Designing Gesture-based Applications for Individuals with Developmental Disabilities: Guidelines from User Studies in India

S. SHARMA, University of Tampere
B. VARKEY & K. ACHARY, Tamana
J. HAKULINEN & M. TURUNEN, University of Tampere
T. HEIMONEN, University of Wisconsin–Stevens Point
S. SRIVASTAVA, Conduent Labs India
N. RAJPUT, InfoEdge India Ltd.

Gesture interaction provides a multitude of benefits to individuals with developmental disabilities; from enhancing social, motor and cognitive skills, to providing a safe and controlled environment for simulating real world scenarios. As gesture-based applications gain ground in the special education domain, we study their potential in the Indian context. Together with Tamana, an NGO in New Delhi, we have been conducting a series of exploratory user studies since October 2013. This includes the design and evaluation of three gesture-based applications to impart social and life skills to individuals with developmental disabilities. The *Kirana* application employs socially appropriate gestures to teach the life skill of buying day to day items from a local Indian grocery. *Balloons* promotes joint attention skills through collaborative interaction. HOPE improves motor coordination and social and cognitive skills, with increasing levels of difficulty. Based on studies with these applications, this article presents guidelines for designing gesture-based applications for individuals with developmental disabilities. The guidelines focus on (a) designing applications that cater to a larger group of individuals to encourage collaboration and inclusion, for instance, providing easy and controllable transitions between different task levels, and balancing interaction and content complexity; (b) addressing the challenges in conducting research in this domain, with respect to ethical and procedural decisions; (c) designing for technology acceptance within the Indian context, for example, by following a collaborative and stakeholder inclusive approach, and addressing apprehensions towards technology adoption. These guidelines aim to benefit other practitioners working in this domain, and especially in the educational technology context of India.

CCS Concepts • **Human-centered computing~Gestural input** • **Human-centered Computing~Empirical studies in accessibility** • *Human-centered computing~Empirical studies in interaction design*

Additional Key Words and Phrases: Gesture Interaction Design; Design for all; HCI; Individuals with developmental disabilities.

ACM Reference Format:

Sumita Sharma, Blessin Varkey, Krishnaveni Achary, Jaakko Hakulinen, Markku Turunen, Tomi Heimonen, Saurabh Srivastava, and Nitendra Rajput. 2017. Designing Gesture-based Applications for Individuals with Developmental Disabilities: Guidelines from Studies in India. ACM Trans. Access. Comput. X, X, Article X (Dec 2017), X pages.

DOI: 10.1145/1234

1. INTRODUCTION

Interactive technology offers several advantages and possibilities for individuals with developmental disabilities, including (a) a controllable input stimuli ([5],[16],[28],[29]), (b) a multisensory and safer learning environment ([2]-[4],[32]), (c) customization for individualized learning goals ([2],[3],[7],[16],[17]), (d) structured, predicable and consistent learning environments ([2],[3],[16],[28]), (e) controlled modifications or difficulty levels ([2],[3],[7],[16],[17]), (f) assistance in generalization between scenarios ([2],[3],[27]-[29]), and (g) self-paced repetition of learning activities ([16],[17],[28],[29],[32]). Moreover, applications can simulate real world scenarios in a safe and controlled environment, and the learning is potentially translatable from the virtual to real world, facilitating self-efficacy ([8],[23],[30],[34],[36]).

This article presents guidelines for designing and developing gesture-based applications for individuals with developmental disabilities. We consider individuals with developmental disabilities as individuals with cognitive challenges, including Down syndrome, ADHD, and high-functioning autism. Specifically, we address the potential of designing for users in India, where there are several challenges towards adoption of novel technologies – from expense and catering to a limited group of people, to difficulties with integration and maintenance. Furthermore, individuals with developmental disabilities have limited access to these technologies, and therefore the digital divide – technical and economic barriers towards technology access – is more pronounced. Thus, it is important to substantiate the potential of such technologies and build a stronger case for their adoption. Collaborating with different stakeholders, for example, educators, therapists and caregivers, and following a user-centered design approach, can potentially increase technology acceptance. Moreover, by taking into account the cultural implications of the environment, socially acceptable interactions should be designed [4].

The guidelines for addressing the above issues are based on our work on gesture-based applications with a nonprofit organization, called Tamana, at their Nai Disha School since October 2013. Tamana is dedicated to providing young adults with developmental disabilities physical, emotional, and functional independence by imparting vocational training and skill development. Our collaboration with Tamana has produced several educational applications. In this article, we present the exploratory user studies of three applications: *Kirana* teaches life skills by breaking down complex tasks that require social, mathematical, and decision-making skills; *Balloons* promotes joint attention skills through social collaboration; *HOPE* improves motor coordination, social and cognitive skills. This article is an extension of our prior work on *Kirana*, which was presented at the ASSETS 2016 conference [34]. In this article, we present results from the two aforementioned additional studies and summarize the findings from all the studies as set of guidelines. The main contributions in this article are:

Guidelines for gesture-based applications: We combine our results and experiences from user studies with *Kirana*, *Balloons*, and HOPE to present guidelines for designing gesture-based applications for individuals with developmental disabilities, with a focus on three aspects: designing the application, designing for research in this domain, and designing for the Indian context.

Redefining inclusivity in interactive technologies: The *Kirana* application ([34],[35]) and our previous study with *Balloons* [33] focused on a small group of individuals. With HOPE, we designed for *inclusion-within*, that is, an application that not only caters to a larger group of individuals, but also encourages collaboration. This is motivated by Nai Disha School's conscious efforts towards inclusion of students, within their schools and mainstream society. Therefore, user studies with *Balloons* and HOPE presented in this article include a larger spectrum of individuals with development disabilities.

In this article, first an overview of the studies is presented (section 1.1) followed by related work in this domain (section 2). Then, the applications and their evaluations are described starting with *Kirana* (section 3), *Balloons* (section 4), and HOPE (section 5). Finally, combining the findings and experiences from all the evaluations, a collated set of guidelines for practitioners in this field is provided (section 6).

1.1 Research Overview

The findings presented in this article are derived from the design journeys, exploratory user studies, and the experiences of the school specialists involved with the work of three applications:

Kirana: The *Kirana* application ([34],[35]) was designed to teach the life skill of buying items from a local Indian grocery store by breaking down the tasks of social, mathematical and decision-making skills, into smaller subtask. It employs pointing with dwell time for selection of items, and eliminates the need of object manipulation gestures by animating item movements between the different subtasks.

Balloons: The *Balloons* application was designed for children with medium-low functioning autism [33], to promote social interaction via shared experiences. It employs pointing with dwell time for selection, and two users can interact with the application simultaneously. In this work, *Balloons* is used to introduce the participants to gesture-based interaction and the concept of colors.

Figure 1: (left) A participant using *Balloons*, and (right) a moderator assisting a participant with the pointing gesture (pictures adapted from [33])

HOPE: HOPE focuses on improving motor coordination, social and cognitive skills through a series of spatial reasoning tasks, employing a gesture sequence of grab, drag, and drop. In this work, the tasks are focused on matching colors – building on the concepts in *Balloons*, while introducing a complex sequence of gestures for selection and object manipulation.

Nai Disha School's psychologists evaluate each student annually, or biyearly, and maintain a record of their developmental progress, which includes IQ and SQ (social quotient) scores. IQ is measured using an Indian adaptation of Wechsler's intelligence scale for children [41], called Malin's Intelligence test [21]. SQ is measured using the Vineland social maturity scale, which is a psychometric assessment designed to measure the social competence of an individual. It measures eight categories of behavior: "self-help general, self-help eating, self-help dressing, locomotion, occupation, communication, self-direction, and socialization" [10]. Our work with *Kirana* focused on a narrower group of individuals towards the severe end of the developmental disability spectrum (IQ < 50, SQ < 60). In studies with *Balloons* and HOPE, participants were divided into two groups based on their IQ and SQ, whereby group one participants are on the severe end of the developmental disability spectrum (IQ and SQ < 50) and group two participants are towards the moderate-mild end of the spectrum (IQ and SQ > 55). A total of 32 participants are a part of the work presented in this article, and *Kirana* participants are not a part of the *Balloons* and HOPE evaluations. Summary of participant information for each study is presented in the table below:

Table 1: Participant information for all the studies

Participant Information	Kirana	*Balloons* and HOPE	
		Group 1	Group 2
Number of Participants	18	9	5
Age (in years)	M=26, SD=5.4	M=29, SD=7	M=22, SD=4
Gender (Male, Female)	M=14, F=4	F=9	M=4, F=1
IQ	M=46, SD=11	M=42, SD=11	M=63, SD=20
SQ	M=58, SD=16	M=41, SD=10	M=57, SD=15

The guidelines presented in this article are therefore based on several different variations of gesture design, application content and participant profile, providing a well-rounded perspective towards designing gesture-based applications for individuals with varying levels of developmental disabilities.

2. RELATED WORK

In the following, we discuss gesture-based interaction for individuals with developmental disabilities and its benefits vis-à-vis the theory of embodied cognition, and present an overview of assistive technologies for Indian children.

2.1 Gesture interaction for Individuals with Developmental Disabilities

Employing gesture-based, embodied interaction for social, therapeutic and educational applications for individuals with developmental disabilities has gained momentum in recent years. Such embodied learning paradigms are based on the theory of *embodied cognition*, whereby cognition is situated within the environment and learning also occurs through bodily interaction with the environment ([19],[38]). The embodied learning paradigm is extensively studied in neuroscience and cognitive sciences ([9],[14],[18],[37]), and is adopted by research in educational technologies within the human-computer interaction (HCI) domain [11].

