Augmented Reality-Based Procedural Task Training Application for Less Privileged Children and Autistic Individuals

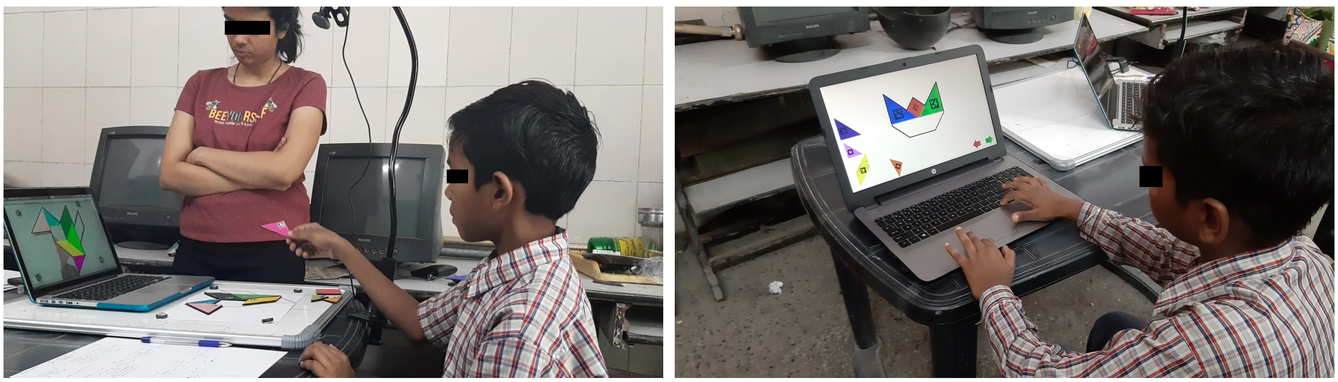
Katyayani Singh
IIIT Delhi
katyayani15046@iiitd.ac.in

Ayushi Srivastava
IIIT Delhi
Ayushi16025@iiitd.ac.in

Krishnaveni Achary
Tamana NGO
krishnaveni.achary@gmail.com

Arindam Dey
University of Queensland
a.dey@uq.edu.au

Ojaswa Sharma
IIIT Delhi
ojaswa@iiitd.ac.in



(a) AR-based instruction

(b) Desktop-based instruction

Figure 1: We compared an Augmented Reality-based instruction for procedural tasks with a desktop-based and live demonstration-based instructions in a less privileged population.

ABSTRACT

In this work, we evaluate the applicability of using Augmented Reality applications in for enhanced learning experiences for children from less privileged backgrounds, with a focus on autistic population. Such an intervention can prove to be very useful to children with reduced cognitive development. In our evaluation, we compare the AR mode of instruction for a procedural task training, using tangram puzzles, with live demonstration and a desktop-based application. First, we performed a within-subjects user study on neurotypical children in the age group of 9 - 12 years. We asked the children to independently solve a tangram puzzle after being trained through different modes of instruction. Second, we used the same instruction modes to train autistic participants. Our findings indicate that during training, children took the longest time to interact with Desktop-based instruction, and took the shortest time to interact with the live demonstration mode. Children also took the longest time to independently solve the tangram puzzle in the Desktop mode. We also found that autistic participants could use AR-based instructions but required more time to go through the training.

CCS CONCEPTS

• **Human-centered computing** → **User studies**; **Mixed / augmented reality**.

KEYWORDS

augmented reality, autism, training, instruction, children, less privileged

ACM Reference Format:
Katyayani Singh, Ayushi Srivastava, Krishnaveni Achary, Arindam Dey, and Ojaswa Sharma. 2019. Augmented Reality-Based Procedural Task Training Application for Less Privileged Children and Autistic Individuals. In *The 17th International Conference on Virtual-Reality Continuum and its Applications in Industry (VRCAI '19), November 14–16, 2019, Brisbane, QLD, Australia.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3359997.3365703

VRCAI '19, November 14–16, 2019, Brisbane, QLD, Australia

ACM ISBN 978-1-4503-7002-8/19/11...$15.00
https://doi.org/10.1145/3359997.3365703

1 INTRODUCTION

Autism Spectrum Disorder (ASD) is characterized by the deficit of a few core skills including social communication and interaction. Genetic findings suggest that higher-order association areas of the brain that normally connect to the frontal lobe are partially disconnected during development for people with ASD [Geschwind and Levitt 2007]. There is no known cure for autism, however regular behavior therapy and training programs seem to help. There are different therapies and interventions used by physicians and caregivers to assist people with ASD. Among these interventions there are efforts that used augmented reality (AR) and virtual reality (VR) technologies.

