# Behavioral Learning and Imitation for Music-Based Robotic Therapy for Children with Autism Spectrum Disorder

Rachael Bevill, Srineil Nizambad, and Chung Hyuk Park, *Member, IEEE*
The George Washington University, Washington, DC.
{rbevill, skn5, chpark}@gwu.edu
Myounghoon Jeon
Michigan Technological University, Houghton, MI
mjeon@mtu.edu
Ayanna M. Howard, *Senior Member, IEEE*
Georgia Institute of Technology, Atlanta, GA
ayanna.howard@ece.gatech.edu

***Abstract*— Children with autism spectrum disorder (ASD) often are developmentally behind their peers in terms of emotional communication and social skills. Current research indicates that key features of robots make them suitable "assistants" for therapy sessions targeted at improving these skills. We propose a robotic framework for social and emotional scenario-based training for children with ASD, in which music and robotic behaviors will be fused while children's reactions are seamlessly analyzed with multi-modal sensors and robotic learning is applied to increase engagement. Through this framework, we aim to provide more enhanced treatment and intelligent progress assessment for the children and clinicians.**

## I. Introduction

The *Diagnostic and Statistical Manual of Mental Disorders – Fourth Edition (DSM-IV)* defines autism as "the presence of markedly abnormal or impaired development in social interaction and communication" [1]. Children with autism may have trouble with interaction and communication. They may have difficulty understanding the emotions of others, and with displaying emotions of their own. Children on the autism spectrum may also experience sensory input differently than neurotypical children – for example, lights may seem brighter, to the point of being painful. Such a situation is referred to as sensory overload. For children with autism, especially with non-verbal conditions, it is crucial to being comfortable and being able to avoid sensory overload situations when expressing their emotions, whether through facial expressions or gestures.

Robotic therapy is being researched with great interest as a viable method for a therapy for children with autism. Children with ASD find robots more approachable than human strangers [5]. Additionally, children in a therapy game that were paired with robot mediators displayed increased shared attention and facial expression imitation than their counterparts that were paired with a human mediator [5]. Children with ASD tend to find the robots more interesting and engaging than a human therapist, making therapy feel more like play than work [10].

Emotions displayed by humans are complex, being a combination of facial expressions, posture, and movement. Inadvertently, a human may display a mixture of multiple emotions. This can also be done intentionally – for example, by using sarcasm. This clash of emotional messages can be confusing for a child with ASD. Due to their simplicity and precise programmable nature, robots excel at separating and "articulating" emotions in comparison.

In our work, we propose the use of an autonomous, interactive robotic framework, with a focus on dance imitation and interactive behavioral play, to stimulate the emotional communication and social interaction skills of children with autism. The details of this approach are described in the following sections. Section 2 explains the relationship between emotion and body language, and how we plan to track this relationship. Sections 3 and 4 discuss previous and on-going research regarding the use of music and imitation in therapy for children with ASD, and explain why and how we intend to incorporate these aspects into our robotic framework. Section 5 describes a literary analysis of current motion learning techniques. Section 6 presents the layout of our preliminary study. Section 7 entails the results of our preliminary study, our analysis and our conclusions. Finally, Section 8 describes our on-going and future work, building off the literary analysis of previous sections.

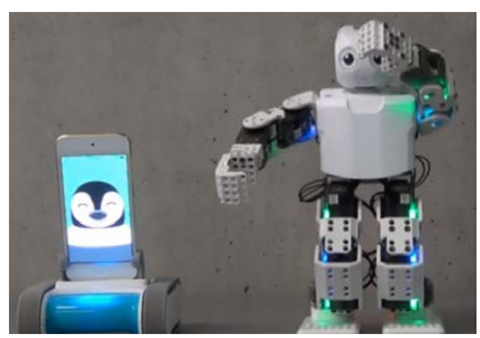

**Figure 1.** Romo (left), our modified penguin character, and Darwin Mini (right), both express the emotion "happy" – one through facial expression, and the other through movement.

## II. Emotion and Body Language

Microsoft Kinect is an RGB-D camera that can motion capture the changes in the dynamics at joints, tracking the positions of about 30 joints in a human subject. We develop a custom human skeletal-motion tracking and evaluation software using Microsoft Kinect SDK and our custom algorithms to track the engagement of the children with our robotic system.