There is considerable research on the benefits of gesture-based interaction for individuals with developmental disabilities. For instance, the Lakeside Center for Autism[1] and Kinems.com[2], have developed applications to improve social, motor and cognitive skills for children with autism. With affordable motion tracking technologies, such as the Microsoft Kinect, several studies have also incorporated Kinect-based applications to match visual facial expressions [7], improve hand-eye coordination, attention and focus [1], detect repetitive behavior or tantrums [13], and for cognitive rehabilitation and exercising [36].

MEDIATE ([25],[26]) is an interactive multisensory environment that uses visual, aural and vibro-tactile stimuli in real time to motivate and engage children with autism. Children are allowed to explore the spaces at their own pace and there is no time limit to the interaction. Results from user studies in multiple European cities have shown high acceptance of its interactive space, as it allows creative self-expression using interactive multimodal applications, creating therapeutic and learning experiences. Likewise, SensoryPaint, a multimodal application that incorporates tangible interaction and whole–body interaction, provides therapeutic interventions to encourage social interaction among children with autism [31].

Furthermore, studies by Bartoli et al. ([2],[3]) showcase the benefits of embodied learning via gesture-based interaction for children with autism. In line with Bartoli's findings, we also purport that learning from gesture-based interaction is applicable to real world scenarios where selection via pointing and moving of physical objects simulates interactions in everyday life. Bartoli et al. [3] also provide design guidelines for gesture-based playful interactions based on their ongoing research with children with autism in Italy. They divide their guidelines as general or goal specific focusing on motor, cognitive, and social skills which are important considerations for children classified with medium-low functioning autism. Our work expands Bartoli's work by extending it to the IDD context (including autism, Down syndrome, and ADHD), and to usage in the context of developing countries.

[1] http://lakesideautism.com/tag/kinect/
[2] http://www.kinems.com

2.2 Assistive technologies in India

In India, there is a culturally misguided attitude towards individuals with disabilities. For instance, in a study of teachers' attitude towards children with disabilities in schools in Mumbai, teachers were more positive and welcoming towards inclusive education if they had prior acquaintance with a person with a disability [24]. Moreover, there is "worrisome vicious cycle of low schooling attainment and subsequent poverty among people with disabilities in developing countries", as reported by World Bank a based on a survey conducted in fourteen developing countries [12]. There is limited research on the benefits of assistive technology in education, even though there are several schools in India that employ technology for children with special needs. One study proposed developing assistive communication technologies for individuals with autism or dyslexia in India [32]. Another, called Jollymate, is a digital notepad for children with dyslexia that emulates a "phonetics system of teaching letter sounds and letter formation" [17]. Tamana has incorporated TOBY Playpad [39], and iPad based learning application, into their educational intervention for children with autism [40].

There are also several challenges in adopting new education technology, especially for IDD in the developing world. For instance, there can be limited access to resources within the environment, like infrastructure and electricity. There is also a growing digital divide where entire communities do not have access to potentially beneficial technologies. In schools, inclusivity and integration between individuals with disabilities and typically developed is low. Moreover, technology is considered too expensive, especially for individual use. There also exist strong cultural barriers for individuals with disabilities, which increases digital exclusion even within technology-capable communities. In fact, studies have shown that cultural, societal and socio-economic factors largely affect "the experience of autism" ([4],[15]). For instance, economically stable and educated parents might provide a mobile phone to a typically developed child but not to a child with autism. However, costs issues are mitigated when emerging technologies become affordable, one device is customized to cater to a larger number of users, and resources are shared. With these changing attitudes, we believe our work to be timely, and well situated within the Indian context. Next, we present the design and evaluation of the *Kirana*, *Balloons* and HOPE applications.

3. *KIRANA* APPLICATION

To inform the design of *Kirana*, we conducted several discussions with different stakeholders at Nai Disha School's annual fair. The fair provided an informal event for researchers to interact with various stakeholders – school staff, educators, students, and their parents – to gain insights into the stakeholders' expectation and challenges towards acceptance of technology. For the complete description of the annual fair's exploratory design work, please refer to the ASSETS paper [34]. From these discussion, we identified one potential application area for gesture-based interaction: simulating real world scenarios. In this case, the focus was on teaching how to purchase day to day items from a local grocery store in Delhi, usually called *Kirana* stores (shown in Figure 2). At Nai Disha School, students who are able to communicate with strangers, have monthly visits to a local store nearby to practice buying day-to-day items, such as toothpaste. However, sometimes a student is socially mistreated by strangers or the shopkeeper. Moreover, not all students are comfortable with the activity and the school specialists are unable to extend these visit to those students. Our research goal was *to validate the potential of gesture-based interaction in the translation of learning of a life skill from an application, practiced in a safe environment, to real world scenarios*.

3.1 Design

In designing *Kirana* [35], we broke down the task of purchasing items into several smaller subtasks: selecting the item to buy (decision making), asking for it from the shopkeeper (social interaction), looking up the price for the item, knowing if the item can be bought with the available cash (arithmetic), handing over the cash (social interaction), calculating the balance (arithmetic), and taking the balance and items (social interaction). In India, all products have an *M.R.P.* (maximum retail price), which is required by law to be printed on the packing, and is usually the price charged by stores in New Delhi. Therefore, the concept of *M.R.P.* needed to be incorporated in the application to make it contextually and culturally relevant. *Kirana's* interface simulates a typical store layout (Figure 2), where customers stand outside the store and point to items they want to purchase.

Figure 2: A *Kirana* store in India

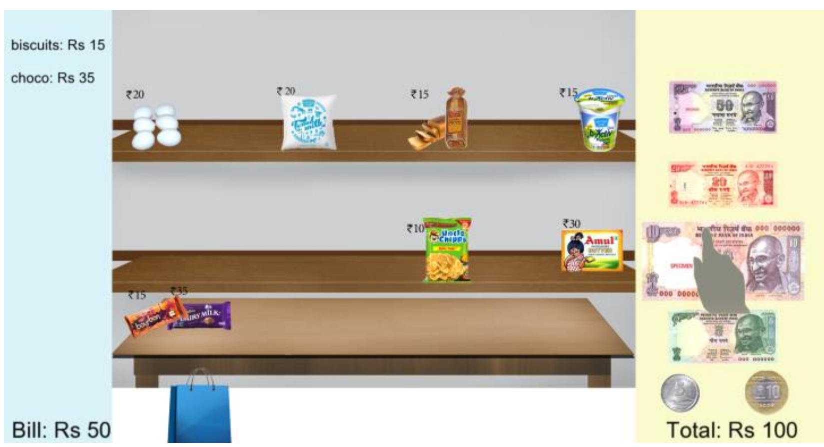

Figure 3: *Kirana*- buying items on the table by paying using a 10 rupees note

The application screen has two shelves behind a table counter-top containing food items that can be bought, as shown in Figure 3. The right side of the screen has the available money with the total amount displayed at the bottom. The left side shows the billing of items as they are selected. The bill format is in line with *Kirana* shopkeepers who provide a written bill for all items purchased. The application has an onscreen hand cursor for interaction, and the gesture for item selection is pointing with a dwell time of 1.5 seconds. When an item is selected, its price is automatically added to the bill on the left. Once there is at least one item to pay for, the participant can select any denomination of money using their right hand. This two-handed selection mechanism was added to encourage increased bodily movement and isolate the item selection from the money selection.

To simplify the interaction, given the complexity of the task, selected items automatically slide to the desired onscreen location, eliminating the need of an explicit drag and drop. The animations of items and money are slow and sequential: there is only one item or money movement, from the shelves or the wallet, to the table at a time. Once these animations are finished, the balance is returned from the table to the wallet, and the bill and total amounts are updated.

Each item and money is accompanied by its spoken name and each transaction process has audio feedback. For example, when the balance is returned, a female voice says *here is your balance*. The application also caters to the various endings: running out of money, or not enough money left to buy an item. For these scenarios, the female voice informs, for example, *you have no money left*. This is followed by a textual *well done* and an audio feedback to indicate the end of the session.

3.2 Application Setup

The application is based on an in-house framework, utilizing the Microsoft Kinect sensor for gesture recognition. The Kinect tracks a user's hand with respect to their body within a three dimensional area, called the *physical interaction zone* [22] or PhIZ. PhIZ is relative to the user, and not the display, making it more flexible to point at larger distances (Figure 4, left). The system responds to the user closest to the Kinect, in a 1-meter by 1-meter area, 1.5 meters away from the Kinect (Figure 4, right). We call this the *active area*, where users can interact with the application using gestures. An onscreen hand cursor helps guide the gesture interaction. A session starts with a verbal welcome message when a user enters the active area, followed by a short music clip and the screen fades out from black to the game screen. While the user is in the active area, the system responds to their gestures as defined by the game. When the user leaves the active area, the game screen goes back to black. A large display was used with a laptop running the application.

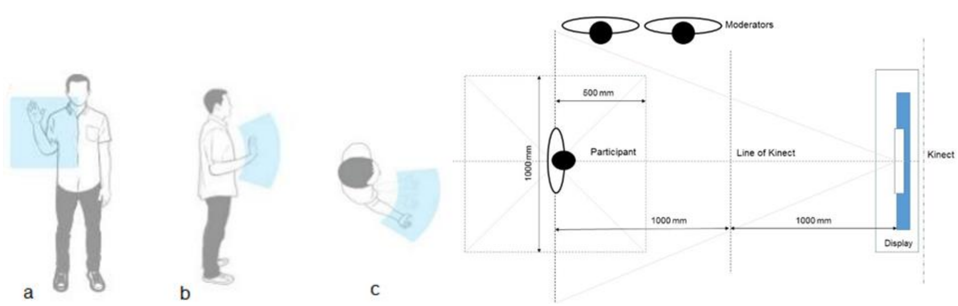

Figure 4: (left) Physical interaction zone area, in blue, as seen from the (a) front (b) side (c) bird's-eye view (adapted from [22]) and (right) setup showing the interaction space

3.3 Participants

18 individuals with developmental disabilities participated in our evaluations (5 individuals with Down syndrome, 3 individuals with autism, 2 individuals with cerebral palsy and an unspecified form of intellectual disability, and 8 individuals with other unspecified forms of intellectual disabilities). There were 14 males and four females aged between 16 to 39 years (*M*=26, *SD*=5.4) with a mean IQ of 46 (*SD*=11) and SQ of 58 (*SD*=16). The participants were recruited from the eighth grade of the school and they (a) understood the concept of left and right, (b) communicated verbally

with the moderators, (c) understood instructions given to them, (d) had an awareness of self and body, and (e) were previously unable to shop independently.