Recent advances in technology like AR and VR provide immersive and informative environments conducive for focused learning. Lee [2012] provides an overview of the use of these technological aids in designing learning systems. Augmented reality has been increasingly used in both education and business settings. Further, educational uses of virtual and mixed reality have been extensively explored by Billinghurst [2002]. Pan et al. [2006] demonstrate by means of several examples that a virtual learning environment can be very helpful in stimulating and enhancing understanding. Within the scope of augmented reality, Dunleavy et al. [2009] show strength and limitations of AR simulations. Their user study reveals that such a technological aid is highly engaging and adds value to the teaching and learning, while simultaneously presenting technological, managerial, and cognitive challenges. Further, such opportunities and challenges of AR in education are documented by Wu et al. [2013] and Dey et al. [2018].

Another interesting aspect of technology assisted training is in cognitive and motor skills development. A number of systems using AR-based and VR-based intervention have been shown to be effective in cognitive development. Henderson and Feiner [2011] evaluate AR user interface to assist users in psychomotor stage of procedural tasks. Marner et al. [2013] measure user performance in a procedural task using AR for assistance. On similar lines, a game based system has been proposed by Muneer et al. [2015] to improve hand-eye coordination and motor skill enhancement. It has been shown that systems based on such immersive technology are capable in rehabilitation treatments. A digital tutoring system by Takacs [2005] utilizes interactive computer graphics for teaching concepts to children and monitors their responses.

However, most of these work was conducted on neurotypical population and without a focus on socioeconomically less privileged population. Our focused users are the children from less privileged backgrounds and people with ASD. Individuals with ASD lack cognitive skills, which affects learning and following instructions.

In this work, with two separate user studies, we explore the effectiveness of AR in procedural task training for young children and how the same systems perform for people with ASD. In particular, we focus on children within an age group of nine to twelve years who have very limited exposure to computers and no exposure to AR-based systems before. The novelty of our comparative user study lies in exploring the outreach and engagement aspect of this technology in inexperienced and less privileged demographics including people with ASD. Our results indicate potential in AR-based training for this age group and also show interesting observations regarding other modes of training. We show that AR-based instructions has its own set of challenges, however it does make the entire experience more engaging compared to classical methods of training and content delivery.

This work is important because this research explores the effectiveness of mundane AR solutions to teach procedural tasks for a less privileged population who cannot afford or even use high-end AR devices such as HoloLens or even a smartphone. Most importantly, this work targets people with ASD who lack effective digital interventions for training that can help them engage more and feel less anxiety. Anxiety is a common factor in young people with autism [Rodgers et al. 2016]. Our main contributions are the insights and design lessons for developing AR-based tools for less privileged children and people with ASD.

The rest of the paper is organized as follows. Next we discuss the earlier work in this direction in Section 2, followed by the description of our training systems in Section 3. We then describe the design of our first user study in Section 4. In Section 5, we present a second user study with autistic individuals using the same training systems. We then conclude by pointing towards the future work in Section 6.

2 RELATED WORK

Spatial reasoning and cognition develops from childhood when kids start to interact with their physical environment. A number of systems for training in mixed reality have been proposed for individuals with already developed cognitive abilities. Dünser et al. [2006] presented a first large-scale user study on training aspects of spatial ability using virtual and augmented reality. The authors use a collaborative tool for 3D geometric construction using a pen and tablet based interaction. The study concludes that AR can be used to develop effective tools for spatial ability training. The other aspect of Augmented reality to enhance learning experience is recently shown to be very effective by Hung et al. [2017]. Their study consists of various teaching materials, including an AR graphic book, to evaluate children's learning abilities.

A *tangram* is an ancient Chinese geometric puzzle that uses basic shapes to create figures. Bohning and Althouse [1997] describe how tangram can be effectively used to teach geometry to young children. In addition to this, authors indicate that the puzzle helps develop spatial sense. Tangram-based AR cognitive therapies for elderly have been proposed in a few studies [Frutos-Pascual et al. 2012; Zapirain et al. 2010] to counter the effect of age-related dementia or Alzheimer. Zapirain et al. [2010] developed a modern AR-based therapy tool using computer vision to provide dynamic feedback and prompts on a computer screen. Their system is meant to assess and improve psychomotor activity in elderly using tangram puzzles. Chou et al. [2013] developed a similar system called ARTangram to assist the young children in learning to solve tangram puzzles. Our system is based on these ideas for procedural training in the AR mode. We make several enhancements to the computer vision based system to make it more friendly and easier for young children so that they can focus more on the task rather than on the system itself.

