Through the Looking Glass: Pretend Play for Children with Autism

Zhen Bai[1], Alan F. Blackwell[2], George Coulouris[3]

University of Cambridge

Figure 1: Participants interacting with the AR system with the car, train and airplane theme

ABSTRACT

Lack of spontaneous pretend play is an early diagnostic indicator of autism spectrum conditions (ASC) along with impaired communication and social interaction. In a previous ISMAR poster [2] we proposed an Augmented Reality (AR) system to encourage pretend play, based on an analogy between imaginative interpretation of physical objects (pretense) and the superimposition of virtual content on the physical world in AR. This paper reports an empirical experiment evaluating that proposal, involving children between the ages of 4 and 7 who have been diagnosed with ASC. Results find significantly more pretend play, and higher engagement, using the AR system by comparison to a non-augmented condition. We also discuss usability issues and design implications for AR systems that aim to support children with ASC and other pervasive developmental disorders.

Keywords: Augmented Reality, pretend play, autism, children.

Index Terms: H.5.1 Multimedia Information Systems

1 INTRODUCTION

Pretend play is a familiar childhood behaviour, in which aspects of the real world are interpreted symbolically or non-literally [8]. Specific varieties of pretend play include: object substitution (e.g. pretending a banana is a telephone); attribution of absent properties (e.g. pretending a toy oven is actually hot); or presence of imaginary objects (e.g. holding an imaginary toothbrush) [23]. Pretend play is closely associated with the development of general cognitive and social skills [26], and is often impaired in children with ASC [22]. As a result, deficits in shared imaginative play have been recognized as a diagnostic criterion in ASC by the American Psychiatric Association [1]. While children with ASC may engage in pretense when instructed to do so, they still find it difficult to develop creative extensions [21].

Several behavioral approaches have been applied to teach autistic children to perform pretend play in order to increase their social interaction with other children in an inclusive play environment. Verbal/physical prompts and adult-led modeling are the most frequently used treatments and experiment results show moderate effectiveness [7]. One major concern, however, is that the child might simply imitate the modeled behaviors, without actually forming an intention of play activity as normal children do [24]. Moreover, these treatments do not emphasize the generation of children's own play ideas, which are expected to be voluntary and flexible instead of directly instructed or modeled. Therefore researchers are still seeking alternative scaffoldings to increase the intrinsic motivation of autistic children engaged in pretend play.

This formed the basis for our proposal to create an AR system to scaffold pretend play [2] (Figure 1 shows children interacting with the AR system in three different vehicle themes). Developmental psychologists note that pretend play relies on dual representations of reality and pretense. For example, Piaget argued that the mental image of an absent object assimilated to a present object is evoked during pretend play [26]. In literalistic childrens' play, the mental representation is closely related to the physical one (e.g. when playing with a physical toy car, only a literal mental representation of a car is involved). In contrast, when a child plays with a block as if it is a car, two different representations are involved - the literal object (the block) and an imaginary mental alternative (the car). AR combines real and virtual by overlaying virtual content on the real world. It naturally assists visualizing the intention of pretense in reality.

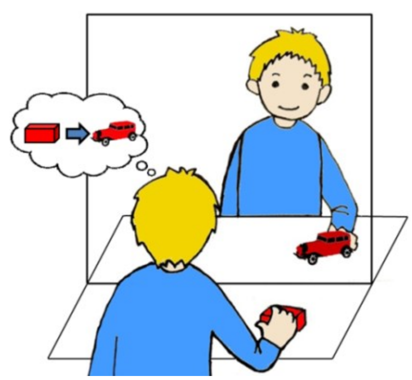


Figure 2: In reality the child holds a block in his hand. In the AR display, an imaginary car overlays on the block.

[1]zhen.bai@cl.cam.ac.uk
[2]alan.blackwell@cl.cam.ac.uk
[3]george.coulouris@cl.cam.ac.uk

IEEE International Symposium on Mixed and Augmented Reality 2013
Science and Technology Proceedings
1 - 4 October 2013, Adelaide, SA, Australia
978-1-4799-2868-2/13/$31.00 ©2013 IEEE

To summarise our previous proposal as shown in Figure 2, the goal of the AR system is to help children with ASC construct a mental representation of pretense by presenting a view of the world in which a simple play object (a wooden block) is replaced by an imaginary alternative (a car). The augmented car tracks the position of the block in the scene, so that the child can manipulate imaginary scenarios that are also visibly represented. The visual rendering of the otherwise invisible imaginary world supports the child to carry out actions in that world, refer to situations, and extend them in novel ways. Before the start of the research reported in this paper, the usability of the AR system was evaluated with feedback from experts and in a pilot study with normally developed children [2] [3].

The remainder of this paper describes the first evaluation of our system with the intended user group of children with ASC. The main contributions are as follows:

- We confirm that the AR system can be used by children with ASC aged 4-7.
- We find that the AR system promotes more episodes and longer duration of play for children with ASC than a non-augmented alternative.
- We find that the AR system results in a higher level of engagement for children with ASC than the non-augmented alternative.

We are able to draw on findings from our empirical study to offer more general design guidance for AR systems intended for related user groups.

2 Related Work

Although we are not aware of an AR system specifically proposed for encouraging pretend play for children with ASC, there is an emerging focus to design AR systems for children with special needs. Several systems aim to encourage social interaction among children with ASC [13] [14] [34] and learning or physical disabilities [9]. Other systems were proposed to teach daily life knowledge [18] [29] or enhance motor and cognitive perception [10] [12]. Radu *et al.* [27] provided a timely review of AR usability issues raised by special developmental capacities of young children. This provides a solid basis for designing AR systems for children. Research mentioned above explicitly and in [27] demonstrates that children with special needs can benefit from carefully designed AR systems to different degrees. These systems consider play as a vehicle to achieve a target behavior or knowledge. Our research, on the other hand, aims to tackle the potential capability of AR technology in facilitating the conceptualization and exercise of pretend play.

With regard to our theoretical goal of promoting symbolic activities for children with ASC, Herrera *et al.* [16] proposed a Virtual Reality (VR) system to teach symbolic transformation of objects seen in real life scenarios such as supermarkets. The transformations were demonstrated in an embedded video format. Positive effect in symbolic understanding was reported based on a small group experiment involving two children. We believe that compared to this VR system, AR technology can further encourage imaginative activities by allowing children to interact directly with physical props, as they do in natural play.

3 System Description

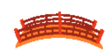
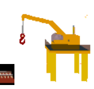
3.1 System Design

The AR system has been designed on the metaphor of a mirrored view of reality enriched with AR augmentations, as demonstrated in Figure 3. We chose the mirror metaphor since (1) it allows interacting with the system without wearing or holding the display equipment; (2) it is hands-free which allows for bimanual manipulation of the toys; (3) viewing oneself in the mirror is typically a familiar and comfortable experience for children; and (4) it provides a shared and consistent visual experience for all concurrent users.