3.4 Procedure

The evaluation consisted of four phases, starting in April 2014, and were moderated by two special educators. This reduced any anxiety and complex social dynamics the participant might experience with the unfamiliar researchers. The educators were also better equipped to understand the nuances and implications of the participants' behavior and reactions to the application or its tasks. They were also responsible for and involved in all technical educational interventions within the school, and had participated in the design and development of the application.

In the phases I and III, manual mathematical tests were conducted for pre and post evaluation. These tests aimed to ascertain the mathematical ability (addition and subtraction) of the participant and included single-digit subtractions, 2-4 digit additions, and single digit multiplications. The tests were evaluated by the educators and a score out of 10 was provided for each participant. In order to understand the learning offered by the application, participants were not taught mathematical concepts in other class sessions during the evaluation.

In phase II, *Kirana* was installed in a classroom and sessions were conducted for three weeks, with every participant interacting once per week. The participant was given a grocery list by the educator and a fixed amount of money from which they had to use to purchase the items. Each participant was asked to complete the following tasks per session:

- You need to select items for yourself for breakfast from *Kirana* using your grocery list. For example, bread, chips, milk and biscuits.
- You have a budget of 100 Rupees to pay for the items you selected. Also check the balance returned by shopkeeper.

In phase IV, participants visited an actual *Kirana* store near the school and were asked to buy several items. One of the educators observed participant behavior with respect to the three sub-tasks: decision making (choosing an item from the list to buy), social interaction (talking to the shopkeeper and asking for an item), and mathematical ability (knowing how much to pay and balance to expect).

Data Gathered: Data from the sessions included automated system logs with task times, items bought, and monetary transaction details, moderator observations and a behavioral analysis to record positive and negative emotional behaviors or signals.

3.5 Results

Phase I and III

Figure 5 shows the mathematical scores from phases I and III for each of the participants. We observed an average improvement of 8.3% in mathematical test scores from phase I to phase III. However, four participants showed a decrease in performance. A score of 3 indicates knowledge of numbers from 1-10 and familiarity with money mainly notes, while a score of 6 indicates knowledge of numbers from 1-100, familiarity with notes and coins, and simple two-digit additions and single-digit subtractions. A score of 10 would ideally mean that the participant has knowledge of numbers from 1-100, is familiar with notes and coins, and is comfortable with mental arithmetic including single-digit multiplication.

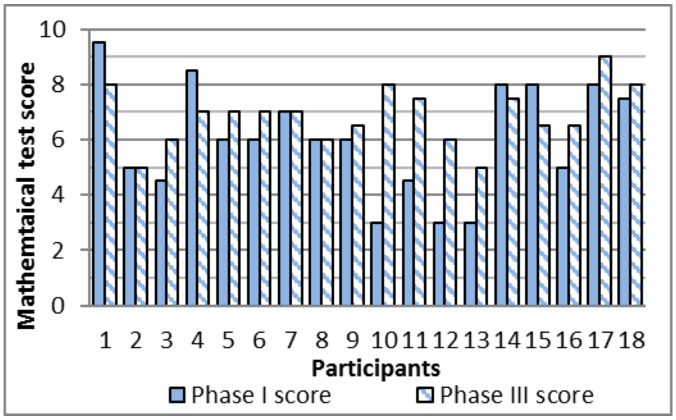

Figure 5: Phase I mathematical test scores with improvements seen in Phase III for participants P1 – P18

Phase II

A total of 218 items were purchased during the evaluation. During the first week, the participants spent an average of 9.3 minutes (SD=4.5) per session buying while during the second and third week they spent 6 minutes (SD=4.4) and 4.5 minutes (SD=2.2) respectively. Consequently, participants spent an average of 2 minutes 34 seconds per item (SD=1.39) during the first week, 1 minute 43 seconds (SD=1.94) during the second week and 1 minute 13 seconds (SD=0.57) during the third week. This reduction in transaction time per item indicates that participants became comfortable with the transactions and interactions with the application. We note here the limitation in measuring task time; we cannot differentiate the time taken for decision making and that of the actual interaction. The participants spent an average of 71 Rupees during the first week, 80 Rupees during the second week and 63 Rupees during the third week.

Phase IV

Actual store visits in phase IV provided an ecologically valid way of translating the learning from virtual to real world scenarios. 12 of the 18 participants went to a local store with an educator. Of the six participants that did not take part in phase IV, two had moved to another school and four were assessed as needing more practice with the application. It should be noted this was the first time the 12 participants had come to an actual store by themselves. The educator observed the participants interact with the shopkeeper and buy items from a given list – similar to the task in phase II.

Decision-making: Participants were able to connect the tasks during the visit to the tasks in *Kirana*, and purchase items from the given shopping list. The moderator observed a high degree of cooperation among the participants and they helped each other out wherever possible. The most popular items during the visit were cold beverages due to the warm Indian summer.

Social interaction: Participants who were feeling shy would observe the others interact with the shopkeeper before approaching the store. One participant made the valid observation that everyone should take a bill from the shopkeeper so that the purchasing amounts can be checked by an adult later. Surprisingly, three usually nonverbal participants were very responsive and able to interact with the shopkeeper.

Mathematical ability: As expected, the arithmetic aspect of calculating costs and balance was highly dependent on one's mathematical ability. Participants with a higher mathematical score (greater than 6) were comfortable with the concept of money and transaction involving arithmetic. Two participants gave all money with them to the shopkeeper: one out of nervousness and the other had problems reading the decimal place in price because of the small font size on the packet. All the participants were able to locate an item's M.R.P., understand that money is exchanged when purchasing the item, and expressed a positive attitude towards purchasing. Several participants understood the relation between items and costs such that they only selected items they knew they would be able to purchase with the amount of money they had.

Overall, we observed improvements in mathematical ability, based on the teacher's assessment, and translatable learning from the application to an actual local store. We believe *Kirana* did the necessary groundwork in preparing individuals with developmental disabilities for purchasing items of need from local stores, thereby teaching them a valuable life skill that promotes self-efficacy.

4. *BALLOONS* APPLICATION

Balloons [33] focuses on promoting the social activity of shared attention, or joint attention, between two users. It is important to motivate individuals with developmental disabilities towards such social interaction and collaboration as counter measures against social isolation. Joint attention is defined as a shared experience over a common goal or object, and is linked to language acquisition and social interaction in the later stages of the neurological development of a child [6]. For individuals with autism, encouraging joint attention is a crucial step towards social inclusion. Our goal for *Balloons* was *to allow different explorations of the challenges in learnability of joint attention using gesture for selection - pointing with dwell time*.

4.1 Design

Similar to the design of *Kirana*, we followed a user-centered design approach to design the application, involving twenty-three stakeholders: nineteen school specialists, three parents, and two high functioning children with autism. First, we conducted semi-structured interviews with the specialists, where they shared their own educational intervention guidelines as compiled over several years of hands-on work. Based on the common themes that emerged from the interviews, we conducted two focus group discussions. The end result is a list of design requirements for developing an application's interaction, interface, and goals, which are:

- Motivate and encourage social interaction even for those who have limited vocabulary, for example, via joint attention.
- Provide adequate rewards to encourage children to interact, regardless of their performance, for example, applause from the system.

- Provide a well-defined start and end of a session in order to allow the children to understand when a task is over.

- Encourage socially acceptable gestures, as gestures such as kicking or jumping, which might be fun for gameplay, can be dangerous and socially isolating outside of the session.

- Provide multimodal feedback, both visual and auditory, as children have varying levels of visual and auditory preferences, sensitivities, and capabilities.

- Include colorful graphics with smooth visual movements and illustrations to captivate attention for longer time spans.

- Empower the child by involving them in the decision making.

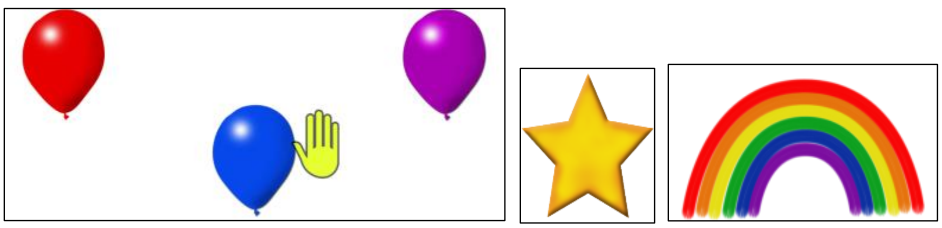

Figure 6: (left) application screen with a participant's right hand pointing near the blue balloon, (center) a star reward, (right)

Designed to address these requirements, *Balloons* provides a simple way to promote joint attention via gesture-based interaction for individuals with developmental disabilities. It consists of three colored balloons (Figure 6, left) that can be selected by pointing and holding the hand steady on a balloon (dwell time) for three seconds. An on screen hand cursor shows the position of the hand on the screen in real time. Two users can interact with the application at the same time. If one user selects a balloon, a star is rewarded (Figure 6, center). If both users select the same balloon together, they are rewarded with a growing rainbow and pleasant music (Figure 6, right). The application aims to encourage social interaction between the participants. The use of a screen to mediate interaction between participants reduces the social awkwardness of face to face social communication or making direct eye contact, which is especially challenging for individuals with autism. The application was built on the same framework (section 3.4) as *Kirana*.