Procedural training has been taken a step further by designing rehabilitation systems for autistic children. Robot assisted tangram puzzle game has been developed by Bernardo et al. [2016] in an attempt to engage the children and increase their attention span. Lin et al. [2016] studied applicability of tangram with an AR display on children with disabilities. The participants demonstrated increased ability to complete puzzle game when compared with corresponding paper-based techniques. A tablet-based tangram game has been shown by Malisova et al. [2015] to increase social skills and communication. These systems and studies clearly indicate great potential that AR-based procedural task training holds for rehabilitation and cognitive development.

2.1 AR and VR Training Applications for People with ASD

Training ASD population using AR and VR technologies has been explore in multiple studies. Earlier studies evaluated the acceptance of these immersive technologies for people with ASD [Newbutt et al. 2016a,b]. For training, McMahon et al. [2016] developed an AR-based application for science vocabulary learning by gamifying the learning process. The AR vocabulary intervention produced a positive impact on student mastery of the science vocabulary terms through its combination of real-world and digital content. Trivilini et al. [2011] introduce multi-modal interactions in a 3D educational environment for navigation using WigWag signs. While users' perception and cognitive comprehension of what they were doing was not completely clear, but they did understand the simple multi-modal icons shown by the system (listening, visualizing and watching). For what concerns the "WigWag" sings, not all the involved users were able to completely follow them in the system. On similar lines, [Bouck et al. 2014] created a VR-based system for solving mathematical problems for students with Autism. In addition to these individual systems for teaching concepts, we found one interesting research that addressed content creation for curriculum development. The authoring tool created by [Vullamparthi et al. 2013] is capable of performing demand-based rendering of augmented reality content. The users can create lessons using their e-Learning framework.

Neuromuscular development and training are integral part of rehabilitation programs in ASD. We found a number of approaches in the literature to introduce AR and VR based tools to augment existing rehabilitation therapies. Engagement through graphics and games is an interesting approach to teach social communication [Takacs 2005]. Chan et al. [2016] proposed a dolphin training game for children with ASD to direct special focus on psychomotor skills and hand-eye coordination. In another approach, Cai et al. [2013] developed a virtual dolphinarium to let children interact with the virtual dolphins. The Mobis tool developed by Escobedo et al. [2014] helps children with autism stay focused with tasks. A mobile device running the AR tool acts as a visor to display annotated digital content. Other similar AR-based tools include ones from Tentori et al. [2015] that combines physics game engine for cognitive training, and from Nubia et al. [2015] that displays annotated pictograms and images for matching. There exist interesting AR and VR systems based on games have been designed to help such children develop cognition and motor skills [Lakshmiprabha et al. 2014; Muneer et al. 2015; Wang and Reid 2013].

We found two studies on the use of AR to support pretend play in autistic children. The system developed by Bai et al. [2012, 2013] incorporates object substitution via three common pretense themes of car, train, and airplane. Their AR system stimulated appropriate behavior of the virtual object when the corresponding physical object is moved by the child.

Overall, in the past literature, it is clear that there is a significant interest in the use and acceptance of AR-based training applications for people with ASD. However, none of the earlier studies attempted to teach procedural task training for autistic individuals from a less privileged background where affordability of the solution is also a serious concern. Our current work, explores the use of a simple AR-based tool for procedural task training for children with less privileged background and people with autism. Next we discuss the design of the training tools.

3 DESIGN OF THE LEARNING SYSTEMS

Our work consists of an AR and a VR based tangram puzzle solving activity. A participant interacts with two modes in each system - the first mode includes hints and the second mode is devoid of any hints. Both learning systems have two levels of difficulty - the easy level consisting of 7 uniquely colored tangram pieces, and the difficult level consisting of 9 uniquely colored tangram pieces.

3.1 Augmented reality-based learning system

The AR based system consists of:

- physical pieces of the puzzle enhanced with AprilTag [Olson 2011] to enable tracking of each piece, and
- a feedback system built in the Unity game engine to detect puzzle pieces and display appropriate hints on a computer screen.