Figure 3: A child is interacting with the AR system.

We chose vehicles as the play theme because researchers have observed that autistic children often show an obsessive interest in machinery [5]. Taking three of the most popular vehicle types, car, train and airplane, we developed three play scenes. Each scene integrates three types of augmentation intended to encourage successively more complex pretend play behaviors. The vehicle related augmentations are summarized in Table 1.

Table 1. Summary of AR objects augmented on blocks (Blk), box (Box) and in the environment (Evt).

	Blk1	Blk2	Blk3	Box	Evt1	Evt2
Car	 car	 school bus	 petrol	 school	 bridge	 dusty
Train	 train	 carriage	 light	 train station	 track	 crane
Plane	 airplane	 helicopter	 stairs	 hangar	 runway	 fire

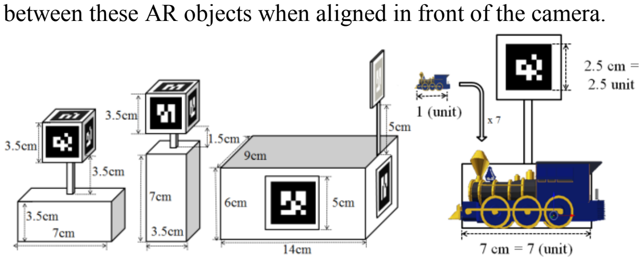
The first type of behavior is spontaneous engagement with the system. The corresponding augmentation, wherein we overlay vehicles on the blocks, is designed to encourage basic actions towards the substituted object (e.g. drive the block on the table). To increase engagement with the system, additional visual stimuli are added to the vehicle overlays, such as spinning propellers and rotating tires. The second type of behavior is the development of more complex, situationally appropriate play ideas involving multiple augmented toys. The augmentations provide vehicle-related props that encourage these actions and ideas (e.g. drive the block into the train station, or fill the block with petrol). The third type of behavior is to mix non-augmented toys into the play scenarios, thus extending the augmented play ideas on to non-augmented, open-ended props. The types of additional props provided include pen lids, cotton balls, popsicle sticks, a square of felt, and other similar nondescript items. Such items are frequently used in pretend play experiments. It is easier for the child to inhibit the original function of these objects when performing object substitution. To encourage this third type of behavior, we developed a series of virtual props that are compatible with the vehicle theme. For instance, in the airplane theme, the helicopter is in fact a rescue helicopter, and we provide a virtual fire with a cry for help so as to encourage the child to act out a rescue scene, hopefully involving some non-augmented props such as pen lids to play the role of those in need of rescue. The experimenter can dynamically switch between vehicle themes and show/hide augmentations registered in the environment (Evt1 and Evt2 in Table 1).

3.2 System Implementation

Marker-based tracking is commonly used in AR applications (e.g. [15][25]). The two primary concerns which informed our decision to use marker-based tracking are flexibility of object choice and avoidance of hand occlusion. Unlike model-based tracking which requires pre-built 3D models, marker-based tracking can easily extend the choice of objects to be tracked. Marker-based tracking can also limit the impact of hand occlusion by offsetting the marker placement from the main body of the object. The system is based on a locally modified version of Goblin XNA [19] and the ALVAR tracking library [20]. Augmentation jitter was a problem in early prototypes which we minimized by applying the double exponential smoothing method in Goblin XNA.

The configuration file that informs the system of the AR and virtual objects contains three scenes (car, train, and airplane). Each scene contains two lists of objects: (1) AR object list maintaining a list of AR objects, each one associated with a fixed number of markers and a virtual object. Among all detected markers of the same AR object, the system will select the one with the biggest area and retrieve its position and orientation. It then superimposes the coupled virtual object on the AR object accordingly. (2) Evt object list maintaining a list of virtual objects to be augmented on the table surface (e.g. train track). Those virtual objects are statically registered with a calibration marker instead of a physical marker placed on the table in order to avoid occlusion (e.g. moves a block along the train track). Prior to the experiment we place the calibration marker in the middle of the table and record its transformation matrix, and then remove the marker. During the system runtime, virtual objects in the Evt object list will be registered according to the position and orientation of the calibration marker. The experimenter can show/hide/switch the Evt objects by pressing hotkeys with a mini Bluetooth keyboard.

The installation of markers on each AR object is illustrated in Figure 4(a). The way it is designed is to keep a high degree of freedom in tracking and minimize the chance of marker occlusion between these AR objects when aligned in front of the camera.



(a) (b)

Figure 4: The marker installation and unit mapping between physical and virtual objects.

The dimension and position of the virtual object is designed to approximate that of its associated AR object. Figure 4(b) illustrates the process of aligning the virtual unit in the 3D modeling environment with the physical unit. When creating a new virtual object, we made its X dimension in one unit. Knowing the physical size of the marker, we can make one unit in the virtual environment equal to one centimeter when configuring the marker. The final step is to scale the virtual object by the length of the physical object, and translate its position accordingly to match the center of the physical object. All the information in the final step is configured in the configuration file that will be loaded by the AR system when it launches. This enables the system to automatically adjust the dimension of the virtual object according to the physical shape of its coupled AR object.

4 Expert Feedback and Pilot Study

We demonstrated the prototype system to several psychology experts and received much positive feedback. We also brought the system to a local autism event and two children with ASC aged 3 years 9 months and 2 years 11 months tried out the system. Both children explored the system by manipulating blocks in front of the camera. Neither of them carried out meaningful play due to their young age and severely impaired play behaviors (mainly engaged with sensorimotor play like mouthing or banging), but this provided insight into the potential user group that the system is usable for.

A pilot study was conducted to test whether normally developed children in the target age group could interact successfully with the AR system. The study invited four neurotypical children as subjects, in the age group 4 to 5 (two boys, two girls, average age 58.6 months) because children of this age are highly engaged with pretend play, while potential usability issues of AR systems dedicated for such low age groups remain largely unexplored.

The subjects had no difficulty performing object manipulation in the AR scenes, using both hands to perform simple manipulations such as grasping, moving, rotating and positioning an AR object relative to another object (for example parking the airplane in a hanger). They had slightly more difficulty locating augmented objects relative to another object represented entirely virtually, as for example in moving a block visualized as a car over a virtual bridge. We hypothesize that this was due to the absence of haptic feedback and to difficulty with depth perception in the mirrored AR view.

Overall the pilot study confirmed that hand-eye coordination of neurotypical subjects in the chosen age range is sufficiently developed to enable them to use the AR system. The subjects spent as much time playing in the AR setup as in the non-augmented one and they reported somewhat greater satisfaction with their play in the AR situation. This judgment was supported by both the participants and their parents.