Laban Movement Analysis (LMA) is a well-known methodology for categorizing and interpreting the differences in human movement. We look at the "effort" category of LMA, which provides quantitative data through its use of the subcategories weight, space, time, and flow. These categories are used to define various aspects of movement. We measure these categories by sorting the Kinect data into a weight term, time term, and motion units.

Eq. (1) below describes the weight term, where the average joint torque for the total number of joints measured $N$ is calculated depending on the radius, acceleration, and mass of each joint $i$ tracked along with the angle between the force generated and the axis of rotation [8]. Eq. (2) presents the weight term, which is the average of each angular velocity quantity (ω) at each joint $i$ measured [8] for $N$ joints.

$$LMA_{Time} = \sum \frac{\tau_i}{N} = \sum \frac{r_i * F_i * \sin(\theta_i)}{N} \qquad (1)$$

$$LMA_{Weight} = \sum \frac{\omega_i}{N} \qquad (2)$$

The weight term in Equation 1 is measured in frames per second (fps), and corresponds to the force of movement exerted by the body. It indicates movement which is either strong or light, in LMA phraseology. The time term in Equation 2 is also measured in fps and indicates the "rhythm" of a movement – that is whether a movement sequence is sustained or sudden. Motion units are a way of measuring the amount of motional activity in different movements/gestures, such as waving a hand, based on set threshold differences between angular joint velocities during the movement period. A movement that generates greater motion units indicates more varying velocities and more sudden movement nature and the pattern of the motion units plot over time provides an indication of the movement's smoothness.

We will partly draw on Laban Movement Analysis to understand the body language of subjects as they interact with the social robots. A number of rising studies indicate strong correlations between kinematic features, including LMA features, and emotional states. [9] Emotions can be identified using typical movement dynamic markers. For example, a person who is sad may present this through slow, sullen movements. This would appear as a small average velocity, low acceleration, and low frequency of acceleration. A person who is angry may move aggressively or erratically. This would appear as movement with high acceleration, high frequency of local acceleration, and high average movement velocity. These kinematic features can all be tracked via Kinect.

## III. Music Therapy and Autism

Music may also serve as a key component of successful therapy for autistic patients. Children with ASD tend to have low eye contact, and difficulty focusing. This leads to difficulty in interacting with their peers. Studies show that music can promote interpersonal responsiveness in ASD patients [7].

In another study on autism and improvisational music therapy, where the child either interacted with musical instruments or with toys, the researchers found that the musical sessions evoked more frequent and longer reactions of "joy," "emotional synchroncity", and "initiation of engagement" behaviors in the children than in toy play sessions" [7]. The children were also more receptive to the therapist's requests during the music sessions than in the toy-play sessions. Music therapy may therefore be an ideal setting for practicing emotional communication.

## IV. Imitation and Autism

Imitation is a crucial form of learning. From birth, and through the early stages of development, children learn by imitating – especially concepts like speech, appropriate social responses, and how to convey feelings. Children learn how to communicate through imitation. For children with autism, these milestones are often delayed.

In therapy sessions, "Imitation and turn-taking games are used…to promote better body awareness and sense of self, creativity, leadership and the taking of initiative" [11]. Studies have shown that using imitation in therapy sessions can elicit increased levels of participation in autistic children [4].

In one study, the experimenter performed three sessions of interactive play with patients with ASD. The experimenter could choose the same object as the child and imitate their method of play, choose the same object but perform different interactions, or choose a different object and perform a different method of play. The study concluded that the session with imitation of both object and style of play "resulted in greater frequency and duration of object manipulation" [12].

Another study found that social behaviors, such as vocalizing, eye contact and proximal behaviors improved consistently after each session of imitative play [6].

Overall, imitation appears to be another promising method of therapeutic intervention for children with autism. We propose using imitation in our work in conjunction with music therapy, by programming our robotic system to imitate the dance moves of the child, in order to stimulate engagement and emotional communication.

## V. Motion Learning

Just like humans, robots are starting to learn through imitation as well. The following is an evaluation of existing motion learning methods, as well as our proposals for what methods would be suitable or adaptable for our robot and intended goals.

Motion learning, also referred to by a handful of other names including Robot Programming by Demonstration and Learning by Imitation, "explores methods to teach a robot new skills by user-friendly means of interaction" [3]. In some situations, it makes sense to hard-code the responses of robots. However, this is not the case for robots whose purpose is to interact with humans. No two humans will react in exactly the same response. While we could pre-program the robot with a set of standard dance moves, it would be impossible to predict and emulate every dance move the future patients could make. This is the reason that motion learning is necessary.

