BACKGROUND

Autism spectrum disorders (ASD) and autism affect many people in the world. According to the U.S. Centers for Disease Control and Prevention, 1 in 110 American children are on the autism spectrum (Autism Speaks, 2012). Astonishingly, more children this year will be diagnosed with autism and ASD than with cancer, diabetes, and pediatric AIDs combined. More than 3 million individuals in the U.S. and 10 million individuals world-wide are diagnosed with ASD and autism.

ASD and autism are terms that cover a larger group of brain developmental disorders, all of which have, in varying degrees, symptoms of social interaction and communication difficulties as well as repetitive behaviors. Other symptoms include lack of emotional or social reciprocity and lack of empathy (Autism Speaks, 2012).

Our system, Kinexpressions! endeavors to teach kids with autism about one particular form of non-verbal communication: Facial Expressions.

DEFINING THE USER

Based on our first few interviews, we first started the design process in class by defining the following user profiles:

1. Young children (5-10) with mild to moderate autism or ASD
2. Occupational and speech therapists

These profiles manifested into more specific user personas:

1. Janie is seven years old with mild autism. She has poor social skills and does not enjoy being in groups. She is on the high functioning end of the autism spectrum. She excels in one particular talent: mathematics. She also has normal motor functions.
2. Dave is a speech therapist who works with elementary age students in a school for children with special needs. He also works at a therapeutic play center on the weekends. He has spent the last 10 years in special education. Dave is familiar with different assistive technology, but is not a fan of the iPad since he prefers more hands-on, pair activities that help encourage his students to work collaboratively.

INITIAL DESIGN PROCESSES

During the brainstorming step, we worked from both a top-down perspective and a bottom-up perspective. For our first couple of brainstorm sessions, we decided to focus top-down by thinking mostly about tangible user interfaces. We thought such technologies would increase collaborative play between children and we each had a personal interest in exploring the space.

Figure 1.0 Brainstorm session

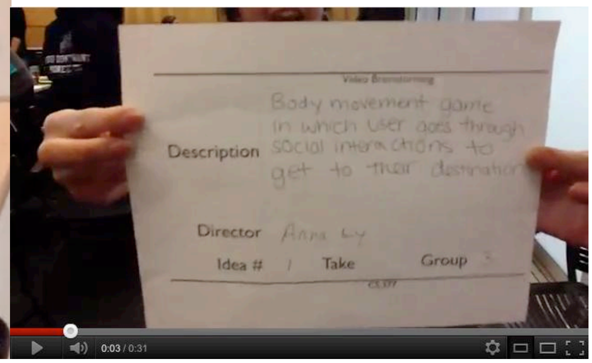
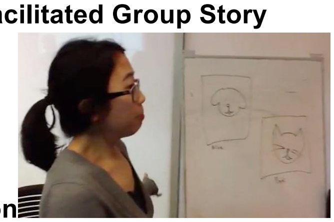
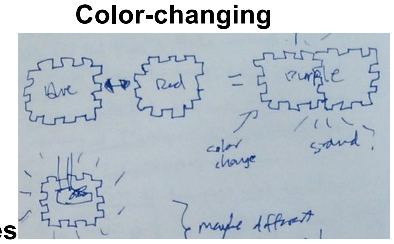
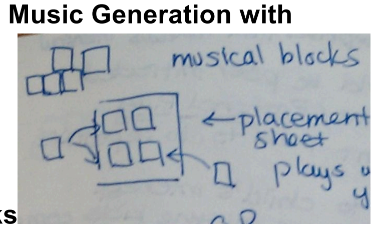


Figure 2.0 Video Prototyping

Next, we took our favorite ideas and turned them into video prototypes. We wrote down brief interaction points for the different design concepts and created quick 30 second or less video sketches.