Several suggestions for improvements to the system design emerged from our observation of play by the pilot study subjects. (1) We noticed that the AR props presented in the AR system were rather simple and might limit pretend play. To improve this, we added additional situational cues in each scene (e.g. school bus/building, train station, rescue helicopter and fire). (2) One participant in the pilot study was very interested in how virtual objects were shown on the display. Considering that autistic children are likely to be interested in computer technology, we added a familiarisation session before the main tasks. The participant sees virtual rectangles in different colors augmented on the AR objects. They are then allowed to explore freely for up to five minutes to get familiar with the technology and minimize potential “wow” effects. (3) We chose to keep colors consistent for physical props of the same type to avoid color matching play; (4) We replaced props using ‘interesting’ materials with similar ones made of plainer material (e.g. hair rollers covered by velcro were replaced by kitchen towel rolls) to avoid simple manipulation out of pure sensory curiosity.

5 EXPERIMENT DESIGN

We designed a within-subject experiment to examine the positive effects of the AR system in promoting pretend play for young children with ASC, compared with a non-computer setup. The experiment consists of two conditions: AR and non-AR. The order of the two conditions is counterbalanced among subjects. There is a short break between the two conditions. In each condition, there are three tasks and the order is randomized. The null hypotheses of the experiment are:

H_{0A}: There is no significant difference in the frequency of pretend play between the AR and non-AR conditions.

H_{0B}: There is no significant difference in the duration of pretend play between the AR and non-AR conditions.

The design of the experiment is largely informed by the rich literature of empirical studies in autism research ([4] [7] [11] [21] [22] [30] [35]). Previous research divides the level of prompt of pretend play into two categories: elicited and instructed. In the elicited prompt scenario, the experimenter encourages the participant to play with available props, without giving specific pretense ideas. An example prompt is "*Show me what you can do with these*". In the instructed play scenario, the experimenter makes verbal or physical prompt by asking the participant to perform/mimic specific actions, such as "*park the car (toy) in the garage (shoebox)*". Research has indicated that non-specific elicitations increase pretend play of children with ASC to some extent while specific instructions prompt them to produce as many pretend play episodes as children in the control group [21]. Instructed prompts, however, always require caution in interpreting a child's behavior as pretend play because the child may just make an "intelligent guess" [6] when asked to carry out certain pretend play actions with limited available props. We adopt the elicited prompt strategy in the design of our experiment, because the AR system is intended to encourage open-ended pretend play without detailed instruction of play actions.

5.1 Participants

Twelve children formally diagnosed with ASC or Asperger Syndrome aged 4-7 participated in the study, 10 male and 2 female. Participants are recruited via the Cambridge Autism Research Center parent mail-list, the newsletter of the Cambridge branch of National Autistic Society, and local autism events. The experiment is approved by the University of Cambridge Ethics Committee. All participants were remunerated with an age appropriate educational gift.

We visited participants' homes prior to the experiment to collect information about their autism and language conditions. We use the Childhood Autism Rating Scale, 2nd edition (CARS2) based on parent interviews and direct observation to inform participants' autism severity. We also evaluate their verbal mental age using the British Picture Vocabulary Scale, 3rd edition (BPVS3) since research indicates that pretend play of children with ASC is closely correlated with their language comprehension [35]. Table 2 shows the participants' information.

Table 2. The summary of participants' information

	Chronological Age (months)	Verbal Mental Age (months)	Autism Severity
Mean	82	73*	33.3 (mild-to-moderate)
SD	11.09	17.82	6.34
Range	53 - 93	45 - 104	22.5 - 41.5

* One participant was not able to complete the BPVS3 test

Based on the CARS2 parent interview, the level for "object use in play" among participants is between mildly and moderately inappropriate (except one participant who was reported as age appropriate). The levels of pretend play frequency at home are frequent (3 participants), sometimes (4 participants), seldom (4 participants) and never (1 participant). Nine participants attend mainstream primary and reception class with special assistance. One was being home educated when the study took place. One attends a special school and the other one is in a special class for autism and learning difficulties affiliated with a mainstream school. All participants are familiar with computer devices. Most of them use computers on a daily basis.

5.2 Apparatus

The apparatus in the AR condition contains a 24-inch monitor, a Logitech webcam Pro 9000 (field of view 75 degrees), a mini Bluetooth keyboard, a table (45*90*45cm), and play materials. There are two types of play materials including AR objects (three foam blocks and a cardboard box with markers attached) and a set of non-AR physical props (three cotton balls, two paper tubes, three popsicle sticks, three pen tops, three strings and a piece of cloth). The detailed description of the setup is shown in Figure 5.

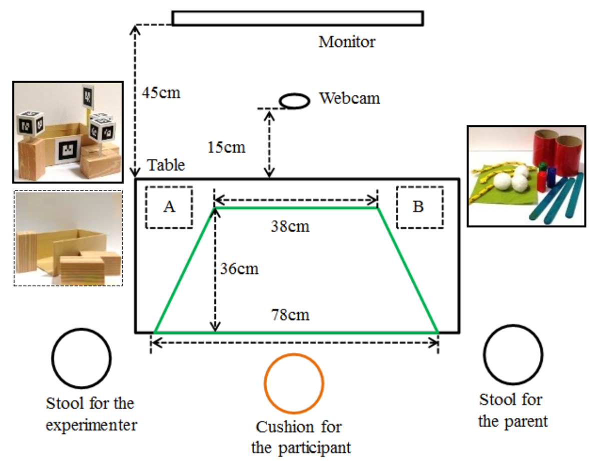


Figure 5: The physical setup of AR and non-AR conditions.

The AR objects are located in area "A" and the non-AR props are located in area "B". In addition, we taped out a trapezoidal area on the table to emphasize the range of the camera view. The computer connected to the monitor and webcam is located in another room next door to avoid potential distraction to the participants. The non-AR setup contains the same table and physical props, plus blocks and a box of the same dimensions but without markers. In both conditions we asked the participant to play within the taped area.

Play materials commonly used in previously mentioned studies include conventional toys (e.g. toy vehicles or human figures) and non-toy objects (as non-AR props in our experiment). The difference between these two types is that conventional toys facilitate functional play (which refers to actions that are appropriate use of toys according to its designed function) while non-toy objects encourage nonliteral use in pretend play. In our AR condition, the child can see that s/he is in reality holding a block, and the visual augmentation reinforces non-literal mental representation of pretense (e.g. a car) to the block. On the contrary, there is no such non-literal mental representation involved in functional play with conventional toys, thus it is out of the scope of our research question. Additionally, research shows that functional play with conventional toys occupies a large amount of time during children's elicited play sessions. In order to maximize pretend play within a short period of time while keeping the play props comparable between the AR and non-AR conditions, we did not include conventional toys as experiment materials.

