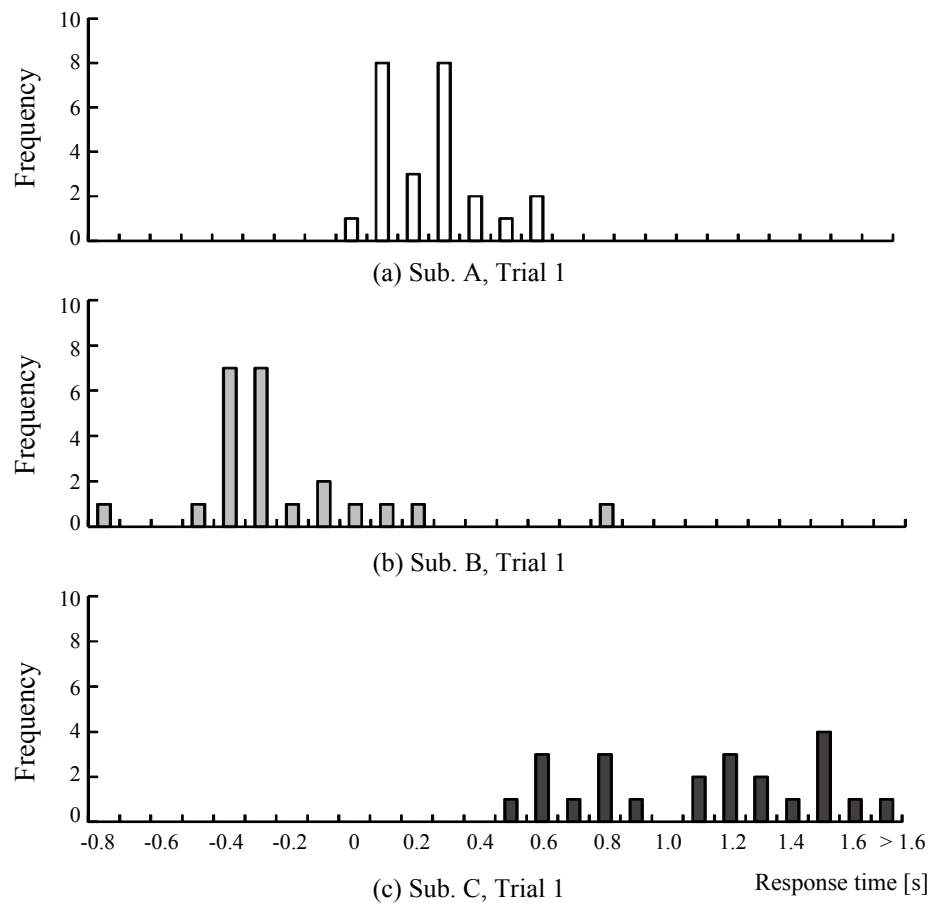


Figure 12. Histograms of time lag between therapist's instructions and children's responses. *Mary Had a Little Lamb*