The four video prototype concepts born from the brainstorms were:

Facilitated Group Story Creation  *Figure 3.0a* This application facilitates a creation of a story amongst a group of children. A child gets a picture and must tell a story based on the picture. The game then prompts the other children to complete the story using other pictures that pop up. This encourages better social interaction and collaboration. (video)	**Color-changing Tiles**  *Figure 3.0b* This is a tangible user interface that encourages collaborative play. Children can construct items out of the tiles that change color when combined with each other (red + white = pink, etc).
Interactive Social Story Quest  *Figure 3.0c* The quest is a game that helps teach the child social norms and interactions. For example, the child must shake the wizard's hand to get the trophy.	**Music Generation with Blocks**  *Figure 3.0d* This is a tangible user interface that encourages collaborative play. The child works with other children to create a song by moving and placing interactive blocks on a board. (video)

After creating these video prototypes, we had a series of interviews with Michelle Fong at Associated Learning & Language Specialists, several therapists at Hope Technology School, as well as family and friends who have children with autism. After hearing all of the specific stories from our interviewees, we soon discovered how children with autism really struggle with expressing emotions both verbally and physically through the face and body, displaying empathy for others, and understanding social cues. The therapists also gave us additional design considerations such as be careful about feedback because children “just want to click on it to get the feedback instead of actually learn” (M. Fong, personal communication, January 27, 2012).

With this information, we switched to a bottom-up approach for our brainstorms. Instead of looking at tangible user interfaces, we focused on the key takeaways from our interviews. After brainstorming some ideas around emotions and expressions, we decided to flush out one of our design concepts. This included developing more detailed interaction points and sketching out storyboards. This culminated into the MotionEmotion (later titled Kinexpressions!) video prototype (see *Figure 4.0),* which we presented during our mid-term presentation. We then did user testing in the form of walkthroughs of the video with potential users including a director of a local center for children with developmental disorders, as well as design experts and students.

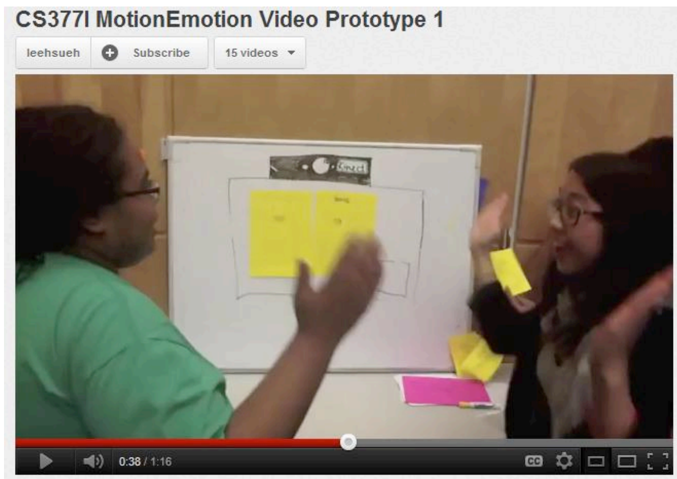


Figure 4.0 MotionEmotion Video Prototype

REDESIGN

Our project changed according to the feedback that we received. This feedback stemmed from the walkthroughs, video prototype, and tangible prototype exercises we conducted. During our first structured walkthrough, there were 3 main interaction points: Turning it on, selecting an option, and matching an emotion. In the case of the latter two, we created branching prototypes/alternative interactions to figure out how to optimize the system interactions and for each branch we generated multiple options.

We storyboarded and video-prototyped various interactions for the task of constructing a facial expression by manipulating a face canvas on the screen. The main three options for a selection technique were as follows: swiping to see each face section individually, holding your hand over a pallet of eyes or another facial feature until it appeared on the face, or dragging the eyes from a pallet to the face. We decided not to use the swiping to see each face section as we received feedback that allowing kids to see one option at a time may make it difficult for them to see the overall outcome. Additionally, there may be a face that matched that they did not see because they chose an option before seeing all of the available ones. We also decided not to use dragging the eyes to the face during the wizard of oz exercise. This is because it added an additional number of unnecessary steps and we wanted our users to mainly deal with activities that were central to our goal. While holding their hand over a selection did make the users work harder, they were building their hand motor skills which is another goal that we set out to accomplish. For these reasons, we settled on having the child hold his/her hand over the object in order to make a selection.

We also generated a surplus of ideas for matching an emotion. These included things like seeing a picture and then mimicking it, watching a short video and then selecting the emotion that the character felt, or watching a video and then building the emotion that displays how the character felt. Since we got positive feedback with these different options, we decided to integrate them together. This is unlike our procedure of cutting options until we had selected the best for “selecting an option” but it also maximizes the effect of the project by creating an effective and coherent flow.

Originally we wanted to design something that allowed interaction between multiple kids, perhaps more than one on the spectrum. However, after interviewing both parents and therapists, we decided to focus more on therapists and the activities they do with children on the spectrum. Rebecca Berry at Developmental Pathways mentioned how difficult and complex it would be to have activities involving multiple children on the spectrum. In fact, most integrated play groups only have one child on the spectrum interacting with typically developing peers.