An ideal *Balloons* session would require two participants to stand in front of the system, facing the screen. As a first step, both participants would identify their respective on-screen cursor, and then one would prompt the other to select a specific balloon together. The prompt can be verbal, by saying the color, or visual, by moving one's cursor to towards a specific balloon (both are instances of social interaction). The other participant would then also move their cursor towards the balloon. In this way, if both participants select the same balloon within the three second dwell time and are rewarded a rainbow (joint attention).

4.2 Participants

Fourteen individuals with a wide range of developmental disabilities participated in our evaluation (3 individuals with Down syndrome, 1 individual with high functioning autism, 1 individual with ADHD, 1 individual with cerebral palsy and an unspecified form of intellectual disability, and 8 individuals with other unspecified forms of intellectual disabilities). There were 10 females and 4 males aged between 18 to 41 years (M=26, SD=7). The participants were recruited from two different classes in the school:

Group one (n=9), with an average IQ of 42 (SD=11) and SQ of 41 SD=10). They (a) did not understand the concept of left or right, (b) communicated with a vocabulary of 50-150 words, (c) could not follow instructions, (d) had an awareness of self and body, and (e) did not participate in the *Kirana* study.

Group two (n=5), with an average IQ of 63 (SD=20) and SQ of 57 (SD=15). They (a) did not understand the concept of left or right, (b) communicated with a vocabulary of 150-500 words, (c) could follow instructions, (d) had an awareness of self and body, and (e) did not participate in the *Kirana* study.

4.3 Procedure

Balloons was set up in a classroom at Nai Disha School using a projector, a laptop and a Kinect (similar to the one in Figure 7) for ten days in February 2017. Each participant had two sessions with the application. The sessions were moderated by a male educator who is working on technology interventions at Tamana and he also moderated the *Kirana* study. However, the participants of this study were unfamiliar with him and therefore, group one's special educator (female) was also present during all their sessions. This created a comfortable and familiar social environment for the participants, who were mostly female. The participants in group one attended the sessions in pairs, i.e., in each session two participants would be present, where one participant interacted, the other watched on, and vice versa. For group two, because of its small size, sessions were conducted with the entire group present at the same time. For both groups, each session was about fifteen minutes long, with three breaks to reduce fatigue and enquire if the participant is comfortable and wants to continue. The moderator initiated communication with the participant to build a rapport with them, asked them to name the colors of the balloons, and whether they found the interaction interesting and enjoyable. In this way, the moderator made sure that the participants were willingly participating in the study.

First session: The aim of the first session was to confirm that each participant is comfortable using gestures for interaction and working with the moderator. Each participant was first asked to count the number of balloons they saw on the screen. Then they were asked to raise their right hand, during which the moderator stood behind the participant to provide support in case they experienced any fatigue. If they did, the moderator would gently support the elbow of the participant and enquire if they were comfortable (similar to Figure 7, right). This was also repeated for the left hand. The moderator then stood in the front of the screen, in view of the Kinect, and demonstrated the pointing gesture using his right and then left hand. The participant was asked to repeat the same by standing in front of the screen and identifying their hand on the screen (right hand cursor as shown in Figure 6), while the moderator supported them. The primary role of the participant in the first round of sessions was to keep the hand steady for the 1.5 seconds dwell time and select a total of five balloons.

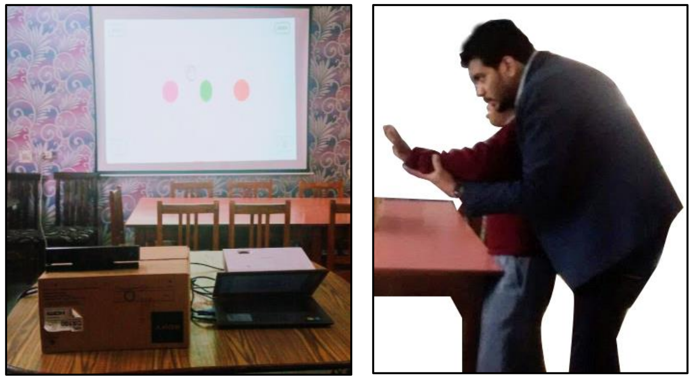

Figure 7: (left) picture of the setup and (right) moderator guiding a participant by supporting their elbow

Second session: In the second session, the participants were asked to independently select seven balloons. The moderator stood behind to support the participant by gently holding their elbow if they mentioned any fatigue.

Since the participants selected balloons with the moderator standing behind them, they were identified as one entity by the Kinect and a star was rewarded. The application was used as an introduction to gesture-based interaction, without focusing on its two-person interaction. However, we note that this usage of the *Balloons* still encourages the social skill of joint attention, that is, a shared experience over a common object. Moreover, this allowed us to prepare the participants for the HOPE study, which required single person interaction.

Data Gathered: We collected the total session time for each participant for each session, the number of balloons that were selected with assistance, and the number of balloons they selected by themselves. The moderator also noted down observations after each session for each participant, namely, the type of assistance required and if there were any other challenges.

4.4 Results

As expected, group one required more time and assistance in completing the tasks during the two sessions with *Balloons*. To compare the task times between sessions, we normalized the data gathered with respect to number of balloons selected per session, which were five in session one and seven in session two. We also define a completion score, which takes into account how many balloons were selected by the participant independently and how many required assistance.

Types of assistance

For each session there were varying levels of assistance, which can be categorized as verbal or physical, with two types of verbal prompts, one for the content and one of the gesture. Physical assistance was provided by the moderator by standing behind the participant and supporting their elbow. Initially, this included supporting the elbow to guide the participant towards the appropriate content on the screen and to reduce interaction fatigue. For the subsequent sessions, the moderator supported the participants' elbow to reduce interaction fatigue but provided no guidance. The type of assistance provided in each session is mentioned in Table 2.

Table 2: Type of assistance in sessions

Session	Verbal Assistance		Physical Assistance	
	Prompts for Content	Prompts for Gestures	To guide hand movement	To reduce fatigue
Balloon I	Yes	Yes	Yes	Yes
Balloon II	Yes	No	No	Yes

Completion Score

We define the completion score (as a percentage) for each participant for each session to understand their performance at the task. This was calculated by:

*Completion Score = [(number of balloons selected by self) + 0.5 * (number of balloons assisted)] * 100 / total balloons per session*

By giving less weightage for balloons that were selected with assistance, the completion score provides a useful metrics to compare how well the participants are able carry out the gesture and complete the task. Figure 8 shows the completion score for each group and session. As expected, Group 1 has a lower completion score than Group 2; in the first session for Group 1, all participants required assistance with the gesture for selection. While for Group 2, in the second session, all participants were able to complete the selection task without any verbal or physical assistance.

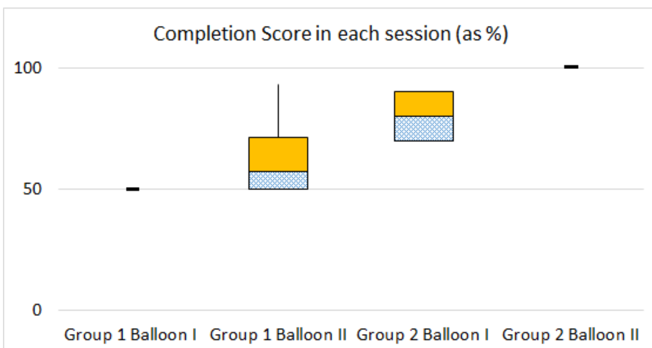

Figure 8: Completion score for *Balloons*, with boxplots showing range, median and quartiles

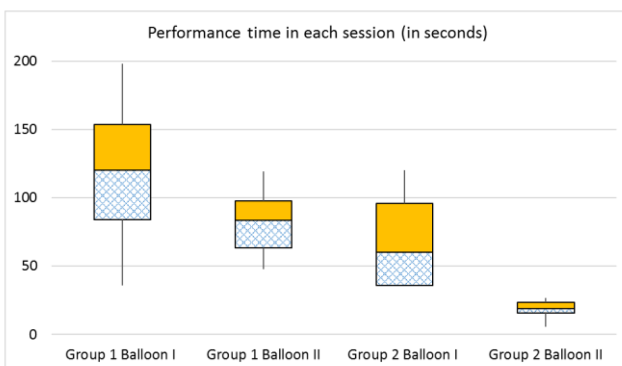

Figure 9: Performance time for *Balloons*, with boxplots showing range, median and quartiles

Performance Time

We calculated the performance time for each participant for each session. We define the performance time (in seconds) as the time taken for each individual task, i.e., time taken to select a balloon (individually or assisted). This was calculated as:

Performance time = total time for a session / total balloons per session

where the total balloons selected for that session is equal to five for session one and seven for session two. With respect to performance time, Group 1 participants required more time for each session as compared to Group 2 (Figure 8). While both groups saw reduction in performance, the reduction was more drastic was Group 2.

Observations

Group 1: In the first session, except for two participants, seven found it difficult to follow the instructions to raise their right hand. They were all able to identify the colors blue, red and purple. After the moderator re-demonstrated the instructions (raising right, left, and finally both hands), six participants were able to repeat them, while three could not. All participants were assisted by the moderator by gently supporting and nudging their elbows, and holding their hand steady. It was also noted that participants took time to communicate with the moderator and build a rapport. Participants also initially found it difficult to identify their hand on the screen. By the end of the first session, however, they were able to relate to the onscreen cursor.