The physical puzzle pieces stick to an iron board with the help of small disc magnets affixed at the back of each puzzle piece. The shape outline of tangram puzzle is printed on a transparent sheet and is placed on the board. A web-cam stationed at a height of about 40-50 cm from the table top captures top-view of the board. A feedback system built using the Unity game engine is coupled with the physical tangram, and provides a virtual view of the puzzle. The screen displays another outline of the puzzle that coincides with the physical outline seen in the camera image. Once a piece is detected in the camera image, the correct position of that piece is displayed on the screen. The user has to interpret orientation and position of the corresponding physical piece from the screen display and place it on the board.

Our feedback system identifies each piece with the help of a unique AprilTag attached to each piece. For identification and tracking of pieces we use the AprilTag implementation available from within the ViSP (Visual Servoing Platform) [Marchand et al. 2005] cross-platform library that allows prototyping and developing applications using visual tracking and visual servoing techniques. The detected tag ID of a piece is converted into the associated visual hint in Unity and displayed on the screen.

3.2 Desktop-based learning system

The Desktop-based learning system is a simple game built in Unity. The outline of tangram puzzle is displayed in the center of the screen and the pieces that need to be assembled are displayed outside of it. The user can click on a piece to move and rotate it using keyboard controls. When the user clicks on a piece, the associated hint is displayed on the screen. The user has to move a piece to its correct position inside the outline and fix its orientation to match the hint. In order to maintain uniformity in the design of the pieces across different modes of learning, virtual pieces in the desktop game are also decorated with their corresponding AprilTag.

4 USER STUDY I: NEUROTYPICAL CHILDREN

To evaluate the effects of different instruction modes on the knowledge retention of young children, we ran a *within-subjects* user study involving 12 children.

4.1 Participants

We recruited 12 children (six boys and six girls) from a local school where students commonly come from a low socio-economic background. We chose this school as the students have limited access to expensive technologies at home such as laptop or desktop computers that will enable us to explore the true impact of the technology driven instruction modes without being confounded by prior knowledge of technology usage. Ages of the children ranged between 9 and 12 years (M=10.3, SD=0.9). Only four of the children reported solving puzzles before. None of the children had laptop or desktop computers at home, only two of them reported using a desktop computer before, and nine of them reported playing games on smart phones of other family members. None of them ever used or heard of any AR or VR applications.

4.2 Experimental Task and Procedure

The experimental task was divided in two phases. First, in the **training** phase, they had to use one of the instruction modes to solve a tangram puzzle with hints. They were allowed to use the instruction mode for as long as they wanted to solve the puzzle. However, once they solved the puzzle they were not allowed to use it anymore for further practices. Second, in the **solving** phase, they had to solve the same puzzle with real tangrams without any hint. The second phase was to measure their knowledge retention from the first phase. We allowed a maximum of five minutes in this phase to solve the puzzle as we were mainly interested in their ability to remember the instructions from the first phase. A pilot study showed that if a time restriction was not given, the children will take longer (as much as 20 minutes) to solve the puzzle on their own without relying on the instructions from the earlier phase. Between the two phases, there was a mandatory break of five minutes to make participants remember the instructions before solving it. We had three instruction modes and two levels of difficulties, participants solved both easy and difficult puzzles with the same condition in a row before moving to the next instruction mode.

In between these two instruction modes, the participants had to answer a short subjective questionnaire to rate their experience of the instruction mode just used. The experimenters asked the questions and the participants had to give them a score between one and five where five is the best. The questions are mentioned in Section 4.4.4.

This break between two instruction modes enabled the children to lose their focus from the earlier content of the instruction to the questions that the experimenter asked. Overall, they had to solve six puzzles in the whole experiment as we have described below and it took about an hour on average to complete the experiment per participant.

4.3 Independent Variables

We had two independent variables—both within-subjects—in the experiment.

4.3.1 ***Instruction Mode** → AR, Desktop, In-Person.* We had three instruction modes—AR, Desktop, and In-Person. We counterbalanced the order of this variable using a 3x3 Latin-square.

The AR and Desktop modes are as described in sections 3.1 and 3.2 respectively.

In the in-person mode, the instructions were provided by the experimenter as they would normally receive in a classroom environment. The children picked one tangram and the instructor told them where to place it and the process was repeated until all the tangrams were placed in the correct manner. This level was included in the study as a control condition.

4.3.2 ***Difficulty Level** → Easy, Difficult.* We had two levels of difficulty, In the easy puzzle, there were seven tangrams and in the difficult puzzle there were nine tangrams. Participants always solved the easy tangram first for any instruction mode and then the difficult one with the same instruction mode. Hence, this variable was not counter-balanced or randomized. However, this was done purposefully as in teaching pedagogy commonly easier lessons are taught first before moving to more difficult lessons.