5.3 Procedure

The main experiment procedures and scripts are consistent in both the AR and non-AR conditions as listed below:

1. A brief introduction: the experimenter reminds the participant to "*play inside of the taped area*", "*play with anything s/he likes on the table*" and "*stop after 5 minutes*".
2. Initialize the play materials on the table.
3. Start the task: the experimenter holds one block and asks: "*show me how you can play with this block as a car/train/airplane*", then gives the block to the participant.
4. During the task:
   1) The experimenter shouldn't give any detailed pretend play ideas.
   2) If the participant doesn't attend to playing, the experimenter should encourage the child by saying: "*I want to see more how you can play with the block as a car/train/airplane. Let's try some more*".
   3) If the participant doesn't use any of the physical props, the experimenter should encourage by saying: "*You can play with anything you like on the table*".
   4) The experimenter can prompt a maximum of 3 times. After that the experimenter should ask: "*Do you want to continue with the play or change to another one*".
5. After 5 minutes, the experimenter should wait until the participant finishes with the current play episode and say "*Very good. Now let's stop and put everything back*".
6. Repeat steps 2- 5 for the other two tasks.
7. Ask the participant for feedback at the end of each condition.
8. Interview the parent when both conditions have finished.

In addition to the procedure above, in the AR condition, the experimenter will let the participant try out the AR system in a demonstration mode prior to the actual task. During each task, at around 3 minutes, the experimenter says "*Watch, something will be on the screen*", and then reveals the extra imaginary content (bridge/track/runway) on the screen. At around 4 minutes, the experimenter says "*Watch, something else will be on the screen*", and then switches the imaginary content (dusty effect/crane/fire).

5.4 Data collection

5.4.1 Video Analysis

We analyzed participants' play behavior based on the video footage recorded during the experiment. We set two video cameras in the experiment, one in front of the participant to record the non-AR session, and one in front of the computer screen in the separate room to record the AR session.

We have reviewed the literature of play coding schemes in both general developmental psychology and pretend play with autism. Piaget [26] proposed three developmental stages of play, namely practice, symbolic and rule-based. Based on Piaget's original proposal, Smilansky [33] further developed a play category including functional play, constructive play, dramatic play and game-with-rules. Smilansky's classification has been adopted by the well-established Play Observation Scale (POS) [31] for play behavior analysis. The coding scheme commonly used in autism research (e.g. [4] [11] [21] [35]) includes pretend play, functional play, relational play, simple manipulation/sensorimotor play, and no play. We exclude functional play because no conventional toy is included in the play material as discussed earlier.

In view of the above considerations, we designed a coding scheme that includes five play categories: pretend play, constructive play, relational play, simple play and no play. The definition and examples of each category are listed below:

1. **Pretend Play*** (PP):
   Play actions that are either vehicle appropriate or novel, and involve any of the following features:
   1) Object Substitution: use one thing as something else (e.g. push the block along the table and make the sound "*choo choo*").
   2) Attribution of pretend properties: assign false or absent properties to an object (e.g. make one block talk to another block).
   3) Imaginary Object: imagine the presence of something invisible (e.g. use imaginary water to put out the fire).
2. **Constructive Play** ** (CP):
   Play actions that involve creating an object or a scene with more than one object (e.g. use tube and block to build a train).
3. **Relational Play** (RP)
   The participant manipulates more than one object or a single object in relation to others (e.g. combination, stacking, containing and arranging), but not attending to creating something or pretending something meaningful.
4. **Simple Play** (SP)
   The participant attends to manipulating one object without purposeful meaning (e.g. moving, waving, banging, fingering, mouthing or throwing of a single object).
5. **No Play** (NP)
   Other actions that are not play related.

* The rater codes the action as PP when there is a strong visual, verbal or vocal cue to confirm the pretense, or the rater is quite sure that there is some imaginative feature involved.

** The rater codes the action as CP when the participant is in the process of creating something. If the participant then manipulates the created thing in the way as what it's meant to be, the rater codes the follow up action as PP (e.g. use the monster s/he built to destroy a tower).

We used the video editing tool Camtasia to manually annotate: (1) participant's discrete play actions relating to the play materials; (2) participant's verbal and vocal utterances; (3) experimenter's and parent's talk during the experiment. The first rater (experimenter) coded each action according to the coding scheme. We then invited an independent rater who was not aware of the hypotheses to code 10 out of total 60 video clips (randomly chosen, 5 from each condition) to verify the reliability of the coding scheme used. The inter-rater agreement was satisfactory (Cohen's kappa = 0.75).

5.4.2 Questionnaires

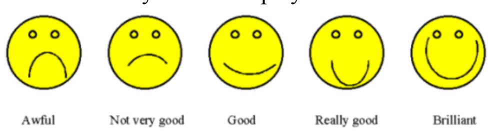
We used both parent questionnaire and participant questionnaire to collect qualitative feedback to evaluate the emotional quality of the participant's involvement in each condition. Given the diverse degree of behavioral disturbance of individuals with ASC, it is considered more reliable to have parents rate for engagement rather than the experimenter. Therefore, we asked each parent to observe the participant playing and rate for his/her engagement in terms of cooperativeness, attentiveness and happy smiling [30] immediately after each experiment session. We also ask the parents to provide overall feedback of the experiment after both sessions were completed. A summary of questions in the parent questionnaire are listed below:

1. Cooperativeness or in-seat behavior
   (Very Good, Good, OK, Poor, Very Poor)

2. Interest or general attentiveness to the play things (Very Good, Good, OK, Poor, Very Poor)
3. Happy smiling involved in play (Frequent, Sometimes, Seldom, Never)
4. Which session do you think the participant enjoyed more? (First session, Second session, Equal, Not sure)
5. In which session do you think the participant was more engaged? (First session, Second session, Equal, Not sure)
6. Do you think the technology will help to promote pretend play for young children with ASC? (Strongly agree, Agree, Neutral, Disagree, Strongly disagree)
7. Do you think the play themes (car, train and airplane) are appropriate? (Strongly agree, Agree, Neutral, Disagree, Strongly disagree)
8. Can you name other play themes in the participant's daily play repertoire?
9. Anything you think could be improved for the computer program?

In addition, we asked the participants questions about their play experience and preference. We included the Fun Toolkit [28], which is a well-known survey method for young children, as part of the questions. The detailed questions are listed below:

1. How much do you like the play?



2. One thing you like about the play?
3. One thing you don't like about the play?
4. (AR condition only) Are there other things you want to be on the screen?
5. Which play is more fun (the one with/without screen)? And why?
6. Which one do you prefer to play with your friend (the one with/without screen)?

6 EXPERIMENT RESULTS

The results of all participants except two are reported in this section. We excluded these two participants' data from the main results because they had difficulty following the experimenter's instructions and were not capable of cooperating during the experiment due to severe impairment in language and joint attention. We will discuss their behaviors separately in the discussion section.