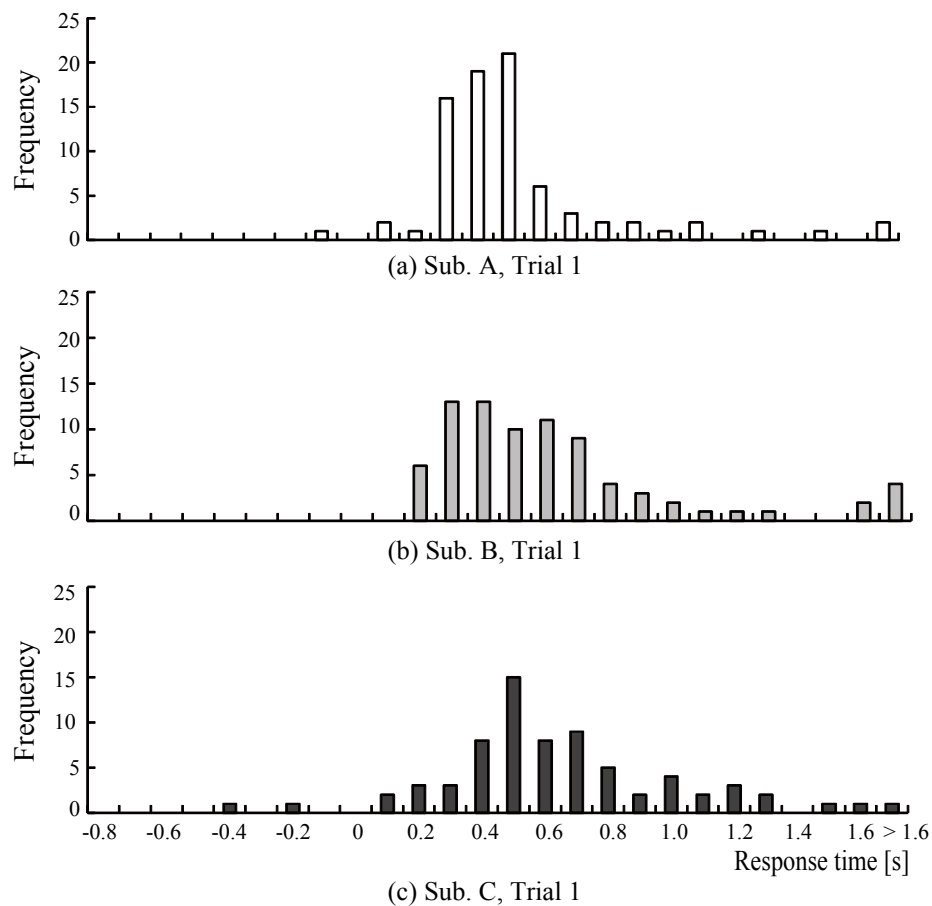


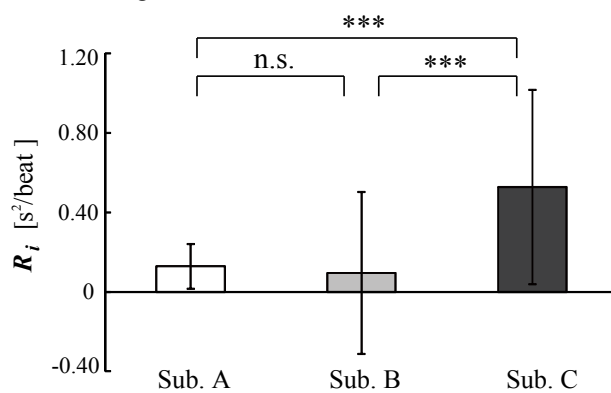
Figure 13. Histograms of time lag between therapist's instructions and children's responses. *"Mz Grandfather's Clock"*

to play the tunes to their own rhythms without reference to the instructions from the therapist.

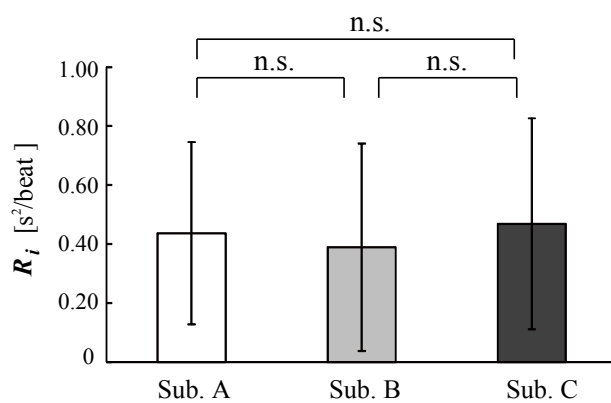
Figures 12 and 13 show histograms of lag time. Figure 12 suggests that Subject A had a relatively consistent lag time in responding to the instructions when playing "Mary Had a Little Lamb." Subject B tended to tap the bell before the instructions were given, while Subject C tended to tap well after instructions were given. The variance in lag time for Subjects B and C was much greater than that for Subject A. However, Figure 13 shows that when the subjects were playing "My Grandfather's Clock," these individual characteristics disappeared for each subject and the distribution of lag time became symmetrical and unimodal. The lag time also increased for all subjects.

Figure 14 shows the average lag time R_i for each subject. Figure 14a indicates a significant difference with a significance level of 0.1 [%] between Subjects A and C as well as between B and C in the simple tune. In contrast, Figure 14b indicates that there were no significant differences among the subjects in the difficult tune. As "My Grandfather's Clock" (which consists of 80 notes) is much longer and more complicated than "Mary had a Little Lamb" (25 notes), all the subjects were more careful when playing the former and this hid the individual characteristics.