All movements, no matter how complex the choreography, can be broken down into basic, core motions called "primitives". All motion can be represented through the scaling and shifting of these primitives. The computational control method of manipulating primitives is called "Dynamic Movement Primitives", or DMPs. In motion learning, the robot first learns the generic primitive, then what parameters can be adjusted, and finally in what fashion these adjustments can occur. Based on feedback from the environment, the robot can adjust the primitive as needed to complete its task. For our research, the robot will need to be able to identify the move sets closest to that of the child's current dance pose, and modify its movements accordingly.

Calinon *et al.* note that it is important to take the physical constraints of the robot into consideration when defining the core features of a move set [3]. For example, OP2 would not be able to jump. However, it could mimic other features, such as waving its arms, or kicking one foot. Knowing these limitations provides you a better understand of the robots capabilities, and allows you to decide how to best "train" the robot. If a robot is not physically able to perform a task, modifications should be made, and also programmed in as the proper response to take instead.

There are two methods for encoding motion learning into a robot: movement demonstration, and kinesthetic teaching [3]. With movement demonstration, the human teacher is tracked by the robot. This can be accomplished with the human wearing motion sensors, for example. With our robotic framework, OP2 will be able to track the human teacher's movements through the use of the Microsoft Kinect skeletal tracking system. After observing and analyzing the human teacher's movements, the robot attempts to mimic the move set. This leads to method number 2, kinesthetic teaching. Once the robot attempts to copy the move, the human teacher will physically manipulate the robot in order to position it in a way that best meets the requirements of the motion intention. This process is repeated, through various environments, until the robot can accurately produce appropriate variations of the move set in all situations.

In mathematical terms, the movement demonstration phase is a series of data acquisitions. The joint angle trajectories are recorded and plotted against time. A latent space of motion is necessary to project these data points on to, and results in a projection matrix. A Gaussian Mixture Model (GMM) is used to define the limits of the motion parameters. From a GMM, we can determine the probability, mean, and covariance of the move set. Next, a Gaussian Mixture Regression (GMR) is used to determine the generic base of the move set, based on the calculations of the GMM.

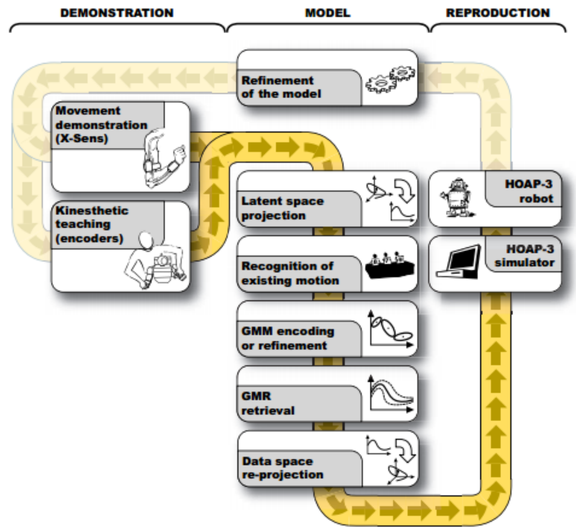


**Figure 4.** Information flow for teaching a robot a new move set. [3]

In terms of updating the robot's understanding of what the ideal move set is, Calinon *et al.* have presented two incremental teaching methods: the direct update method, and the generative method. The direct update method separates the original and corrected performances of the move sets, and evaluates their differences. The generative method, the robot generates an entirely new updated set of directions for the move set, by combining the original and corrected performances. Both methods were successful in reproducing the human teacher's movements.

## VI. Robotic Systems and Interaction Scenarios

Our preliminary studies have focused on two robotic systems, as pictured in Figure 1: Romotive and Robotis Mini. Our Romotive ("Romo"), a rover-style robot utilizing an iPhone, has been modified from its original version. We have created a network-based software architecture, and a custom character that is more appealing to children with autism than the original character. Romo emotes through facial expressions. The second robot model, Robotis Mini "Darwin Mini", presents emotion through body language - posture, gestures and movements.