During our many generative walkthroughs, we assessed our system for different socio-technical principles and added or improved functions in conjunction with this assessment. One particular function that we added to the project was the ability for the child to take and save pictures of themselves portraying a particular emotion. The therapist then may view these photos and see how the child progresses over a period of time. In this way, our system is co-adaptive because the therapist can potentially direct the way that the pictures can be used.

Another function that we added to our system after doing the generative walkthroughs and interviews was using colors to indicate when the right or wrong answer occurs in addition to text feedback. Students from the class mentioned that more scaffolding may be helpful for those who have more moderate than mild autism. The region which was wrong (eyes, mouth, nose) is highlighted while the incorrect answer disappears to indicate to users that they should make a different selection. At first, our prototype only had sound but we wanted the feedback to be multisensory. Thus we also used our peripheral awareness principle to highlight the correct or incorrect answer while leaving the other information on the screen such as the palette of eyes.

While highlighting and color coding allow our users to pick up information easily, another way to do this was by leaving the emotion face that the child constructed up while they were trying to mimic it. Since the child did not need to remember what the face they had constructed looked like they could focus on actually making the emotion face themselves. In this way, we can reduce their cognitive load while learning.

From these exercises we finally decided on our current product Kinexpressions!. As seen above, our feedback from both interviews and walkthroughs indicated that there was a huge need for a product that addressed emotion expression and identification for children with autism. Additionally, we found that many of our earlier ideas had already been created. We wanted to design something new or something that existed but was inadequately explored. Since emotion expression kept occurring as a key theme in the interviews we decided to go with that idea. Moreover, the successes at the Lakeside center for autism showed us that the Kinect could be used effectively for children with autism.

Although many of the techniques were useful to our design, there were some techniques that did not work. For starters, the Introspection exercise failed because we were not part of the intended audience. Moreso, because our intended audience consists of children with autism, there is a risk of applying unwarranted bias. The function interaction table was also difficult for us to do since we were trying to change behavior, rather than the user manipulating a screen.

CURRENT DESIGN

Our current design is a Kinect activity that aims to help children with autism learn empathy through emotions and facial expressions. On one level, the activity is trying to directly instruct the child about what the different emotions are and in what contexts they might arise. At a higher level, the activity is meant to facilitate social interaction between the child and a facilitator such as a therapist, parent, or an older typically-developing peer.

ACTIVITY

At its core, the activity consists of a collection of short videos and an internal library/database of emotions with corresponding facial expressions. The procedural flow of the activity can be summarized as follows:

1. The child watches a video
2. The child identifies the dominant emotion associated with one of the characters in the video
3. Upon successful identification, the child constructs a facial expression corresponding to the emotion
4. Upon successful completion the child tries to express the emotion through his/her own face

A video is selected not by showing a list of videos, but by presenting a choice of people (Figure 5.0a). The person that the child selects is the character that he/she needs to pay attention to in the video that follows (Figure 5.0b).