*** : Significant level 0.1% ($p < 0.001$)
n.s. : Not significant



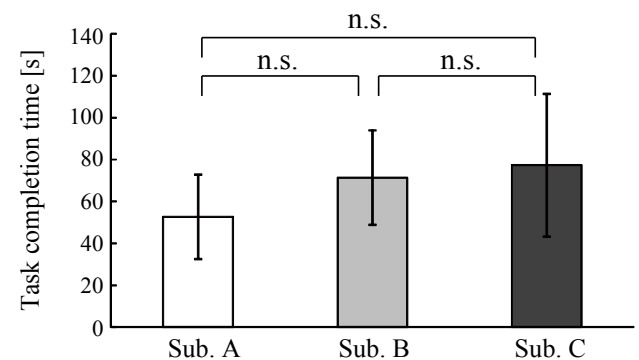
(a) Mary Had a Little Lamb



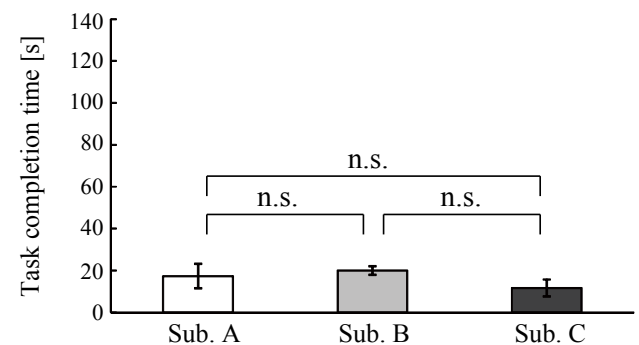
(b) My Grandfather's Clock

Figure 14. Average time lag

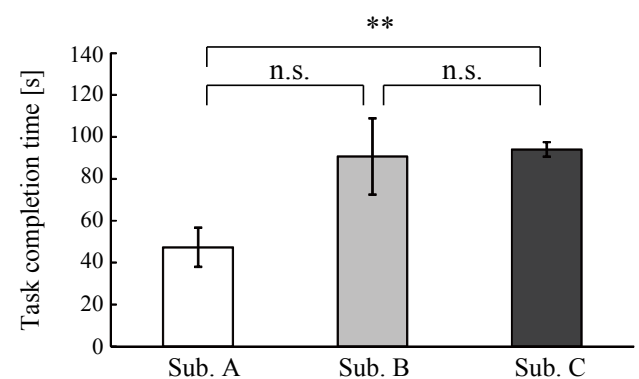
** : Significant level 1.0% ($p < 0.01$)
n.s. : Not significant



(a) Task1 : Preparation of the music bells



(b) Task2 : Changing music bells



(c) Task 3 : Clearing

Figure 15. Average task execution time

As described above, analysis of lag time enables quantitative evaluation of differences among the subjects. In addition, the evaluation results suggest that a simple tune such as "Mary had a Little Lamb," which has fewer notes to play, is suitable for observing individual characteristics.

Evaluation of Ability to Complete Non-Musical Tasks

Figure 15 shows the average task completion time for all trials. As a result of the t -test's assumption of equal variances in average task completion time, no significant differences were found among the subjects for Task 1 or Task 2. In contrast, as shown in Figure 15c, a significant difference with a significance level of 1 [%] was observed between Subjects A and C for Task 3. Here, considering the multiplicity of the problems involved, multiple

comparisons using the Bonferroni method were performed. As Task 1 and Task 2 posed the clear main goal of playing the tunes, all subjects were able to comply with the instructions. In contrast, the purpose of Task 3 was vague, which made Subjects B and C unable to concentrate on the task. The video shows that both subjects were reluctant to complete the task and ignored repeated instructions from the therapist. These results indicate that the proposed indices reflect the characteristic behavior of subjects in both musical and non-musical tasks.

Comparison with ASEBA Scores

Evaluation results from the proposed method were

compared with subjective evaluation results based on ASEBA scores collected from parents and a therapist (Kyoto International Social Welfare Center, Kyoto: Japanese version) [3]. The parents of each subject completed questionnaire CBCL6-18 and questionnaire TRF6-18. Although the content of these questionnaires differs somewhat, both are designed to facilitate investigation of children's behavior.