For our preliminary tests, there were a total of four modes. The child would interact with each of the two robots separately in a demonstration-based behavioral game. For each robot, there would be two rounds: one with a musical accompaniment, and one without. The child's movements during the behavioral game were tracked and recorded using the Microsoft Kinect skeletal tracking system. The child's movements were then analyzed using a combination of LMA and kinematic features based on tracking of the skeletal joint angles. This analysis was used to determine the child's level of engagement with the robots, with and without a musical accompaniment in the background as they played the behavioral game.

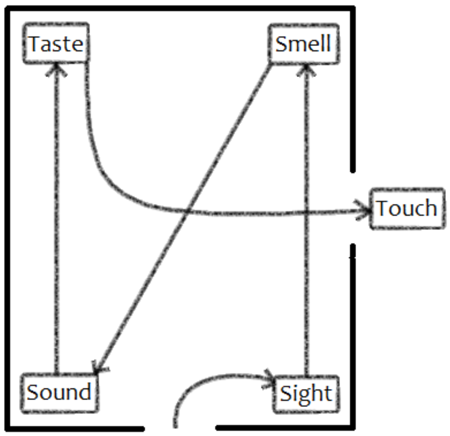


**Figure 2.** A map of the five senses behavioral game "maze".

For the musical accompaniment, as the situations and emotions of the robot changed, the music would correspondingly change to reflect the new emotion. The music chosen is based on an opensource list from YouTube, the free Audio Library. The "emotion" of each instrumental song was sorted based on the original YouTube "mood" category it came from.

For the preliminary study behavioral game, each robot explored a maze containing five stations. At each station was an object that stimulated one of the five senses, such as a light, a flower, and a soft blanket. Depending on the object, the robot gave a positive, neutral, or negative reaction, and encouraged the child to interact with the object as well. This behavioral game is meant to introduce the children to interacting with a robotic framework, and to allow the child to develop empathy for the robots as well. The contents of the behavioral game also show the child how to interpret emotional cues, as well as how to avoid sensory overload when dealing with unfamiliar objects.

## VII. Results and Analysis

Our motion analysis yields promising preliminary results. Three participants engaged in the behavioral game. Data was collected for each. Graphical representations of some of our findings can be seen in figure 3. For the comparison of motion units (MUs) versus time without music (figure 3, top left) and with music (top right), both figures have a y-axis measuring MUs, from 0 to 900, and an x-axis measuring time, from 0 to 450ms. For the comparison of absolute weight term versus time, without music (bottom left) and with (bottom right), both figures have a y-axis measuring weight term, from 0 to .025, and an x-axis measuring time from 0 to 450ms.

In sessions with musical accompaniment, there was more constant increase in MUs than in sessions with no accompaniment. As seen below in figure 3, the session with no music (top left) has choppier, "start-stop" increase in MUs. The session with music (top right) has a much smoother increase. We also found that music-accompanied trials have an increased frequency of participant movement (also seen in figure 3) and increased average participant speed.

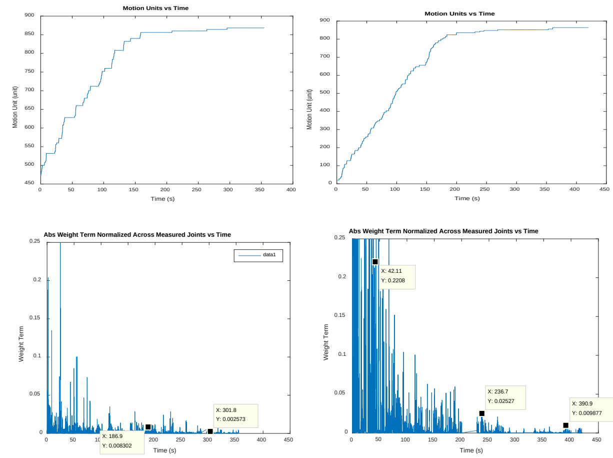


**Figure 3.** Results from the kinematic analysis of our preliminary study (Participant #2), focusing on a comparison on motion units (top) and change in weight term (bottom).

We believe that the results of this preliminary study indicate that incorporating music into our robotic framework interaction can increase the child's level of engagement. If we assume that "start-stop" motion is indicative of nervousness, then the more constant increase of MUs for with-music in comparison to without-music suggest that incorporating music reduces nervousness during the game sessions. Alternatively, a more constant flow of motion could indicate an increased level of engagement, rather than a reduction of fear. Increased patient frequency of movement and increased speed also suggest that musical accompaniment encourages an increase both in calmness and interaction.