We had three easy and three difficult puzzles that we counterbalanced to be used in the experiment. Hence, all instruction modes was performed using the different puzzles equal number of times.

4.4 Dependent Variable

We had both quantitative and qualitative dependent variables. Overall, in the whole study we had 12 (participants) × 3 (Instruction modes) × 2 (Difficulty levels) = 72 data points.

4.4.1 Training Time (seconds). We allowed the participants to use the instruction modes for as long as they wanted until they successfully completed the lessons, i.e. they placed all tangrams in their specific locations to form the target shape successfully. We were interested to measure the time it takes for the children to go through the instructions.

4.4.2 Solve Time (seconds). We calculated the time participant needed to solve the puzzle with physical tangrams that they learned in the training phase. However, there was a maximum allowed time of five minutes per puzzle. Participants had to inform us when they thought they have solved the puzzle successfully.

4.4.3 Correct Placements. We measured how many of the tangrams were placed correctly in the right orientation. Maximum numbers of correct placements are seven and nine for easy and difficult puzzles respectively.

4.4.4 Subjective Questionnaire. Between the training phase and the solving phase we asked five subjective questions to the participants. For two of the questions (ii and iii) participants were asked to give marks out of 5, instead of ratings as they understand that analogy easily. As most school evaluations are carried out in terms of marks, students are familiar that low marks reflect unsatisfactory performance and higher marks reflect better performance.

We asked the following questions in their local language: (i) Will you be able to explain how to use the system to a friend?, (ii) On a scale of 1 to 5 how easy was it to understand the instructions using the system?, (iii) On a scale of 1 to 5 how enjoyable was it to perform the task?, (iv) One thing you liked about the activity?, and (v) One thing you disliked about the activity?

4.4.5 Observations. We video recorded all the sessions for future observations of gestures, words spoken, and behavioral cues.

4.5 Hypotheses

Before running the experiment, we had the following hypotheses.

- **H1**: As the children are used to receive instructions in the in-person mode, they will require less training time in this mode compared to the two other modes.
- **H2**: AR-based mode will have shorter solve time than the Desktop-based mode as AR-based mode enables interaction with the physical tangrams.
- **H3**: For the same reasons as in H2, Desktop-based mode will have least correct placements among the three modes.
- **H4**: In-person mode will be less preferred as both of the other modes enable interaction with the technology that provides a new experience to the children. Also AR-based mode will be preferred over the Desktop-based mode.
- **H5**: In all quantitative dependant variables, difficult puzzles will have a negative effect compared to the easy puzzles.

4.6 Results

Overall, we noticed that Desktop-based instruction required longest training time while in-person required the shortest time. Desktop-based also required longest solve time and it also had least amount of correct placements (Figure 2). Below we describe the results in more details. First, we describe the quantitative variables then the qualitative ones. We ran repeated measured ANOVAs using SPSS v25 on each of the quantitative dependent variables. For post-hoc analysis we used pair-wise comparisons with Bonferroni adjustment. In cases where the assumption of sphericity was not satisfied, we used Greenhouse-Geisser adjustments. Table 1 and Figure 2 provide the high level overview of the results.

4.6.1 Training time. We noticed a significant main effect of instruction modes on training time—F(1.2, 13.6)=45.74, $p < .001$, η_p^2=.81, Observed Power (OP)=1 (Figure 2a). A pairwise comparison showed that Desktop-based instruction look significantly longest than AR-based ($p < .001$) and in-person ($p < .001$) modes. In-person mode required significantly less training time than the AR-based mode (p=.02). There was no significant main effect of complexity but there was a trend—F(2, 11)=4.1, $p = .07$, η_p^2=.27, OP=.45.

While we did not find a significant interaction effect of *instruction mode* $\times$ *difficulty* with the number of participants but there was trend—F(2, 22)=3.28, $p = .057$, η_p^2=.23, OP=.56. In case of Desktop mode the hard puzzles took considerably more time than the easy puzzle, while for the other two modes the difference was as much (Figure 2a).

4.6.2 Solve time. We noticed a significant main effect of instruction mode on solve time—F(2, 22)=46.1, $p < .001$, η_p^2=.8, OP=1 (Figure 2c). A pairwise comparison identified that puzzle solving time following the Desktop-based instruction mode was significantly slowest than AR-based ($p < .001$) and in-person ($p < .001$) modes (Figure 2b). There was no other significant difference found. One fact to note here that none of the participants were able to complete the difficult puzzle within the allowed five minutes following the Desktop-based mode.