Figure 16 shows the results for each evaluation item. The horizontal axis in Figure 16 includes 20 evaluation items integrated from the questionnaires, and the vertical axis includes the scores for all evaluation items. Each

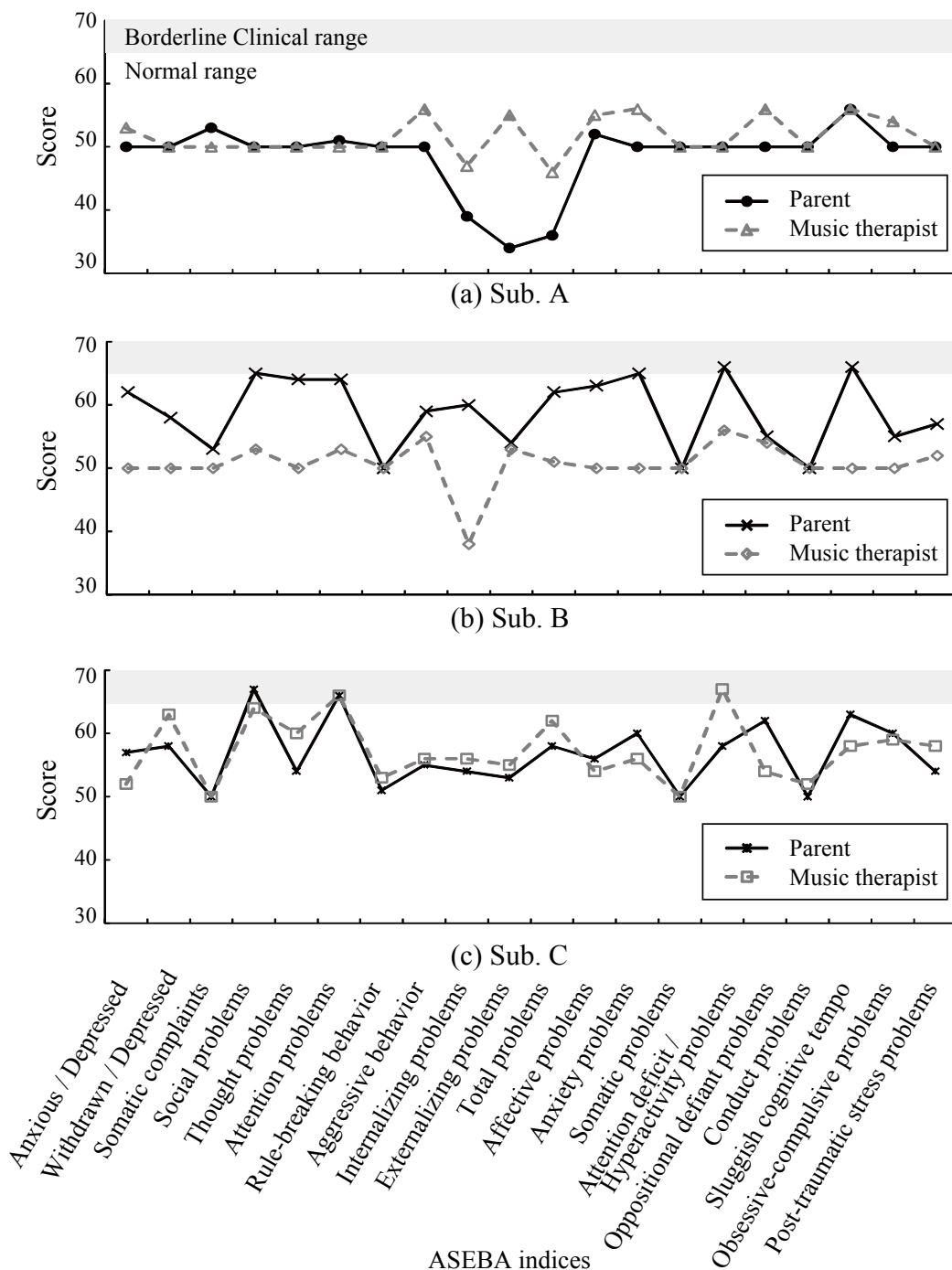


Figure 16. ASEBA scores collected from parents and therapist

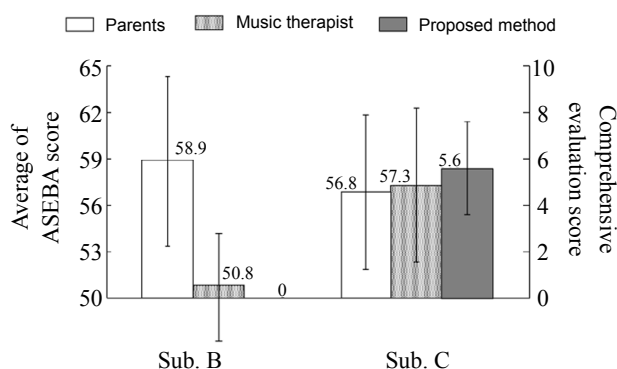


Figure 17. Comparison of average ASEBA scores and number of significant differences with Sub A