6.1 Play Frequency

The action frequency of each play category among participants is normally distributed. The distributions of action frequency (occurrences per minute) in each play category according to the coding scheme are shown in Figure 6. We can see that the mean frequency of pretend play is higher in the AR condition, while the mean frequency of constructive play is higher in the non-AR condition. The figure also shows that the level of relational play, simple play and no play remains similar in both conditions. We conducted a paired *t*-test evaluation and found there is a significant difference in pretend play ($t(9) = 4.66$, $p < 0.01$) and constructive play ($t(9) = -4.91$, $p < 0.01$).

To explore one indicator of the quality of pretend actions produced in both conditions, we further excluded pretend play actions with repeated play ideas. The result shows that there is still a significantly higher frequency of pretend play (t (9) = 2.41, $p < 0.05$) produced in the AR condition (mean = 1.79, SD = 0.68) than the non-AR condition (mean = 1.23, SD = 0.63).

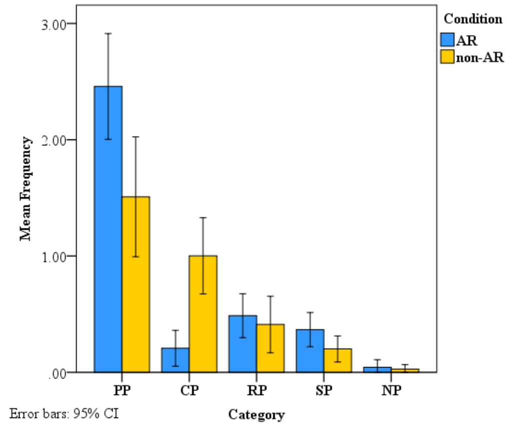


Figure 6: Play frequency (occurrences per minute) in each category.

6.2 Play Duration

The percentage of time spent in each type of play is illustrated in Figure 7. As with the play frequency results, the percentage of time that participants spent in pretend play is significantly higher ($t(9) = 3.25$, $p < 0.01$) in the AR condition, while the percentage of time in constructive play is significantly higher ($t(9) = -3.49$, $p < 0.01$) in the non-AR condition. The differences among relational play, simple play and no play between the two conditions remain non-significant.

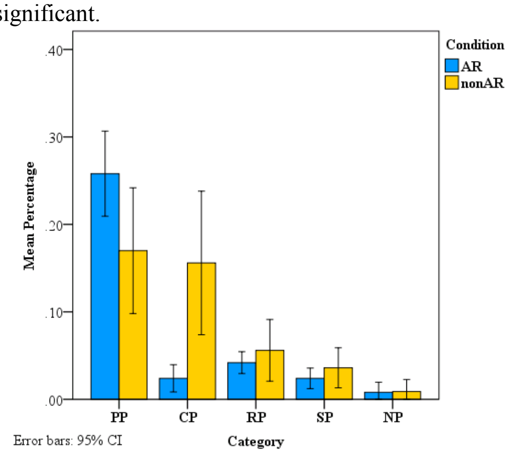


Figure 7: Percentage of play time in each category.

6.3 Engagement and Enjoyment

Figure 8 shows that the mean scores of attentiveness and cooperativeness are between ok and high in both conditions, while the appearance of happy smiling for the children is between sometimes to frequent. There is a marginally significant difference in happy smiling ($z = -1.90$, *Asymp. Sig* = 0.058) using the nonparametric Wilcoxon test. According to the parent questionnaire, eight out of ten parents thought their children were more engaged in the AR condition. One parent thought the participant was equally engaged in both conditions and one thought the participant was more engaged in the non-AR condition. Moreover, we counted how often the experimenter gave verbal prompts ("show me how to play with the block as a …") to encourage the participant to carry on with playing. The experimenter made significantly more verbal prompts ($z = -2.61$, *Asymp. Sig* < 0.01) in the non-AR condition (mean = 0.44) than the AR condition (mean = 0.26) according to the nonparametric Wilcoxon test.

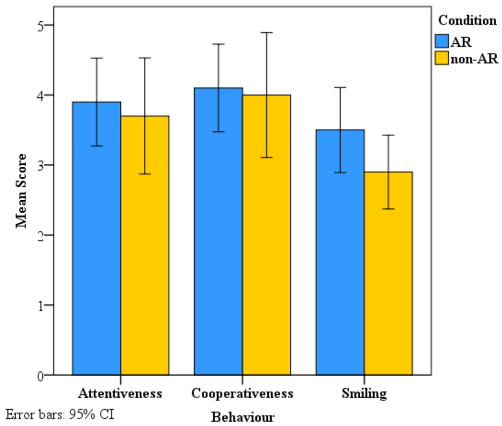


Figure 8: The boxplot of mean score of engagement.

The parents' feedback about participants' enjoyment is aligned with their engagement. The same eight parents thought their children enjoyed the play more in the AR condition. The average score for enjoyment is good in the AR condition and really good in the non-AR condition according to the participants' feedback. Comments about things participants like and dislike in each condition are summarized in Table 3:

Table 3. The summary of participants' feedback of like and dislike of each condition.

Condition	Feedback	
	Like	Dislike
AR	"a different picture on the block" "when the car gets rusty" "car get into a school" "flying" "they all change into different things"	"the train push a lot of things off the table" "no police car"
Non-AR	"when the goodies win" "the party" "make a lot of things using the blocks" "can rescue the car in a box" "shop keeper" "make the airplane crash"	"the baddies broke the plane" "there is no police car, no train station" "didn't know what to do" "the toy doesn't have eyes"

In addition, when asked which one is more fun, nine out of ten participants chose the AR system and indicated that they would prefer the AR system to the non-AR system for play with friends. Reasons explained by the participants include: "The blocks become into different things", "It has a picture", "can see things that is not actually there", "I like seeing myself", and "It's funny".

6.4 Pretend Play Theme

During the experiment, we noticed that the themes of pretend play varied between AR and non-AR conditions even though the participants were asked to carry out the same vehicle theme at the beginning and during each task. To investigate the difference of attending to the vehicle theme indicated by the experimenter, as well as details of play ideas in terms of realistic and novel among those complying with the vehicle theme, we further categorized pretend play actions into three types:

1. **Relevant_Reality**: Actions that approximate realistic behavior of the vehicle which are situationally appropriate.
2. **Relevant_Novel**: Actions that involve the vehicle but are novel instead of realistic.
3. **Not relevant**: Actions that do not involve the vehicle theme indicated by the experimenter.

Table 4 shows examples of representative play ideas:

Table 4. The summary of participants' play ideas in terms of relevance to the suggested vehicle theme.