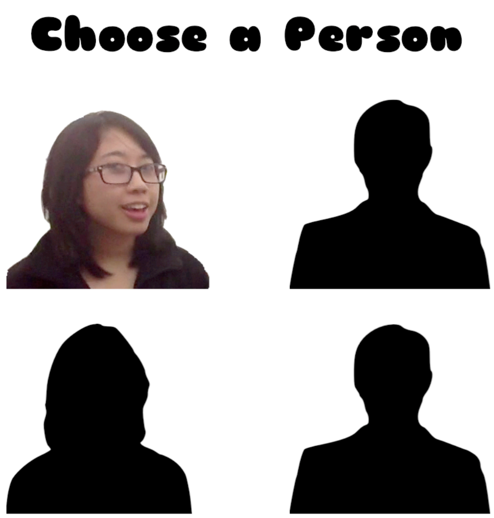


Figure 5.0a

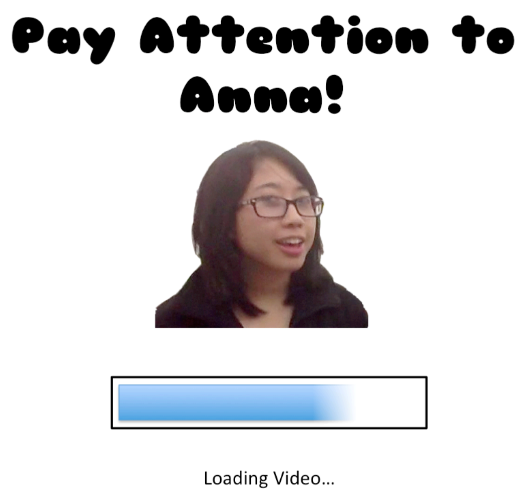


Figure 5.0b

Each video depicts real people or characters in a particular scenario that elicits a certain emotion in the "main" character, which the child needs to empathize with. Real people are used to encourage transfer of what the child is learning to real life. In more difficult videos, the child may need to make a distinction between how the main character in the video feels and how the video itself makes the child feel as an observer. In beginning videos, however, it may be helpful to have the dominant emotion felt by the character coincide with the emotion that is evoked by the video itself.

After the video is done playing, the child is presented with a multiple-choice question of how the main character feels (Figure 6.0). The child has the option to select a choice or to go back and watch the video again.

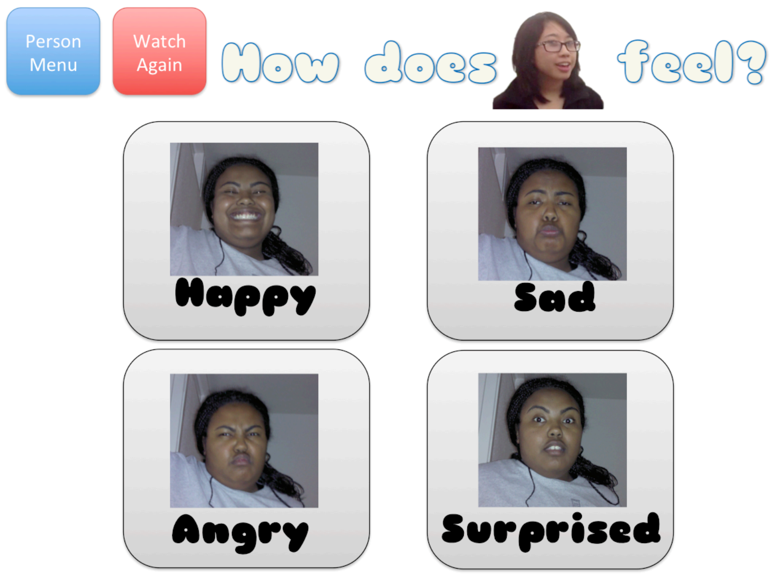


Figure 6.0

When the child successfully selects the correct emotion, he/she is directed to a mini-activity in which a facial expression must be constructed. Initially a blank face “canvas” is shown that is divided into three regions. The activity guides the child to choose between various facial features for each region of the face (e.g. for the eye region, there is a palette which consists of different eyes showing different expressions like angry eyes, surprised eyes, etc.). When the child moves onto the next region of the face, the palette changes to reflect choices for the corresponding region (palette of noses, palette of mouths). Once all three regions are “filled,” a submit button appears which gives the child the option to check whether or not the constructed face successfully matches to the given emotion (Figure 7.0).