score ranges from 0 to 100, and those exceeding 65 are considered characteristic of a developmental disorder. The solid line in Figure 16 represents the evaluation results collected from the parents (CBCL6-18), and the dashed line represents those from the therapist (TRF6-18).

For Subject A, the evaluation results from the parents were roughly consistent with those of the therapist, as shown in Figure 16a. This suggests that Subject A is within the range of typical development. For Subject B, the evaluation results from the parents showed several high scores; specifically, six items were scored high enough to be considered indicators of developmental disorder, as shown in Figure 16b. The scores from the therapist were lower than those from the parents, and all items were within the range of typical development. For Subject C, the parents and the therapist shared common trends in their evaluations except in the case of several items for which the therapist gave higher scores, as shown in Figure 16c. As seen here, correspondence in evaluation results from parents and therapists cannot necessarily be expected because ASEBA evaluation is subjective.

Figure 17 shows the average scores of the ASEBA evaluations and the comprehensive evaluation scores of the proposed method. The resulting score for Subject B was 0 points and that for Subject C was 5.6 points. Interestingly, these outcomes are consistent with the evaluation results from the therapist as shown in Figure 17. This suggests that the proposed method can be used to convert the subjective opinions of the therapist into quantitative and objective indices, and can therefore be used to convey the ideas of the therapist to parents. Moreover, the evaluation results can be utilized in the design of subsequent music therapy sessions. By way of example, as the method indicated significant jerk in the reaching movement of Subject C, subsequent music therapy sessions should involve activities to improve the subject's gross motor skill.

Conclusion

This paper proposes a computer-aided evaluation method for specific music therapy activities and describes a pilot study conducted to determine whether the evaluation results support therapist opinions. The child's behavior was monitored during

the activity using video cameras, and related characteristics were identified via image processing and frequency analysis of audio signals. The child's commitment to music therapy can then be quantitatively evaluated using the proposed indices, which are based on consideration of some of the goals of music therapy.

A child with typical development and two children with developmental disorders (autism spectrum disorder, or ASD), all of whom attended private piano classes, participated in the experiment. Their behavior was evaluated from the viewpoints of exercise ability and response to instructions. For all proposed indices, significant differences were observed between the child with typical development and those with developmental disorders. Comparison of ASEBA scores from the therapist with the results of evaluation using the proposed method showed similar trends between the two. In addition, the parents welcomed the evaluation results and the video footage provided because this information supported understanding of children's behavior during music therapy.

As the proposed method can be used to objectively evaluate children's behavior, it serves as a tool for converting the subjective evaluation of the therapist into quantitative indices and for explaining the basis of subjective evaluation to parents. As the indices reflect response and motor function to clarify children's strong and weak points, they provide the therapist with reference data for subsequent therapy sessions and support recommendations for additional treatment, such as physical therapy.

Based on this pilot study, the authors plan to define more indices incorporating cumulative expertise on developmental disorders, such as metrics for line of sight and sitting posture, using recently developed image processing techniques and motion capture devices.

Source of Funding

This work was supported by JSPS KAKENHI Grant Number 15H01584.

References

1. Baio J. Prevalence of autism spectrum disorder among children aged 8 years - autism and developmental disabilities monitoring network, 11 Sites, United States, 2010: Surveillance summaries; National Center on Birth Defects and Developmental Disabilities, Center for Disease Control and Prevention, United States of America 2014; 1-21.
2. Cohen H, Amerine Dickens M, Smith T. Early intensive behavioral treatment: Replication of the UCLA model in a community setting. *Journal of Developmental & Behavioral Pediatrics* 2016; 27: S145-S155.
3. Achenbach TM. The Achenbach System of Empirically Based Assessment (ASEBA): Development, findings, theory and applications, Burlington, University of Vermont Research Center for Children, Youth & Families 2009.
4. Smith T, Groen AD, Wynn JW. Randomized trial of intensive early intervention for children with pervasive developmental