Relevant		Not relevant
Realistic	Novel	
move the car along the table; move the train into the train station; make the airplane take off from the runway; point the stick at the dusty car and say "water"; put a cotton ball on the train and say "driver"; tap the finger around the car and say "fix the car"; move the train over a stick and say "train track"	move the car in the air and say "climb a tree"; make cotton balls hit the cars and say "angry bird"; make the car "go through" a tube and say "in the black hole"; tap a string around the train and say "poison the train driver"; point the two airplanes at each other and say "how's going"	party; spaceship fight; shopkeeper; make the "monster" step on things; move a string on the table and say "snake"; remove all physical props from the table and say "set off"

Figure 9 shows the percentage of pretend play actions in each play theme category. The mean percentages of total relevant actions including both reality-based and novel-based is significantly higher in the AR condition according to the paired *t*-test ($t(9) = 2.84$, $p < 0.05$). The inter-rater agreement of two raters is satisfactory (Cohen's kappa = 0.85).

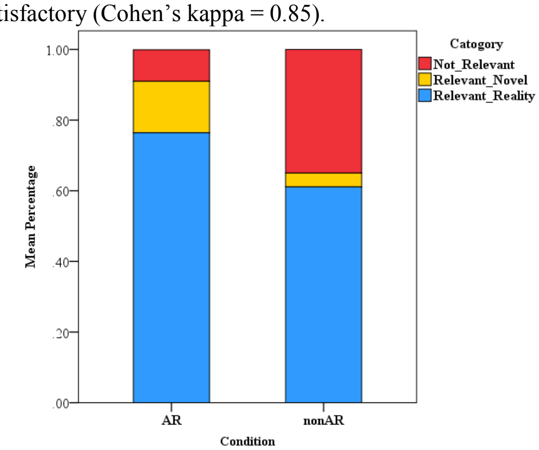


Figure 9: The percentage of actions in terms of relevance.

7 Discussion

7.1 Effectiveness

7.1.1 Quantity of pretend play

The experiment results support that children with ASC can carry out pretend play actions under elicited prompts, which complies with the existing psychology literature. In particular, the results reject both null hypotheses and demonstrate that there is a significantly higher frequency and duration of pretend play in the AR condition than the non-AR condition. This indicates a positive effect of using the AR system to promote elicited pretend play for young children with ASC.

Results also show that pretend play actions surpass other types of play in both conditions. This is largely due to the non-functional feature of the experiment materials and verbal elicitations from the experimenter. On the other hand, participants tend to produce more constructive play in the non-AR condition. This can also be related to the non-functional feature of the play materials, which supports building things up in addition to pretending. Meanwhile in the AR condition, the salient visual indications given by the AR system are more persuasive for pretend play than constructive play.

7.1.2 Quality of pretend play

We analyzed the detailed play ideas to further investigate the quality of pretend play produced during the experiment. First, there is a noticeable difference in how participants followed the elicited play theme in each condition. As shown in Figure 9, in the AR condition pretend play actions carried out by the participants were highly relevant to the vehicle theme of each task indicated by the experimenter. In the non-AR condition, participants tended to carry out less relevant themes. In some extreme cases, the participants ignored the experimenter's suggestions and carried out irrelevant themes (e.g. spaceship, party, shopkeeper, etc.) consistently across tasks. Such intense and inflexible play preference is largely due to a lack of mental flexibility [17] and restricted play interest. It is, therefore, difficult for normally developed peers to join pretend play with an autistic child, where plots are often strictly copied from things seen in movies, games or on TV, as many parents reported in this study. The inclination of following play themes with the AR system can make it an ideal platform to support and regulate shared pretend play among autistic children and their parents or peers. Second, participants produced diverse pretend play ideas relating to the indicated vehicle theme. In particular, the proportion of novel ideas to situationally appropriate ideas is relatively high in the AR condition. This is a very interesting result because children with ASC are constantly found to lack novel pretend acts compared with matched control groups [11] [21] due to general executive deficit in generating novel ideas. One potential explanation of this is that the AR system visually externalizes some internal representation of the suggested pretending theme. As a result participants have "higher bandwidth" in the working memory to get access to relevant internal representations needed to extend the play with novel ideas. One comment from a participant about the AR system is that "… (I) can remember what I might need".

7.2 Engagement

The participants' engagement is relatively high in both setups, which could be related to the structured nature of the experiment. Parents' feedback shows that participants are more engaged in the AR condition. Although we should be cautious interpreting such feedback, the significantly lower frequency of experimenter's verbal prompting also indicates that participants are more engaged in the AR than the non-AR condition. Although the system demonstration session at the beginning of the AR condition is meant to reduce the novelty effect of the AR technology, it is still likely to be one of the motivational factors. This is further indicated by the result that the majority of participants thought the AR session is more fun and preferred to play it with their friends.

7.3 Insight from the excluded cases

We have excluded the results of two participants because their engagement for the experiment was too low to be valid. Both of the participants have severe impairment in language, joint attention and object use in play. The first participant managed to attend to and manipulate the AR objects in some vehicle appropriate ways in all three tasks in the AR condition, but produced fewer actions in total compared with other participants and spent most of the time engaged in simple play or no play, such as banging blocks on the table, wandering around the room, and lying on the floor. In the non-AR condition, the participant did similar simple and non-play actions except without any pretend play in spite of constant direction from the experimenter and the parent.

The second participant could not finish the BPVS3 test due to severe impairment in joint attention. In the AR session, the participant spent most of the time watching the self-image on the screen. When the participant manipulated the AR object, it was rather immature including mouthing, banging and ordering. Other inappropriate behaviors included pulling the marker cube from the block and breaking the stick used to connect the marker cube to the block. In the non-AR session, the participant paid little attention to objects on the table and ran away several times out of the experiment room. The observation shows that the AR system has at least some positive effect in encouraging pretend play for the participant in spite of the severe language delay and poor joint attention. It also shows the challenges of designing computer systems for children near the lower end of the autism spectrum.

7.4 Design Reflections

7.4.1 Re-examine usability

The experiment demonstrates that participants can successfully interact with the AR system, even though seven out of ten participants were reported to have poor fine motor skills including eye-hand coordination tasks (e.g. handwriting). Some usability issues perceived during the experiment include:

1. **Hand over marker**: Participants were told to hold the block instead of the marker cube when manipulating the AR object during the familiarisation session. Even under physical prompting, some participants persevered with holding the marker cube which caused the virtual object to flicker.
2. **Inward/Outward orientation**: As discussed in the pilot study, most participants have to spend extra time exploring the spatial relationship in order to align objects properly with virtual objects registered on the table (e.g. a virtual bridge) due to the inward/outward reversal caused by the mirrored view, similar as discussed in [27].
3. **Limited size of play area**: The taped play area can be very crowded when it is occupied with AR and non-AR objects.