4.6.3 Correct placements. Similar to other dependant variables, we found a significant main effect of instruction mode on number of correct placements—F(2, 22)=33.51, $p = .001$, η_p^2=.57, OP=1. A pairwise comparison found that the desktop-based mode had significantly least number of correct placements than AR-based (p=.001) and in-person (p=.002). There was no other significant effects found.

4.6.4 Questionnaire. We asked five questions after exposure to each of the instruction modes. Below are the findings. First, we wanted to know whether the children would be able to explain how to use the system to their friends. For the AR-based mode, eight out of the 12 participants said that they would be. In comparison, only three participants said they can explain the desktop-based mode.

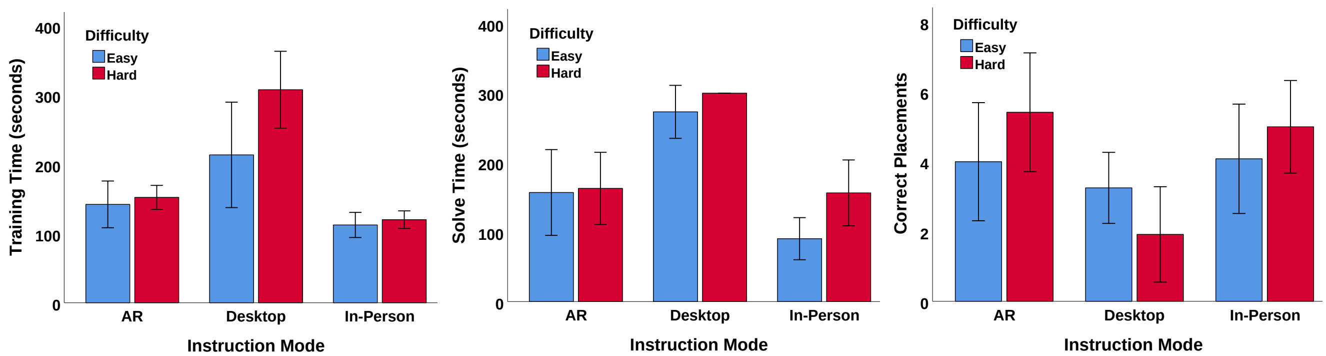
The next two questions asked them to give a score on the understanding the instructions and enjoyment while using the systems. We did not find any significant difference between the three instruction modes.

Then we asked them about their likes and dislikes for each of the instruction modes. For AR-based mode, they liked the fact that the tangrams were detected by the system and that they can move the tangrams physically while looking at the screen. Overall, they were amused that the system were highlighting hints on the puzzle. However, they reported dislikes when the camera was not able to detect the piece easily, either due to low light or not being able to place the tangram in the view of the camera. They also reported dislike when they were not able to rotate the tangram easily in its correct orientation following the hints. For the Desktop-based mode, they liked the ability to click on an object and moving it to the correct location following the hints. However, they disliked that the at times the manipulation was too fast for them and their fingers started to hurt after some time. For the in-person mode, they liked that there was a human guiding them and they could asked questions as required. However, they also reported dislikes when they were not able to understand the instructions and had to ask more.

At the end of the experiment during debriefing, we asked participants which instruction mode did they like the most? Five participants preferred the in-person mode, four preferred the AR mode, and three preferred the desktop mode most.

4.6.5 Observations and Comments. We took notes during the experiment and also video recorded the participants to identify their overall attitude and behavior towards the instruction modes. Below are some of the key observations.

In the AR-based mode, participants took time in understanding hints on the computer screen and translating the on-screen orientation of a puzzle piece to its real orientation of the physical piece.



(a) Training Time (b) Solve Time (c) Correct Placements

Figure 2: Results of the User Study I. Whiskers represent ±95% confidence intervals.

Table 1: User Study I: Mean and Standard Deviation values of the dependent variables.

Mode	Difficulty	Quantitative			Qualitative	
		Training Time (seconds)	Solve Time (seconds)	Correct Placements	Easy to explain	Enjoyable
AR					4.1 (1.2)	4.7 (0.8)
	Easy	142.2 (53.1)	157.0 (97.2)	4.0 (2.7)		
	Hard	152.3 (27.3)	162.9 (81.7)	5.4 (2.7)		
Desktop					4.3 (1.2)	4.0 (1.1)
	Easy	213.7 (119.9)	273.1 (60.0)	3.3 (1.6)		
	Hard	307.8 (87.3)	300.0 (0.0)	1.9 (2.2)		
In-Person					4.2 (0.7)	4.5 (0.7)
	Easy	112.4 (28.5)	90.5 (47.7)	4.1 (2.5)		
	Hard	120.1 (19.9)	156.4 (74.7)	5.0 (2.1)		

As mentioned in the earlier section, a few participants did report that the mental transformation required to map a virtual object on the screen and to its physical object in their hands was difficult. For example, P11 said "I was unable to match the piece with how it looked on screen". As all of the participants were using AR for the first time, initially they took longer to understand how the system works and how to place the tangram pieces (with markers) within the camera's field of view to get the hints on the screen. However, after they understood the concept they were confident and faster in using the system.