In the second session, participants tried the task themselves, with the moderator occasionally supporting their elbow to reduce interaction fatigue. One participant faced particular difficulty in identifying their hand on the screen in the second session, and then lost interest in the task. This was exacerbated by a system misbehavior where the application did not detect their hand correctly. Between session, participants seemed to open up to the moderator, as they would meet him in the school corridor and excitedly enquire about their next session.

Group 2: The sessions were conducted with all participants of group two present at the same time. The first session required minimal assistance and was completed in less than the 15 minutes assigned. However, one participant with autism did not show interest unless assisted. In the second session, participants were able to select seven balloons in under 2 minutes, where moderator assistance was only required by one participant who wanted to play with their friend (one of the observers). They selected balloons together and were awarded the rainbow.

When considered together, quantitative and qualitative data provide a holistic perspective of the potential of gesture-based interaction. Although the participants of Group 1 required more time and assistance, they still completed the tasks and were engaged in the learning process. We believe this feeling of accomplishment is more valuable and rewarding than focusing solely on the actual performance time.

5. HOPE APPLICATION

HOPE is an interactive learning application that promotes motor coordination, cognitive and social skills through a series of spatial reasoning tasks. It addresses two main challenges towards sustainable educational interventions in the Indian context: technology diversity and inclusive learning.

Technology diversity: Indian users have a multitude of old and new technology within the classroom and home environment. An application that supports multiple interaction mechanisms (e.g., mouse, touch screens/tablets, or gestures using a Microsoft Kinect sensor) can be used on a multitude of devices. This enables consistent and sustained usage of the application for individuals to benefit from. For instance, a school that can afford a Kinect can provide gesture-based interventions to their students within the classroom environment, while parents can follow-up with at home sessions on a tablet.

Inclusive learning: Given the context of a special school, it is naïve to think one size fits all when it comes to educational interventions. For instance, within the autism spectrum, it is well known that needs between individuals can be quite diverse. Our previous gesture-based interventions therefore focused on a smaller subset of users, for example, *Balloons* for medium-low functioning children with autism [33], and *Kirana* for individuals with developmental disabilities who were previously unable to shop independently [34]. To be inclusive and cater for a larger spectrum, HOPE provides multiple tasks with increasing levels of difficulty. Participants, parents or educators are able to decide the level they want to start with.

Thus, HOPE provides a sustainable, consistent, and flexible educational intervention with multiple interaction mechanisms. Our research goal for this study was to *explore the challenges in learning complex gestures for selection and object manipulation for individuals with developmental disabilities*.

5.1 Design

The application combines memory, spatial, motor, and cognitive skills based on two broad learning paradigms: matching shapes, colors, and objects, and the concept of numbers and alphabet. Our focus in this article is on the matching tasks and using gestures for interaction. The matching tasks consists of three categories: colors, shapes, and objects. Each category consists of up to twenty screens with increasing level of difficulty. For instance, for matching colors, the first screen consists of three items of the same shape: two green triangles and one red (Figure 10, left). From screen six onwards, the difficulty increases by having two items of the same shape and one of a different shape, as seen in Figure 10, center. From screen eleven onwards, another level of difficulty is introduced as all the items are of different shapes. Figure 10 right shows a black star and heart, and a yellow circle, for which a participant has to match the star and heart, as they are of the same color.

Figure 10: Matching Colors (left) screen 1, (center) screen 6, and (right) screen 11

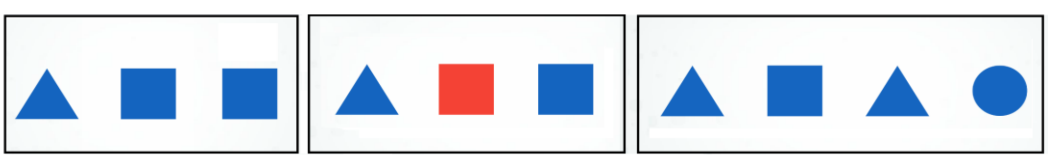

Figure 11: Matching Shapes (left) screen 1, (center) screen 6, and (right) screen 11

Similarly, for matching shapes, screen one starts with three items of the same color (Figure 11, left), while screen six introduces another color (Figure 11, center). From screen eleven onwards, the number of items increase to four (Figure 11, right). For the object matching tasks, the first screen starts with familiar objects and shapes (Figure

12, left), while new shapes and objects are introduced from screens six and eleven onwards (Figure 12, center and right). Each correct response is rewarded with an auditory applause and flying balloons for visual stimuli.

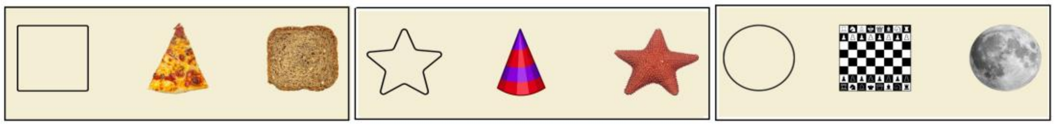

Figure 12: Matching Objects (left) screen 1, (center) screen 6, and (right) screen 11

The matching tasks require a sequence of gestures (grab > drag > drop) to complete. An ideal HOPE session will include a participant interacting with the onscreen content by solving the tasks presented using grab, drag, and drop gestures. For example, the first level of the shape matching screen shows two blue squares and one blue triangle (Figure 11, left), and the following steps are required to correctly perform the matching task using gestures:

1. Grab: A participant has to select one of the blue squares by grabbing – that is start with an open palm (shown in Figure 13, left), move their hand on top of one of the squares, and then close the palm (shown in Figure 13, right). The blue square is then attached to the onscreen hand cursor.

2. Drag: with the palm closed, the participant has to move their hand to the other blue square. The onscreen hand cursor follows the movement path, dragging the selected square with it.

3. Drop: once both squares are almost on top of each other, the participant has to open their palm.

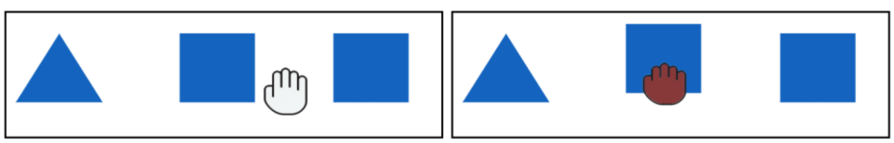

Figure 13: On screen hand cursor in (left) open palm and (right) closed palm for grab

In this way a successful grab, drag, and drop gesture sequence is performed. The application starts with an instruction screen, as shown in Figure 14.

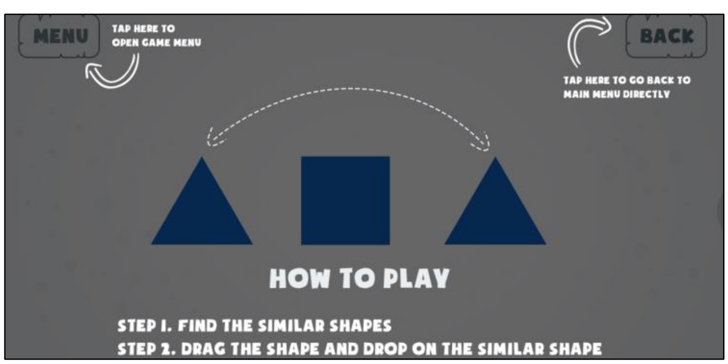

Figure 14: HOPE's how to play screen

5.2 Participants

The participants from the *Balloons* study (section 4.2) took part in the HOPE study.

5.3 Procedure

There were a maximum of three sessions conducted with each participant using HOPE over the span of two weeks in February – March 2017. Similar to *Balloons*, sessions with Group 1 were conducted in pairs and with Group 2 in a group. The task was matching *colors*, an extension to the task in *Balloons*. This task familiarity provided a base to build on, since the gesture (grab, drag, and drop) was now more complex. Verbal prompts were given throughout the sessions to help the participants with the gesture sequence, and also with the content. Physical assistance was provided to reduce interaction fatigue. The procedure for each of the sessions was as follows:

First Session: The first session included a practice session to provide a detailed explanation and demonstration of the grab, drag, and drop gesture. Unlike *Balloons*, here assistance was provided by the special educator who stood behind the participant. The special educator guided and supported the participant by gently nudging their elbow, similar to the process followed by the moderator in the first session of *Balloons* (Figure 7, right). The participant and the special educator stood facing the screen and Kinect (similar to Figure 7, left). The moderator however, stood next to Kinect, with their back to the screen, facing the participant and out of the active area. This allowed the participant to see the moderator's palm and follow the palm-open and palm-close instructions. The participant was asked to focus on the moderator's hand movement and verbal prompts, while being guided by the special educator standing behind the participant. After practicing palm-open and palm-close several times, the complete grab, drag, and drop gestures were taught. The complete practice run and the accompanying verbal prompts are explained in Table 3.

Table 3: Grab, drag, and drop gesture

Gesture	Moderator - action	Moderator – Verbal prompts	Special Educator - action
Ready	Raises his right hand, and then opens his palm	*Raise your hand. Open your hand*	Gently nudges the participant's elbow, moves hand over an object to select it on the screen
Grab	Closes the palm	*Close your hand*	Supports the participant to select the object on the screen
Drag	Moves closed palm sideways	*Move your hand*	Moves the participant's hand to the correct placeholder for the selected object
Drop	Opens the palm	*Open your hand*	Supports the participant to correctly match the object

After the going through the gesture sequence twice, the first color task screen was started. The session consisted of ten screens for matching colors, as shown in Figure 10, for up to five different colors: green, red, blue, yellow, brown. The moderator and the special educator kept repeating the same gesture practice, that is, the participant could look at the moderator's hand movement while the special educator gently guided the student's hand to complete the onscreen task. Once the participant became comfortable with the gesture, they would start looking at the screen instead of the moderator, and at this point the moderator stopped repeating the gesture sequence. Most participants learnt the gestures by the fifth task screen, and would then be able to interact with the content on the screen independently, without the special educator's assistance. To reduce fatigue, instead of breaks, the participants were asked to exercise their hand by moving both hands up and down to relax the muscles.