These usability issues are trade-off results to acquire high tracking accuracy/extensibility, avoid occlusion of placing a marker on the table and high tracking reliability respectively. For the second issue, we may improve the situation by drawing a physical marker (e.g. a dot) on the table to indicate the center location of the virtual objects, and include virtual objects registered on the table surface in the familiarisation session.

7.4.2 Adapt to the real world

The current AR system is designed as an experimental apparatus. Therefore, it only provides a small set of AR props and three fixed play themes. Although preliminary results show a positive effect compared with the non-AR condition, further improvements are needed in order to scaffold pretend play development beyond the laboratory setup. Some improvements are proposed below based on direct observation and parent/participant feedback:

Provide more AR props: Most children with ASC have a very restricted range of play interests. Therefore the availability of their desired play theme can be an important motivation. The participant and parent interview provide a rich set of AR props potentially to be included, such as: superheroes, dinosaur, people, baby, police car/office, ship, animal, emergency vehicle, and characters from popular film/game/TV program.

Assign augmentation to the AR object dynamically: It can encourage children not only to develop ideas about how to pretend, but also proactively choose what to pretend.

Fade out visual effect: In order to gradually bridge the pretend play experience from the AR system to real life scenarios, a fading out mechanism can be implemented for the visual effect, which is based on the most-to-least prompt strategy commonly used in applied behavioral approach.

Enable the user to record the play: Recording is a common feature for storytelling systems (e.g. [32]). Several participants in the experiment, in their spare time browse online videos for game demonstrations (e.g. Minecraft, Super Mario) and two participants particularly mentioned that they would like to share the play they made in the AR session online with other children.

7.4.3 Reflections on system and experiment design

Unlike other AR systems discussed earlier, the AR system we proposed encourages open-ended play. Such open-ended design aims to direct the children to proactively think of divergent play ideas, instead of focusing on exploring specific action-effect rules. In the action-effect style, the play motivation may decline quickly once the user explores all effects. Motivation also remains a challenging issue in the open-ended style when there is minimum adult guidance. Therefore alternative motivational methods should be considered in the future study, such as to enable recording and sharing the play carried out with the AR system.

Based on the current study, we suggest a two-tier approach when designing AR systems and experiments for children with pervasive developmental disorders like ASC.

Design for the target deficit: A thorough literature review on the target developmental disorder is required during the early design phase. Investigating theoretical explanations of the deficit provides critical reference about the potential positive effect of using AR technology. Although the novelty of AR technology is a motivational merit, it is expected to fade out over time. It is therefore advisable to focus on how the unique nature of AR technology can best eliminate the specific deficit from a mechanism level. In addition, both theory and intervention methods provide extensive references for system design and evaluation approach. One caution for knowledge derived from the research corpus is that most results are acquired from controlled experiments and that some well-established studies were conducted decades ago. Such knowledge unavoidably causes deviation from autistic children's behaviors in today's world. One example is that we found many participants in our experiment prefer rule-based video games than traditional toys with more open-ended orientation.

Consideration for general autistic characteristics: While the effectiveness of an AR system is largely determined by its design to reinforce the target cognition/behavior, the usefulness of the system is a prerequisite to achieve that. We summarized a list of pervasive disorders that must be considered when designing AR systems and experiments for children with ASC:

1) **Language delay and joint attention**: it is difficult to explain things like how to interact with the AR system and what one is expected to do during the experiment, to children whose language and joint attention are severely impaired. Depending on the goal of the AR system, whether exploration-oriented or task-oriented, the researcher/designer has to carefully determine the developmental level of the target group.
2) **Restricted interest**: It is difficult to persuade children with autism to take part in activities that they are not interested in. Besides the novelty of AR technology, contents of the AR system have to be appealing in order to keep the children engaged for any length of time, which ideally requires a selection of available AR objects of different types, shape, color, etc. An extreme case during our experiment is that one of the participants got frustrated while manipulating the augmented airplane simply commenting "this is not the airplane I want".
3) **Resistance to change**: most children with ASC follow a strict routine for daily activities and become anxious easily when new activity is introduced such as visiting an unfamiliar environment. In order to eliminate withdrawal caused by the above reasons, we decided to meet the participants before the study by running a home interview, so that the participants can get familiar with the experimenter ahead of time. We also prepared visual guidance (with photos of the experiment location, the setup of the AR system and gifts) to help the parents describe the study to their children. Some parents found this very helpful to reduce stress of the participants and increase their cooperativeness during the study.

7.5 Study Limitation and Future Work

Besides technical restraints discussed in the previous section, there are several study limitations that we expect to explore in future. First, although the preliminary study results demonstrate a positive effect of elicited pretend play with the AR system, the potential generalization effect to improve spontaneous pretend play in real life is still to be examined with systematic intervention methods over a longer period of time. During the AR condition, we observed the participants carry out some unique object substitution actions with non-augmented objects, which did not occur in the non-AR condition. This shows that even in a short period of exposure, participants tend to extend object substitution beyond what the system suggested. We consider this as an indication of potential generalization effect of the AR system which remains to be explored in the future study.

Second, although the familiarization session at the beginning of the AR condition helps participants explore the mechanism of AR, technology novelty may still be one of the motivational factors during our study. Thus engagement of pretend play over longer system exposure is another topic to be probed in the further study

Third, the experiment did not include any conventional toys as play materials (e.g. toy vehicles) intentionally to avoid functional play, thus maximizing the occurrence of pretend play. As a result, the positive effect shown in this study should be interpreted cautiously with potential interference with functional play in a natural play environment with conventional toys.

Fourth, the current study only examined using the AR system for solitary pretend play. Development of shared pretend play involving social context is an important next step. As discussed in the previous section, the AR system naturally supports multi-user interaction and potentially directs mutual play themes among users.

Fifth, the current vehicle play themes are more appealing to boys and most of the participants signed up for the study are boys. Therefore the outcome of the study has a potential gender bias. More girl-friendly and mutual themed AR objects should be added and the effect with girls should be explored accordingly.

8 Conclusion

We presented the design and evaluation of an AR system aiming to promote open-ended pretend play for young children with ASC who have limited language delay. Results indicate a positive effect of increased elicited pretend play in both frequency and duration with the AR system compared with a non-computer setup. Participants were highly engaged with the AR system and produced a diverse range of play ideas. The salient visual effect of the AR system potentially helps participants to focus on carrying out pretend play actions relevant to suggested play themes, rather than persisting with restricted themes regardless of what they have been encouraged to do, as some of the participants did in the non-AR session.

We discussed usability issues of the AR system and possible improvements. Limitations of the study, including not examining the generalization effect to real life and long-term engagement, only evaluating pretend play in the solitary scenario, exclusion of conventional toys and potential gender bias, are to be addressed by future work.