## VIII. On-going and Future work

As we continue to improve the format of our preliminary study, we are also beginning planning of our next phase of research. As the use of both robots and music improve the level of engagement of children with autism during therapy sessions, we aim to take this a level further – from just having music in the background, to having music be a central part of the therapy. We aim to focus on the combination of robotic interactive play and dance with music in order to enhance the child's level of engagement.

In this next phase, we will be introducing a new robot – Robotis-OP2 ("OP2"). OP2 has a greater range of motion, increased vocal capabilities, and higher computing power than our previous two robots. OP2 also has visual tracking capability. Focusing on this higher computing power, our first task will be to program OP2 to mimic the child's movements. This will be done using motion learning. The more OP2 interacts with the child, the better it will be able to copy their movements. Additionally, using Laban Motion Analysis, OP2 will more accurately be able to interpret the child's movement for emotions, and will be able to emotionally react accordingly.

We plan to program OP2 with motion learning functionalities by employing a method called Gaussian Mixture Regression, or GMR. GMR has been used in both 2D and 3D models to create the best path of movement using virtual spring-dampers and multiple frames of reference [2].

## IX. CONCLUSIONS

Our research aims to provide children on the autism spectrum with affordable, effective improvement to their social and communication skills by means of robotic therapy. This robotic therapy could eventually be used in the home, in supplement to school and outside-therapy services.

Our preliminary studies show promising results. Our trial participants showed increased engagement in our behavioral game when presented with a musical accompaniment. In our current work, we are incorporating dance and interactive play into our studies, teaching our robot to mimic movement through motion learning. We plan to tap into the increased engagement levels shown to occur through the use of imitation and music in behavioral therapy sessions.

## ACKNOWLEDGMENTS

This research is supported by the National Institutes of Health (NIH) under grant # 5 R01 HD082914-04 through the National Robotics Initiative.

## REFERENCES


[1] http://www.autism-society.org/dsm-iv-diagnostic-classifications/

[2] Calinon, Sylvain, et al. "Statistical dynamical systems for skills acquisition in humanoids." *Humanoid Robots (Humanoids), 2012 12th IEEE-RAS International Conference on*. IEEE, 2012.

[3] Calinon, Sylvain, and Aude Billard. "Incremental learning of gestures by imitation in a humanoid robot." *Proceedings of the ACM/IEEE international conference on Human-robot interaction*. ACM, 2007.

[4] Dawson, Geraldine, and Alexandra Adams. "Imitation and social responsiveness in autistic children." *Journal of abnormal child psychology*12.2 (1984): 209-226.

[5] Duquette, Audrey, François Michaud, and Henri Mercier. "Exploring the use of a mobile robot as an imitation agent with children with low-functioning autism." *Autonomous Robots* 24.2 (2008): 147-157.

[6] Field, Tiffany, et al. "Children with autism display more social behaviors after repeated imitation sessions." *Autism* 5.3 (2001): 317-323.

[7] Kim, Jinah, Tony Wigram, and Christian Gold. "Emotional, motivational and interpersonal responsiveness of children with autism in improvisational music therapy." *Autism* 13.4 (2009): 389-409.

[8] Masuda, Megumi, Shohei Kato, and Hidenori Itoh. "Emotion Detection from Body Motion of Human Form Robot Based on Laban Movement Analysis." Principles of Practice in Multi-Agent Systems Lecture Notes in Computer Science (2009): 322-34. Web.

[9] Morita, Junya, Yukari Nagai, and Tomoyuki Moritsu. "Relations between Body Motion and Emotion: Analysis based on Laban Movement Analysis."

[10] Ricks, Daniel J., and Mark B. Colton. "Trends and Considerations in Robot-assisted Autism Therapy." *2010 IEEE International Conference on Robotics and Automation* (2010): n. pag. Web.

[11] Robins, Ben, et al. "Robotic assistants in therapy and education of children with autism: can a small humanoid robot help encourage social interaction skills?." *Universal Access in the Information Society* 4.2 (2005): 105-120.

[12] Tiegerman, Ellenmorris, and Louis Primavera. "Object manipulation: An interactional strategy with autistic children." *Journal of autism and developmental disorders* 11.4 (1982): 427-438.