- disorder. *American Journal on Mental Retardation* 2000; 105: 269-285.
5. Mahoney G, Perales F. Relationship-focused early intervention with children with pervasive developmental disorders and other disabilities: A comparative study. *Journal of Developmental & Behavioral Pediatrics* 2005; 26: 77-85.
6. Boxhill EH. *Music therapy for the developmentally disabled*, NY. Aspen Publishers, Inc 1985.
7. Thaut M. *Rhythm, music and the brain: Scientific foundations and clinical applications*, NY. Routledge 2005.
8. Malone AB. The effects of live music on the distress of pediatric patients receiving intravenous starts, venipunctures, injections and heel sticks. *Journal of Music Therapy* 1996; 33: 19-33.
9. Kim J, Wigram T, Gold C. The effects of improvisational music therapy on joint attention behaviors in autistic children: A randomized controlled study. *Journal of Autism and Developmental Disorders* 2008; 38: 1758-1766.
10. Wigram T, Gold C. Music therapy in the assessment and treatment of autistic spectrum disorder: Clinical application and research evidence. *Child: Care, Health and Development* 2006; 32: 535-542.
11. Oshima C, Ito N, Nishimoto K, et al. An accompaniment system for healing emotions of patients with dementia who repeat stereotypical utterances: Toward useful services for elderly and people with disabilities. 9th International Conference on Smart Homes and Health Telematics 2011; 6719: 67-71.
12. Kurizuka Y, Miyake Y, Kobayashi Y. Walking support system based on musical exercise therapy. *SICE 2004 Annual Conference* 2004; 3: 2693-2697.
13. Gillberg C. Hyperactivity, inattention and motor control problems: prevalence, comorbidity and background factors. *Folia Phoniatica et Logopaedica* 1998; 50: 107-117.
14. Thelma MP, Jan PP, Hay DA. Fine and gross motor ability in males with ADHD. *Developmental Medicine and Child Neurology* 2003; 45: 525-535.
15. Yamamoto J, Kusumoto C. Development and support for autistic spectrum disorder. *Cognitive Studies* 2007; 14: 621-639.
16. Rogers SJ, Williams JHG. Imitation in autism: Findings and controversies, in imitation and the social mind and typical development, Sally J. Rogers and Justin H. G. Williams eds., NY. Guilford 2006: 277-309.
17. Boso M, Emanuele E, Minazzi V, et al. Effect of long-term interactive music therapy on behavior profile and musical skills in young adults with severe autism. *The Journal of Alternative and Complementary Medicine* 2007; 13: 709-712.
18. Green D, Baird G, Barnett AL, et al. The severity and nature of motor impairment in Asperger's syndrome: A comparison with specific developmental disorder of motor function. *Journal of Child Psychology and Psychiatry* 2002; 43: 655-668.
19. Johnston LM, Burns YR, Brauer SG, et al. Differences in postural control and movement performance during goal-directed reaching in children with developmental coordination disorder. *Human Movement Science* 2002; 21: 583-601.
20. Mari M, Castiello U, Marks D, et al. The reach-to-grasp movement in children with autism spectrum disorder: Philosophical Transactions of the Royal Society B. *Biological Sciences* 2003; 358: 393-403.
21. Cook J, Saygin AP, Swain R, et al. Reduced sensitivity to minimum-jerk biological motion in autism spectrum conditions. *Neuropsychologia* 2009; 37: 3275-3278.
22. Flash T, Hogan N. The coordination of arm movements: an experimentally confirmed mathematical model. *The Journal of Neuroscience* 1985; 5: 1688-1703.
23. Uno Y, Kawato M, Suzuki R. Formation and control of optimal trajectory in human multijoint arm movement. *Biological Cybernetics* 1989; 61: 89-101.
24. Harris CM, Wolpert DM. Signal-dependent noise determines motor planning. *Nature* 1988; 394: 780-784.
25. Migita R, Shimatani K, Shibasaki T, et al. A markerless monitoring system for behavioral analysis of infants using video images. *Japanese Journal of Human Growth and Development Research* 2014; 65: 1-7.

Correspondence to:

Zu Soh,
Institution of Engineering,
Hiroshima University,
1-4-1 Kagamiyama Higashi-Hiroshima,
Hiroshima, Japan.
Tel: +81-82-424-5763
E-mail: sozu@bsys.hiroshima-u.ac.jp