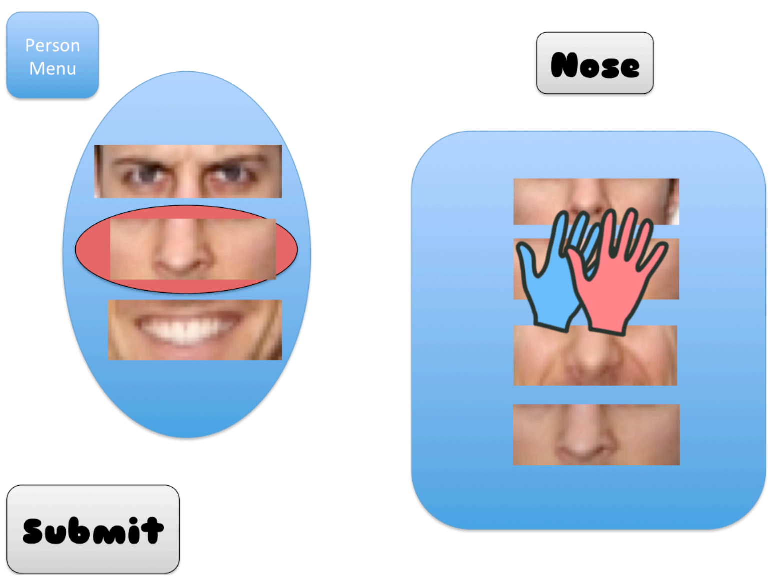


Figure 7.0

This purpose of this mini-activity is to help the child analyze and parse facial expressions into smaller and thus more manageable regions. There is a slight tension here between the value of holistic evaluation and reducing clutter – we do not want the child to feel overwhelmed with the many ways a whole face might look and thus we have them work on one region at a time, yet it is conceivably difficult to discern whether a particular feature is correct without looking at the entire face. Perhaps this tension is also something that the child must grapple with and learn.

The final portion of the activity, after the child successfully completes the face, is for the child to practice expressing the emotion using his/her own face. A text and visual prompt is shown to encourage the child to bring up the camera (Figure 8.0a). A side-by-side comparison shows the child's face real-time with the correct facial expression that was constructed in the previous activity (Figure 8.0b). The purpose of this step is to actively engage the child in becoming familiar with emotions by being able to recognize them in his/herself.

Figure 8.0a

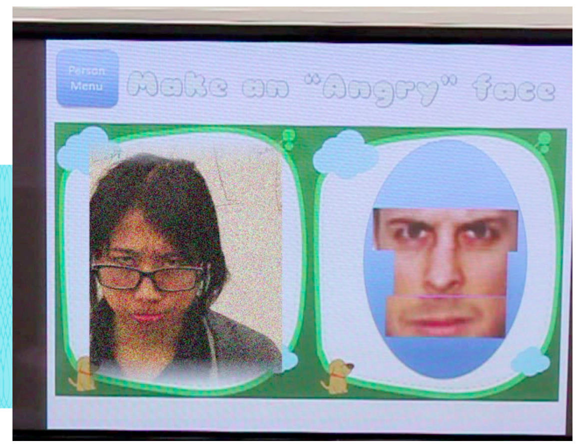


Figure 8.0b

Once the child has taken his/her picture, the photo is stored in a "library/catalog" of emotions. These photos can serve as additional data for the therapist to compare/contrast specific emotions and to track progress of a particular child over time.

The activity consists of a relatively fixed flow of sub-activities. While there is not much flexibility to explore, this kind of rigid procedure may actually work better for our target users, since individuals with autism and many other mental disorders find routines and predictable structure and events soothing. Unpredictable and unexpected behaviors and responses can provoke anxiety and outburst.

The flow of activities also aims to help internalize emotions through a sequence of exercises that gradually increase in precision and explicitly. The video scenarios are meant to establish connection with the child's existing experiences, and specific emotions are extracted and formed from them so that the child can explicitly express them outwardly.

USER INTERACTION

The primary form of interaction in the current design relies heavily on hand movement to control a virtual cursor on the screen. This interaction is used to select and confirm choices, which is inherent in all of the sub-activities. It also helps refine the child's hand-eye coordination and motor skills.