We summarized guidance for designing and evaluating AR systems for children with ASC and general pervasive developmental disorders based on existing literature. Moreover, our study demonstrated a procedural approach to exploring the potential of AR technology in stimulating specific cognitive activities like pretend play for challenging user groups like young children with ASC.

ACKNOWLEDGEMENT

We would like to thank psychology and technology experts, including Dr. Nicola Yuill and colleagues in the ChatLab, Dr. William Farr, Dr. Joseph Newman, Dr. Elisabeth Hill, Dr. Judith Good and local autism event organizers for giving invaluable feedback as the study progresses. We also would like to thank the Cambridge Autism Research Center and National Autistic Society Cambridge branch for helping advertise the recruitment. In particular, we would like to thank all participants' families for taking part in the study and the two independent raters who kindly helped with the video coding.

REFERENCES

[1] American Psychiatric Association DSM-5 development: http://www.dsm5.org/

[2] Bai, Z., Blackwell, A. F. and Coulouris, G. Making Pretense Visible and Graspable: An augmented reality approach to promote pretend play. *In Proc. ISMAR 2012*, 267- 268.

[3] Bai, Z., Blackwell, A. F. and Coulouris, G. Can we augment reality with "mental images" to elicit pretend play?: a usability study. *In Proc. CHI Extended Abstract 2013*, 1-6.

[4] Baron-Cohen, S. Autism and Symbolic Play, *British Journal of Developmental Psychology* 5 (1987), 139-148.

[5] Baron-Cohen, S. and Wheelwright, S. 'Obsessions' in children with autism or Asperger Syndrome: a content analysis in terms of core domains of cognition. *British Journal of Psychiatry* 175 (1999), 484-490.

[6] Baron-Cohen, S. Instructed and elicited play in autism: A reply to Lewis & Boucher. *British Journal of Developmental Psychology* 8 (1990), 207.

[7] Barton, E. E. and Wolery, M. Teaching Pretend Play to Children With Disabilities: A Review of the Literature. *Topics in Early Childhood Special Education* 28, 2 (2008), 109-125.

[8] Bergen, D. The Role of Pretend Play in Children's Cognitive Development. *Early Childhood Research & Practice* 4, 1(2002).

[9] Brederode, B., Markopoulos, P., Gielen, M., Vermeeren, A. and Ridder, H. D. pOwerball: The design of a novel mixed-reality game for children with mixed abilities. *In Proc. IDC 2005*, 32-39.

[10] Casas, X., Herrera, G., Coma, I. and Fernández, M. A Kinect-based Augmented Reality System for Individuals with Autism Spectrum Disorders. *GRAPP/IVAPP 2012*, 440-446.

[11] Charman, T. and Baron-Cohen, S. Brief report: prompted pretend play in autism. *Journal of Autism and Developmental Disorders* 27, 3 (1997), 325-332.

[12] Corrêa, A. G. D., Ficheman, I. K., Nascimento M. D. and Lopes, R. D. D. Computer Assisted Music Therapy: a Case Study of an Augmented Reality Musical System for Children with Cerebral Palsy Rehabilitation. *In Proc. ICALT 2009*, 218-220.

[13] Escobedo, L., Nguyen, D. H., Boyd, L., Hirano, S., Rangel, A., Garcia-Rosas, D., Tentori, M. and Hayes, G. MOSOCO: A mobile assistive tool to support children with autism practicing social skills in real-life situations. *In Proc. CHI 2012*, 2589-2598.

[14] Farr, W., Yuill, N. and Hinske, S. An augmented toy and social interaction in children with autism. *In Proc. IDC 2010*, 104-125.

[15] Fiala M. Magic Mirror System with Hand-held and Wearable Augmentations. *In Proc. VR 2007*, 251-254.

[16] Herrera, G., Alcantud, F., Jordan, R., Blanquer, A., Labajo, G. and Pablo C.D. Development of symbolic play through the use of virtual reality tools in children with autistic spectrum disorders. *Autism* 12, 2 (2008), 143-157.

[17] Hill, E. L. Executive dysfunction in autism. *TRENDS in Cognitive Sciences* 8, 1 (2004), 26-32.

[18] Hong, H., Jeong, H. Y., Arriaga, R. I. and Abowd, G. D. TriggerHunter: Designing an Educational Game for families with Asthmatic children. *In Proc. CHI 2010*, 3577 – 3582.

[19] http://goblinxna.codeplex.com/

[20] http://virtual.vtt.fi/virtual/proj2/multimedia/alvar

[21] Jarrold, C., Boucher, J. and Smith, P.K. Generativity Deficits in Pretend Play in Autism. *British Journal of Developmental Psychology* 14 (1996), 275–300.

[22] Jarrold, C. A Review of Research into Pretend Play in Autism. *Autism* 7 (2003), 379-390.

[23] Leslie, A.M. Pretense and Representation: The Origins of "Theory of Mind''. *Psychological Review* 94 *(*1987*)*, 412–26.

[24] Luckett, T., Bundy, A. and Roberts, J. Do behavioural approaches teach children with autism to play or are they pretending? *Autism* 11, 4 (2007), 365-388.

[25] Oda, H., Lister, L. J. White, S. and Feiner, S. Developing an Augmented Reality Racing Game. *In Proc. INTETAIN 2008*, Article No. 2.

[26] Piaget, J. Play, dreams, and imitation in childhood. New York: Norton, 1962.

[27] Radu, I. and MacIntyre, B. Using Children's Developmental Psychology to Guide Augmented-Reality Design and Usability. *In Proc. ISMAR 2012*, 227- 236.

[28] Read, J. C. and MacFarlane, S. Using the fun toolkit and other survey methods to gather opinions in child computer interaction. *In Proc. IDC 2006*, 81 – 88.

[29] Richard E., Billaudeau V., Richard P. and Gaudin, G. Augmented Reality for rehabilitation of cognitive disabled children: a preliminary study. In Proc. *Virtual Rehabilitation 2007*, 102–108.

[30] Riguet, C. B., Taylor, N. D., Benaroya, S. and Klein, L. S. Symbolic play in autistic, Down's, and normal children of equivalent mental age. *Journal of Autism and Developmental Disorders* 11, 4 (1982), 439-448.

[31] Rubin, K. H. The play observation scale (POS). University of Waterloo (1989).

[32] Ryokai, K. and Cassell, J. StoryMat: a play space for collaborative storytelling. *In Proc. CHI EA 1999*, 272-273.

[33] Smilansky, S. The effects of sociodramatic play on disadvantaged. Preschool children. New York: Wiley (1968).

[34] Tartaro, A. and Cassell, J. Authorable virtual peers for autism spectrum disorders. *In Combined Workshop on ECAI 2006*.

[35] Ungerer, J. A. and Sigman, M. Symbolic Play and Language Comprehension in Autistic Children. *Journal of the American Academy of Child Psychiatry* 20, 2 (1981), 318–337.