In the desktop-based mode, participants struggled with using track-pad to move the pieces. They took more time getting used to the interaction with the desktop system as opposed to other modes of learning. It is an interesting finding as our participants had almost no experience of using computers—which is a hard to find population these days—a fair comparison between AR mode and Desktop mode was possible, where AR mode was found to be easier to learn. We also noticed that in this mode participants tended to remember larger tangram pieces more than the smaller ones. Majority of the participants took more time interacting with the harder level as compared with the easy one. We also found that despite scoring low with the desktop, a few participants reported having fun in this mode.

In the in-person mode, primarily the participants found it difficult to understand the placement and orientation of tangram pieces that were not surrounded by any other pieces due to lack of references. The participants rated this mode favourably since the feedback was being provided constantly till they got the orientation of any piece correct. However, it was difficult for the instructor to describe the exact position of a piece that was not surrounded by other pieces.

4.7 Discussion

In this user study, we found a few interesting observations that are unique to the population of children that participated. Here we discuss the results in relation to the hypotheses we had prior to the experiment.

Our first hypothesis stated that the children will require less training time using the in-person mode. This hypothesis was accepted as we noticed in-person training mode was significantly faster than the other two instruction modes. As it was the first time that the participants interacted with a mouse or a trackpad, the training time in the Desktop-based mode was significantly higher compared with that of the other modes. It also shows that AR can perform better than the Desktop or laptop based instructions for a less privileged population who do not have adequate exposure to the modern technologies.

Our second hypothesis stated that the children will solve the puzzles faster after getting instructions from the AR-based mode over the Desktop-based mode as AR enables them to interact with the physical objects. This hypothesis was accepted as we noticed significant difference between the AR-based mode and the Desktop-based mode for the solve time. Desktop-based mode was in particular worse for difficult puzzles. We believe this performance issue was caused due to the usability challenges the children faced while using this mode. They had to focus their attention more on the usability of the Desktop-based system than on the instructions for solving the puzzle.

In our third hypothesis we expected that the Desktop-based mode will have least correct placements due to the challenges of learning using this mode. This hypothesis was accepted as well. This could be attributed to the fact that participants focused more on the movement of pieces on the screen rather than remembering their exact positions.

These children did not have any prior experience of working with computers, touch devices, and AR. Before the user study, in out fourth hypothesis, we expected that the participants would prefer the digital instruction modes due to their novelty over the in-person mode. However, since these children are very well familiar with the in-person teaching instruction in a classroom environment, the in-person instruction mode turned out to be the most preferred. Further, constant feedback in the in-person mode also played a positive role, whereas in the other modes the participants could only receive visual feedback. Hence, this hypothesis was not accepted.

In out fifth hypothesis, we expected the hard puzzles will have negative effect on all performance matrices. However, we cannot accept this hypothesis as we did not notice a significant effect of puzzle difficulty. It is possible that the difficulty levels were not that different to cause an effect. Other studies in the VR domain have found that human is capable of adapting to increasing challenges without decreasing cognitive performances [Dey et al. 2019].

5 USER STUDY II: INDIVIDUALS WITH AUTISM

Following the first user study we tested the same systems with people with autism in collaboration with a not-for-profit organization (school hereafter) specializing in training autistic individuals. This user study was more formative in nature than summative. Our main motivation for this user study was to explore how, and if at all, this simple AR-based instruction mode can be used in training individuals with ASD and how such interventions can be designed to make their learning experience more engaging and enjoyable?