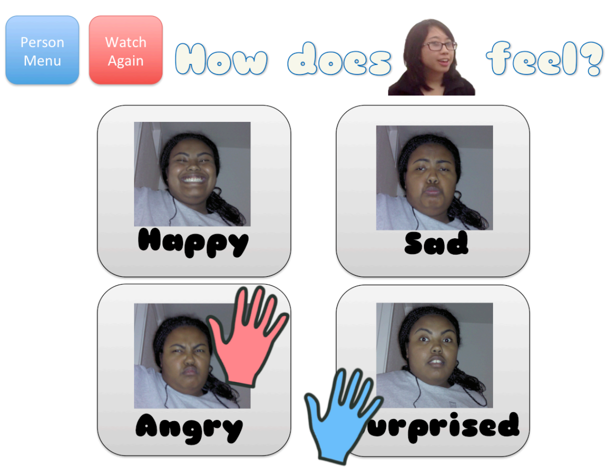


Figure 9.0

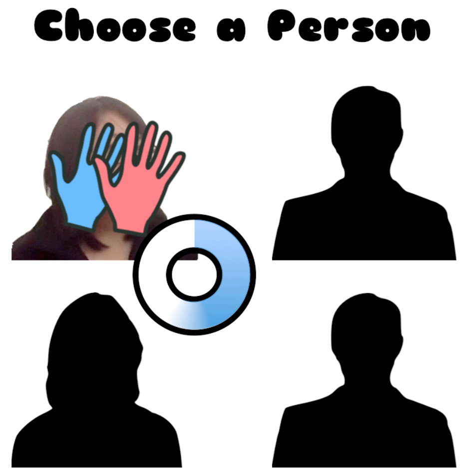


Figure 10.0

We drew inspiration from the Xbox Kinect interaction, which uses a hovering mechanism to allow the user to make a selection. Instead of “clicking” on a menu item, for example, the user must move the hand cursor to hover over it. A small ring appears under the hand and gradually becomes filled up, indicating the time that the hand has been hovering over the item (Figure 10.0). Once the ring is filled, the selection is confirmed. We decided to utilize this interaction because not only does it serve as a technical substitute for clicking, but because it also prevents haphazard and aimless selection, which often occurs in tablet apps. With the Kinect, a child must exercise more deliberation and motor control in making each decision, which we speculate will encourage more thinking.

As mentioned earlier, the user interaction involved in the execution of the activity procedure is intended to facilitate social interaction and collaboration between the child with autism and a facilitator, most likely a therapist. In order to accomplish this, we decided to make every selection/decision in the activity be a paired decision. What this means is that the child and the facilitator both have virtual cursors on the screen, and in order for a selection to be confirmed, *both* cursors must be hovering over the item for the necessary duration.

This interaction is introduced and taught at the start of the activity. The task is to have both the child and the therapist’s hand cursor match to an outline of a hand on the screen (Figure 11.0). The purpose of this introductory interaction is two-fold:

1. Familiarize the child with the interaction, since it is used throughout the activity
2. Enforce that the activity can only be completed if two people are present

Figure 11.0

The rationale behind enforcing this “co-selection” is to encourage the child and the therapist to self-negotiate and collaborate. If the child wants to choose a particular item, he/she must communicate the intention to the therapist, possibly expressing why he/she wants to make that specific choice. Similarly, if the child is inactive, the therapist can propose the child to try making certain choices and having the child move his/her cursor to match the therapist’s. With this enforced, the child can practice verbal communication and listening at almost every interaction.

Design Evaluation and Socio-technical Principles

Another run-through of our project using socio-technical principles revealed additional opportunities for redesign. One possible breakdown that we had was the power going out. If the system was able to automatically transfer photos by either email or usb we could show the pictures to the child and have them replicate the face they see. Additionally, the therapist could make a face for the child to copy. This graceful possible handling of a breakdown is good because it shows positive ways a user could use our system in unforeseen circumstances (Situated Action Principle).

Other examples of socio-technical principles affecting our design is co-adaptation. For the future, we want to expand the system beyond training and into monitoring the progress of the child. For example, after the kids save their pictures, the therapist may use it to monitor how well the child did in comparison to pictures that were taken several weeks ago. This also might contribute to distributed cognition, in which the therapist does not have to remember details of the child’s face and emotions since they will be saved and recorded in the system.

Conclusion

Through the process of design and redesign, we were able to come up with an idea that could potentially help children with autism learn important social skills. We were able to talk with therapists about what they work on with these kids and share ideas with them. With their advice and feedback, we came up with Kinexpressions!, a Kinect-based activity that helps teach empathy through emotions and facial expressions.

Going forward, we would like to further refine our idea and begin development and implementation using actual hardware and software. Our idea will continue to evolve, since in its current state there is nothing in the activity that leverages perhaps the most powerful affordance of the Kinect: skeletal tracking. This affordance opens a huge potential for incorporating body language into the activity, which is another form of non-verbal communication that children with autism need to become familiar with.

We are also in the process of working with developers and researchers at Microsoft involved with the Kinect, and hope to procure a Kinect sensor for development purposes. There are many initial technical decisions that must be made, such as which Kinect sensor to use (Xbox or Windows) and which SDK to use (official Microsoft SDK or open source alternatives), as well as the complexities of implementation itself.

We also hope to hold a participatory design workshop with therapists and parents in the community to generate more ideas and insights from the people who spend so much time with our target user population. The Kinect is already beginning to be incorporated in therapy sessions using off-the-shelf games and software. Hopefully by the end of the year, we can jumpstart the development of dedicated and specialized activities using the Kinect for children with autism. Through everyone’s efforts and expertise, we hope to create a solution that is effective and can be used by therapy centers and clinics that serve the autism community.