Second and third sessions: In the second and third session, the participants were already familiar with the grab, drag, and drop gestures. Therefore, there was no elaborate explanation or demonstration of the gesture and the moderator provided the assistance by standing behind the participant. The moderator supported the participant's elbow or hand, in case of fatigue, while the special educator observed from a distance. The second and third sessions included the same screens as the first with an additional eleventh screen, which was the most difficult and was added to increase the complexity for participants who were able to complete all ten screens, to keep them engaged (Figure 10, right). These sessions were not run with Group 2 participants since they were able to comfortably and correctly complete the task in their first session.

Data Gathered: Similar to *Balloons*, for both groups, we collected the total time per session per participant, and the number of screens that were assisted, and number of screens completed by the participant independently.

5.4 Results

In the evaluation, the level and type of assistance played a significant role. After the practice gesture sequence, Group 2 participants completed the ten task screens without any assistance. Meanwhile, Group 1 participants completed all the screens with some assistance in sessions one and two. However, because of the minimum level of assistance provided in session three, Group 1 participants completed between 5-11 screen with the longest performance time. Therefore, learning from session to session varied widely among the participants.

Types of assistance

Similar to *Balloons*, for each session there were varying levels of assistance as described in Table 4. This assistance was mainly provided to Group 1 participants.

Table 4: Levels of assistance in HOPE

Session	Assistance Verbal		Assistance Physical	
	Prompts for Content	Prompts for Gestures	To guide hand movement	To reduce fatigue
HOPE color I	Yes	Yes	Yes	Yes
HOPE color II	Yes	Yes	No	Yes
HOPE color III	No	No	No	Yes

Completion Score

We define the completion score (as a percentage) for each participant for each session to understand their performance at the task. This was calculated by:

*Completion Score = [(number of screens completed by self) + 0.5 * (number of screens assisted)] * 100 / total screens attempted in that session*

Figure 15 show the completion score for each group and session. Since Group 2 participants were able to complete all screens in their first session, no further sessions were conducted with them. For Group 1 participants, the completion scores vary between 50 - 100%. For instance, a 50% completion score in session one means that the participant was able to complete all the screens but required physical and/or verbal assistance for each. In session three, Group 1 participants completed between 5 – 11 screens.

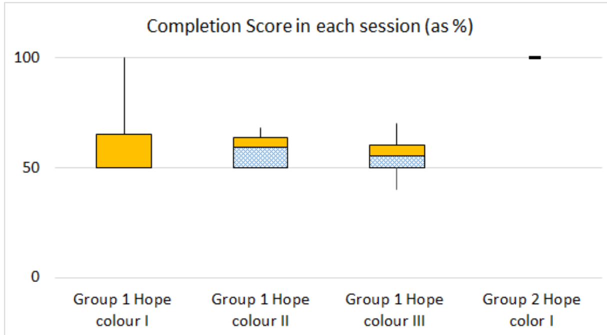

Figure 15: Completion score for HOPE

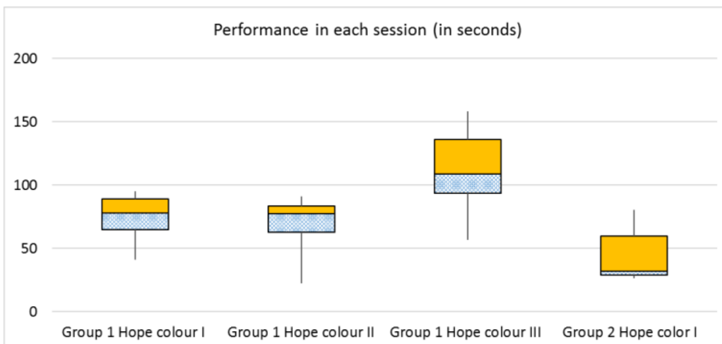

Figure 16: Performance time for HOPE, with boxplots showing range, median and quartiles

Performance Time

We calculated the performance time (in seconds) as the time taken to complete one screen (individually or assisted):

Performance time = total time for a session / total screens completed in that session

where the total number of screens completed in each session varied for each participant.

Figure 16 shows the performance time for the three sessions of Group 1 and one session of Group 2. Overall, Group 2 participants were quicker to learn the new gesture and complete the color matching tasks as compared to Group 1. Moreover, in session three, Group 1 participants required more time to complete the gesture and task as there was no verbal or guided physical assistance provided.

Observations

Group 1: The moderator observed that by the 5th or 6th screen, the participants were able to familiarize themselves with the gesture sequence, and could complete it with only verbal prompts, thus without the need to look at the moderator standing in front of them. By the end of the first session, participants were able to focus on the content of the screen. In the subsequent sessions, the gesture sequence was performed with verbal prompting (explained in Table 2), while the moderator stood behind the participant to support their hand, to reduce interaction fatigue. Since the task content was similar to *Balloons*, the participants did not face any issues with executing it correctly. This made the prime focus of the sessions on learning the gesture sequence, as the content was considered relatable. However, we note here that most participants were new to the colors brown and yellow.

The grab, drag, and drop gesture in HOPE increased the complexity of the tasks, especially in comparison to *Balloons*. In *Balloons*, since the focus was on pointing, a half opened palm did not cause any interaction difficulty. However, a known technical issue with the Kinect is that it recognizes an open palm as closed if it is below a certain height, or facing away from the Kinect. Therefore, very specific hand movements were required to execute the gesture sequence – initiation with an open palm, selection by closing the palm, dragging by moving a closed palm across the screen, and dropping by opening the palm. Group 2 participants were able to understand this gesture complexity and complete the task in their first session. The technical limitation of the Kinect did not affect them. Group 1 participants faced challenges in performing the gesture sequence and maintaining a steady hand throughout their first session, which we believe was exacerbated by the technical limitation.

In their second session, Group 1 participants had a better understanding of the gesture and improved their performance time. However, in session three, not all Group 1 participants were able to complete the entire set of 11 screens, and without verbal and physical prompts for the gesture or task, required more time to complete each task screen. Therefore, Group 1 participants would benefit from more sessions with HOPE color matching tasks, while Group 2 participants can graduate to the next learning challenge, say of matching shapes. This support for learning diversity in HOPE establishes the potential of designing an application to cater to a large group of individuals, that is, inclusion-within, one of main research goals for this work.

6. DESIGN GUIDELINES AND DISCUSSION

We presented several gesture-based applications and their evaluations with individuals with developmental disabilities in India. We investigated the use of pointing with dwell time for selection with two applications (*Balloons* and *Kirana*). We explored the challenges in gestures for object selection via grab and object manipulation via drag and drop in the HOPE application. Although complex, the gesture sequence of grab, drag, and drop for interaction was found to be learnable by individuals with developmental disabilities, in the context of a simple problem-solving task (match objects by color). On the other hand, *Kirana* combined a complex set of tasks, including mathematical calculations, decision making, and social interaction, with simple pointing gestures. Our findings indicate that there needs to be balance between task and gesture complexity, when employing gesture-based interaction.

In the following, we consolidate all our findings as guidelines that highlight the opportunities and help overcome some challenges in designing gesture-based applications for individuals with developmental disabilities, and for practitioners working in the area. We categorize the guidelines as those (a) specific to the application design, (b) specific to the research study design (i.e., ethical considerations and conflicts that might arise), and (c) specific to the Indian context. Furthermore, our guidelines adhere to the principle of *inclusion-within*, that is, designing applications that cater to diverse learners, and nurture collaboration among users. Moreover, individuals with developmental disabilities form a wide spectrum of users, who, even when grouped under the same diagnosis, have diverse abilities and preferences. The guidelines aim to respect these differences and preferences when it comes to designing interactive applications, for instance, by providing multimodal feedback, having tasks that steadily increase in difficulty, and finding the right balance between gesture and task complexity.

6.1 Guidelines for Application Design

Provide a clear start and end of gameplay

It is extremely important to provide a clear start and end of gameplay to avoid ambiguity and confusion during interaction. For example, in *Balloons* and *Kirana*, when a participant enters the active area of the Kinect, i.e., the area most suitable for interaction, the screen turns white from black to reveal the virtual objects. Similarly, when the participant moves out of the active area, the screen turns black to indicate the end of interaction, that is, the application no longer responds to gestures. This can also assist in reducing excessive attachment with the application and discourage repetitive gestures and behaviors, a common concern when designing for individuals with developmental disabilities, especially autism [25]. *Kirana* also caters to various end-of play scenarios such as not enough money left to buy an item, or running out of money or items to buy. In each of these scenarios, the reason for the end is stated, followed by a visual and aural "well done" feedback to conclude the session. Therefore, each session with the application has a clear start and end, which provides structure to the interaction.

Provide feedback with multiple modalities

Individuals with developmental disabilities find it easier to select items that have both visual, e.g., a change in size and auditory, e.g., name of the item said aloud, feedback. This reinforced feedback with multiple modalities helps overcome visual or auditory impairments, if any, and provides multiple stimuli for attention ([2],[3]). In *Kirana*, each visual object, upon being selected, also had an auditory feedback. In HOPE, the moderator provided verbal prompts for gestures and content, to assist the participants. Going forward, it would be better to incorporate the verbal prompts in the application, with the option to control the volume. Overall, providing multimodal feedback allows for diverse stimuli that makes an application suitable for a larger group of participants.

Offer rewards and positive reinforcements

One way to encourage and motivate social or physical interaction is through continuous rewards and positive reinforcements. In *Balloons*, every selection is rewarded with a star (individual selection), or a growing rainbow with playful music (jointly selected). In HOPE, every correct answer is rewarded with flying balloons and an auditory applause. Physical rewards such as a high five, clapping, and other harmless activities that are of interest and liking of the user can be used as rewards. A reward can also be a time-bound session with another favorite game to further maintain a participant's engagement level. Furthermore, it is important to add reinforcements like "*shabash*" (well done in Hindi), "well done", or an applause, regardless of the correctness of the task, to encourage and motivate participants to continue playing. Therefore, rewards and positive reinforcement should be designed for continuous encouragement and motivation.

Evolving task difficulty levels with easy transitions

Applications should provide varying levels of difficulty with enough repetitions, to allow the educator or caregiver to customize learning goals for each individual and cater to a large group of participants [2]. Moreover, it should be easy to repeat a level or specific task, and to also transition to the next level. For instance, as shown in Figure 10, in HOPE the matching colors task start with three items of the same shape: two green triangles and one red. From the sixth screen, the difficulty increases by having two items of the same shape and one of a different shape. Then, from the eleventh screen, another level of difficulty is introduced by having all items of different shapes. In this way, the level of difficult increases every five screens. Moreover, participants can start from any screen, that is, there is no need to *unlock* the difficult levels. This enables participants to select content that is interesting and suitable, without having to go through disinteresting or disengaging content first. Therefore, applications that introduce increasing levels of difficulty in their tasks, should allow for easy transitioning between levels, and enough repetition of each level.

Provide serial and structured content

Designing content that is structured and follows a series of related steps provides a way to break down complex tasks into manageable steps. For instance, in *Kirana*, gesture interaction simulated the real world scenario of buying groceries from a local store. This was achieved by designing interaction that is socially and culturally appropriate and breaking down complex tasks that require social, mathematical, and decision-making skills. Moreover, the purchasing animations were slow and sequential. In HOPE, matching tasks are structured based on level of difficulty, with enough repetition of tasks for each level. Furthermore, designing serial and structured content also supports designing a clear start and end of gameplay, and evolving task difficulty levels with easy transitions. We recommend complex tasks to be broken down into a sequence of smaller achievable steps. If the animation and other visual media content also follow the same step by step order, individuals with developmental disabilities can follow the progress of the tasks with ease.

Simulate real world scenarios

Gesture-based applications offer the potential of simulating real world scenarios in a safe and controlled environment. This can be achieved in two ways: by incorporating socially and culturally appropriate interactions, and by designing the content to match a real world activity. Moreover, using free form body gestures provides a mechanism for more inclusive social interaction and team work, especially when compared to touch screens that have limited form factor and surface area to support multiple users. This inherent nature of embodied interaction [38] makes for an immersive and engaging experience, which can be utilized to encourage real world social interactions. In *Kirana* and *HOPE*, even though used by a single active user, the interaction was visible to the moderator and others who could then prompt and encourage the participant, thus involving them in the activity. Overall, an application can be designed to simulate a real world scenario through their content or interaction, or both.

Balance gesture and content complexity

Interactive applications incorporate two learning paradigms in one task – the content and the method of interaction. When designing application with varying levels of difficulty, it is important to make incremental changes to either the content or interaction method. For instance, in HOPE, the interaction gesture was the same for different tasks, such as matching color, shapes, or objects. In this way, the learning of the gesture sequence from matching colors is potentially translated to other more complex matching tasks. Moreover, the concept of selecting a color was translatable from *Balloons* to HOPE, when the gesture for selection increased in complexity, from pointing with dwell time to grab, drag, and drop. Therefore, we believe that it is important to incrementally add to the task complexity, through either content or gesture, to allow for incremental learning.

6.2 Guidelines for Research Study Design

Design for diverse learners for inclusion-within

A school environment caters to diverse group of learner, separated either on the basis of chronological age or cognitive abilities, to provide learner appropriate goals. Likewise, gesture-based applications for individuals with developmental disabilities have usually focused on a narrow group of learners. However, Nai Disha School has identified a growing need to create a sense of unity between diverse groups of learners. With HOPE we redefined inclusivity and designed an application that could cater to students within the whole school and nurture collaboration, instead of competition. This provides an additional learning opportunity: how to accept the strengths and challenges of other students, i.e., *inclusion within*. Our guidelines for application design are also attuned to this goal.

Promote self-efficacy

Applications that promote self-efficacy should employ gestures and interactions that are socially and culturally relevant, and acceptable for the specific life-skill. This is particularly relevant in the developing world where social awareness about individuals with disabilities varies greatly. Moreover, the socio-cultural norms of a society dictate the degree of inclusivity and integration of individuals with developmental disabilities ([4],[15]). Individuals with developmental disabilities, such as autism, display a tendency towards repetitive behavior and teaching socially unacceptable gestures might increase their isolation, even if the gesture is fun during gameplay [2]. In this respect, *Kirana* is an application that is culturally relevant and socially acceptable to the community. We believe it did the necessary groundwork in preparing individuals with developmental disabilities for purchasing items of need from local stores, promoting self-efficacy.

Address socio-ethical challenges

Conducting research with individuals with developmental disabilities poses several ethical challenges. First is the question of a control group, as is common with user research. For Nai Disha School, it is important to include each student in activities they have the potential to benefit from. Thus, it is against their ethics in creating a control group that is denied any form of participation. Second, answering the research hypothesis might require a certain level of moderator restraint to collect comparable data. One example of this is the time taken per task. Moderator assistance or taking several breaks, would not allow for comparable task times. However, it is more important to provide assistance to the participant when requested or required. In HOPE, our initial research aim was to look at the learning curve of the grab, drag, and drop gesture sequence, for which it was important to provide limited assistance. Yet, the moderator quickly realized the need for several levels of assistance, and provided the same. Third, participants have a tendency to get attached to a specific application, and they show an interest to continue after the study sessions are over, which Nai Disha School gladly facilitates. Therefore, any resolution of conflict between the research procedure and the participants' interest should always be in favor of the participants.

Provide assistance for gestures

Using gestures for interaction can induce fatigue and require uncomfortable motor movements. Therefore, for individuals with varying motor and physical strengths, it is important to provide assistance to reduce fatigue. During the HOPE sessions, the moderator provided physical assistance to avoid gesture interaction fatigue. In the initial sessions, when participants were learning the gestures, the educator supported their elbows and gently guided hand movements while standing behind them. Once a participant was comfortable with the gesture sequences, the educator continued to provide elbow support to reduce interaction fatigue. This can also be achieved by using a stand, which does not affect the motion recognition technology. It is thus important to provide support and assistance for interaction methods that require motor or physical exertion.

6.3 Guidelines for the Indian Context

Design for technology acceptance

We followed a collaborative and stakeholder inclusive approach for our work, from *Kirana*, to *Balloons*, and HOPE. Since the applications were designed and developed with educators, therapists and parents, who provided valuable insights and experience, there was no resistance towards technology acceptance. The applications were integrated within the classroom environment with ease. Additionally, we identified ways to address the challenges in introducing new technological interventions for individuals with developmental disabilities. First, economic barriers for technologies that are too costly can be overcome by designing applications that can be integrated within schools, and can be personalized and customized for use by a larger group of individuals. Second, resource constrains can be overcome through collaborations between schools, universities, and industry partners. Third, the digital divide can be reduced by spreading awareness of the benefits of technology within schools and to parents. In India, gesture-based technology is rare and novel, and thus its acceptance remains largely unexplored. Our findings suggest that following a user-centered, collaborative, and stakeholder inclusive design approach can reduce the challenges in technology acceptance.

Address technology fears

During the course of our work at Nai Disha School, several parents enquired about the harmful effects of technology mediated educational interventions, from being addictive to possible sources of radiation. As researchers and practitioners that advocate technology, one needs to be very clear on what it means to use technology. For instance, we explain to parents that their children use technology at the school for 15-30 minutes a day, and that the usage of technology is always monitored and moderated. Moreover, most of our interventions employ gesture interaction, which does not require holding of any devices. In fact, the Microsoft Kinect motion sensor works best if the participant is between 2-3 meters away from the device. Individuals with developmental disabilities, specifically with autism, get attached to games or devices, which can socially isolate them even further. Therefore, it is important to discuss technology apprehensions that parents and care givers might have, including technology addiction. This becomes especially relevant when there is potential for misinformation and misguidance around socially repressed topics, such as autism and developmental disabilities, which is common in the Indian context.

Give control to the educators

One important way to increase technology acceptance, is to provide control to the special educator. As discussed already, due to the wide range of abilities of an individual, the educator or moderator should be able to customize the application, and its levels and screens, based on individual capabilities and interest of each user. While we believe providing control to the educator is important in the Indian context, it is universally applicable as shown in previous research ([2],[3],[33]). Moreover, in India there is a particular need to personalize learning goals to cater to a large number of individuals, to increases technology acceptance and cost-effectiveness. For instance, in *Kirana*, the educator guided the social interaction and the task by giving a shopping list. This enabled the educator to customize the learning, even on the fly. Moreover, the number, type, and price of items, and the total amount of money in the wallet could be changed easily between sessions. In HOPE, the educators had access to all the screens, encouraging them to first understand the sequence and complexity of the task. This could help them identify the most suitable starting level for each student. Furthermore, when designing an inclusive and customizable application, including a detailed curriculum of the tasks and learning goals can be incorporated as a guide. This guide can be made accessible from each screen with a small tab, something already available in the TOBY Playpad application [39].

7. CONCLUSION

This article presented several studies with gesture-based applications for individuals with development disabilities in India. The applications included *Kirana*, which teaches life skills to individuals with developmental disabilities by breaking down complex tasks, *Balloons*, which promotes joint attention skills, and HOPE, an application for improving motor coordination, joint attention, and cognition. Based on the exploratory user studies with these applications, we presented design guidelines in favor of gesture-based systems for enabling educators and caregivers across the globe to impart social, motor and life skills to individuals with developmental disabilities. Furthermore, our guidelines focus on applications that are inclusive and cater to a larger group of individuals to nurture collaboration, and help overcome challenges towards technology acceptance and adoption in India.

For individuals with development disabilities, technology provides a safe, controlled, and predictable learning environment, as shown by research all across the world. Gesture-based interaction provides an additional leverage by allowing for culturally appropriate social interactions and collaborations, reducing social isolation while also providing a means for self-expression. Involving different stakeholders, for instance, parents, therapists, and special educators, in the designing process reduces the barriers towards technology acceptance and adoption. Furthermore, applications that can be customized for different individuals, and diverse learning goals, cater to a larger user group, enabling efficient use of resources. Going forward, we believe that our work, and these guidelines pave the way for designing applications that can be used across schools in different cities, in India and globally.

8. ACKNOWLEDGMENTS

Our thanks to the school for their valuable time and guidance, and a special thanks to all the participants. We also like to thank Sanna Grönlund for developing the graphics for the *Balloons* and *Kirana* applications.

REFERENCES

[1] Alyssa Alcorn, Helen Pain, Gnanathusharan Rajendran, Tim Smith, Oliver Lemon, Kaska Porayska-Pomsta, Mary Ellen Foster, Katerina Avramides, Christopher Frauenberger, and Sara Bernardini. 2011. Social communication between virtual characters and children with autism. Artificial Intelligence in Education. Springer Berlin Heidelberg, 7-14.
[2] Bartoli, L., Corradi, C., Garzotto, F., & Valoriani, M. (2013, June). Exploring motion-based touchless games for autistic children's learning. Proc. Interaction Design and Children (IDC) (pp. 102-111). ACM.
[3] Bartoli, L., Garzotto, F., Gelsomini, M., Oliveto, L., & Valoriani, M. (2014). Designing and Evaluating Touchless Playful Interaction for ASD Children. Proc. Interaction Design and Children (IDC)
[4] Boujarwah, F. A., Hong, H., Abowd, G. D., & Arriaga, R. I. (2011). Towards a framework to situate assistive technology design in the context of culture. In The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility (pp. 19-26). ACM.
[5] Chang, Y. J., Chen, S. F., & Huang, J. D. (2011). A Kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities. Research in developmental disabilities, 32(6), 2566-2570.
[6] Tony Charman. 2003. Why is joint attention a pivotal skill in autism? Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 358(1430), 315-324.
[7] Christinaki, E., Triantafyllidis, G., & Vidakis, N. (2010). A gesture-controlled Serious Game for teaching emotion recognition skills to preschoolers with autism. In Proceedings of fdg. 2013, 417-418.
[8] Cromby, J. J., Standen, P. J., Newman, J., & Tasker, H. (1996). Successful transfer to the real world of skills practised in a virtual environment by students with severe learning difficulties. In Proc. 1st European Conference on Disability, Virtual Reality, and Associated Technologies (pp. 8-10).
[9] Damasio, A. R. (1999). The feeling of what happens: Body and emotion in the making of consciousness. Houghton Mifflin Harcourt.
[10] Doll, Edgar Arnold. Vineland social maturity scale: Condensed manual of directions. American Guidance Service, 1965.
[11] Dourish, P. (2004). Where the action is: the foundations of embodied interaction. MIT press.
[12] Filmer, D. (2008). Disability, poverty, and schooling in developing countries: results from 14 household surveys. The World Bank Economic Review, 22(1), 141-163.
[13] Niall Firth. 2012. Kinect cameras watch for autism. New scientist, 214(2863), 17.
[14] Gover, M. R. (1996). The Embodied Mind: Cognitive Science and Human Experience (Book). Mind, Culture, and Activity, 3(4), 295-299.
[15] Grinker, R. R. (2008). Unstrange minds: Remapping the world of autism. Da Capo Press.
[16] Ke, F., & Im, T. (2013). Virtual-reality-based social interaction training for children with high-functioning autism. The Journal of Educational Research, 106(6), 441-461.
[17] Khakhar, J., & Madhvanath, S. (2010, November). Jollymate: Assistive technology for young children with dyslexia. In Frontiers in Handwriting Recognition (ICFHR), 2010 International Conference on (pp. 576-580). IEEE.
[18] Lakoff, G., & Johnson, M. (1999). Philosophy in the flesh: The embodied mind and its challenge to western thought. Basic books.
[19] Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by embodiment six precepts for research on embodied learning and mixed reality. Educational Researcher, 42(8), 445-452.
[20] Mäkelä, V., Heimonen, T., & Turunen, M. (2014). Magnetic Cursor: Improving Target Selection in Freehand Pointing Interfaces. In Proceedings of the International Symposium on Pervasive Displays (PerDis '14), 112, ACM,
[21] Malin, A. J. "Indian adaptation of Wechsler's Intelligence Scale for children." Manual. Indian Psychological Corporation. Lucknow (1966).
[22] Microsoft Corporation, Kinect for Windows Human Interface Guidelines v1.8.0, 2013.
[23] Padgett, L. S., Strickland, D., and Coles, C. D. (2006). Case study: Using a virtual reality computer game to teach fire safety skills to children diagnosed with fetal alcohol syndrome. Journal of Pediatric Psychology, 31(1), 65–70.
[24] Parasuram, K. (2006). Variables that affect teachers' attitudes towards disability and inclusive

education in Mumbai, India. Disability & Society, 21(3), 231-242.
[25] Parés, N., Carreras, A., Durany, J., Ferrer, J., Freixa, P., Gómez, D., & Sanjurjo, À. (2005, June). Promotion of creative activity in children with severe autism through visuals in an interactive multisensory environment. In Proceedings of the 2005 conference on Interaction design and children (pp. 110-116). ACM.
[26] Parés, N., Carreras, A., Durany, J., Ferrer, J., Freixa, P., Gómez, D., Kruglanski, O., et al. 2004. MEDIATE: An interactive multisensory environment for children with severe autism and no verbal communication. In Proc. of the Third International Workshop on Virtual Rehabilitation.
[27] Parsons, S., Mitchell, P., & Leonard, A. (2004). The use and understanding of virtual environments by adolescents with autistic spectrum disorders. Journal of Autism and Developmental disorders, 34(4), 449-466.
[28] Parsons, S., & Cobb, S. (2011). State-of-the-art of virtual reality technologies for children on the autism spectrum. European Journal of Special Needs Education, 26(3), 355-366.
[29] Parsons, T. D., Rizzo, A. A., Rogers, S., & York, P. (2009). Virtual reality in paediatric rehabilitation: A review. Developmental Neurorehabilitation, 12(4), 224-238.
[30] Reid, D. (2002). Virtual reality and the person-environment experience. CyberPsychology & Behavior, 5(6), 559-564.
[31] Ringland, K. E., Zalapa, R., Neal, M., Escobedo, L., Tentori, M., & Hayes, G. R. (2014, September). SensoryPaint: a multimodal sensory intervention for children with neurode-velopmental disorders. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (pp. 873-884). ACM.
[32] Sampath, H., Sivaswamy, J., & Indurkhya, B. (2010). Assistive systems for children with dyslexia and autism. ACM Sigaccess Accessibility and Computing, (96), 32-36.
[33] Sharma, S., Srivastava, S., Achary, K., Varkey, B., Heimonen, T., Hakulinen, J. S., ... & Rajput, N. (2016, February). Promoting Joint Attention with Computer Supported Collaboration in Children with Autism. In Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (pp. 1560-1571). ACM.
[34] Sharma, S., Srivastava, S., Achary, K., Varkey, B., Heimonen, T., Hakulinen, J. S., ... & Rajput, N. (2016). Gesture-based Interaction for Individuals with Developmental Disabilities in India. In Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '16). ACM, New York, NY, USA, 61-70. DOI: https://doi.org/10.1145/2982142.2982166
[35] Sharma, S., Srivastava, S., Achary, K., Varkey, B., Heimonen, T., Hakulinen, J. S., ... & Rajput, N. (2016). Kirana: A Gesture-based Market App for Life Skills Learning for Individuals with Developmental Disabilities. In Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '16). ACM, New York, NY, USA, 291-292. DOI: https://doi.org/10.1145/2982142.2982149
[36] Standen, P. J., & Brown, D. J. (2005). Virtual reality in the rehabilitation of people with intellectual disabilities: review. Cyberpsychology & behavior, 8(3), 272-282.
[37] Thompson, E. (2007). Mind in life: Biology, phenomenology, and the sciences of mind. Harvard University Press.
[38] Wilson, M. (2002). Six views of embodied cognition. Psychonomic bulletin & review, 9(4)
[39] Venkatesh, S., Phung, D., Duong, T., Greenhill, S., & Adams, B. (2013, April). TOBY: early intervention in autism through technology. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 3187-3196). ACM
[40] Pratibha Vellanki, Stewart Greenhill, Thi Duong, Dinh Phung, Svetha Venkatesh, Jayashree Godwin, Kishna V. Achary, and Blessin Varkey. 2016. Computer assisted autism interventions for India. In Proceedings of the 28th Australian Conference on Computer-Human Interaction (OzCHI '16). ACM, New York, NY, USA, 618-622. DOI: https://doi.org/10.1145/3010915.3011007
[41] Wechsler, David. "Wechsler intelligence scale for children." (1949).

Received April 2017; revised July 2017; accepted xxx 2017