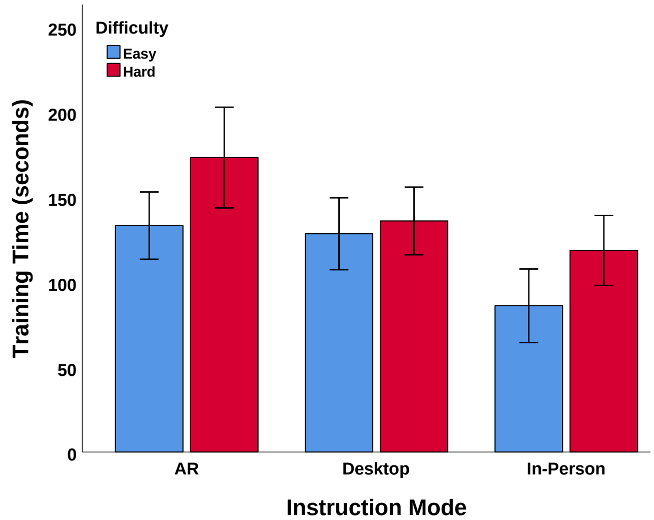


Figure 3: Training time for people with ASD using different instruction modes. Whiskers represent ±95% confidence interval.

The experimental task was primarily to go through the instructions and complete the tangram puzzles. To reduce the load on the participants we asked them go though one instruction mode each day. As we ere interested more on their use of the instructions, we created three different versions of easy (five tangrams) and hard (seven tangrams) puzzles; and these puzzles were counterbalanced ensuring all instruction modes were performed using all puzzles equal number of times. The number of pieces were reduces from the User Study I after consultation with the instructors at the school. Overall, everyday each participant had to complete six puzzles using one instruction mode and over three days a total of eighteen puzzles using all three instruction modes.

There was another difference between the execution of the first user study and this one. Here, the in-person instructions and instruction on the use of other modes were provided by an instructor from the school as opposed to the experimenters in User Study I. However, the experimenters always accompanied the instructor in all sessions and they intervened when the participants required more clarifications related to the description of the system and/or the task.

5.1 Participants

We recruited seven volunteers for this experiment. However, two of them did not attend all of the sessions that resulted in incomplete data. Here we report results from the remaining five participants. Out of these five participants with diagnosed autism, two were diagnosed with down syndrome and three with mild mental retardation. Based on Malin's intelligence scale [Malin 1969], all participants had an intelligent quotient (IQ) of less than 55 (M=42.6, SD=7). Their social quotient (SQ) based in the Vineland Social Maturity Scale [Doll 1953] ranged between 47 and 69 (M=54, SD=9.1). None of the participants solved tangram puzzles before. All of the participants come from low to middle income families. However, all of them have used computers before at the school.

5.2 Results and Observations

As this experiment was formative in nature we only measured how long it took participants to complete the puzzles using the instructions. Although we told them to complete the puzzles without the instructions afterwards but none of them could complete the puzzles perfectly using any of the instruction modes. Figure 3 shows that people with ASD took longest to understand instructions using the AR mode (M=153.1, SD=49.1), while the in-person mode (M=102.3, SD=40.9) was the fastest. While the Figure 3 indicates possible significant differences between the instruction modes, we did not perform any statistical analysis as data from five participants can not be used to reliably test hypothesis. Table 2 shows the numerical values for each instruction mode and task difficulty.

We observed that participants with autism were more familiar with in-person instructions and their relationship with their instructor played a big role in their performance. Whereas for the other two digital instruction modes, where the experimenter also had to intervene, they took longer to process the information.

As the participants had prior experience of using computers they did not face any serious issues with the usability of the system or understanding the instructions in most cases. In this case, the time they took to complete he easy and hard puzzles were not so different, unlike the other tow instruction modes (Figure 3).

The usability of the AR-based instruction mode was affected by the understanding of the marker. First, it took longer for the participants to understand the relationship of the marker on the tangrams to the visual feedback they were receiving on the screen. Several times the participants blocked the markers and did not get instruction on the screen they were expecting. For one participant this issue was predominant, which took them noticeably longer to compete the puzzle. Second, we noticed that some participants were showing the reverse side of the tangram where there was no marker. However, this issue was easily corrected after instructing them and did not affect the overall training time considerably. However, besides these issues, participants were able to the AR-based mode reasonably well and when they got used to it the above issues were occurring less. Most of the participants understood the rotation of the tangrams shown on the screen and could place them correctly, with occasional misinterpretations, which is a positive finding as the spatial relationship of virtual and physical objects are often challenging for people with autism [Ohta 1987]. None of the participants showed any frustration or agitation when using the AR-based tool and in some cases, towards the start, they were positively surprised to see the AR instructions and repeatedly smiled.

Another behavioral thing we noticed was the participants' reliance on their instructor. Whenever, they struggled to figure out the correct placement of a tangram, they handed it over to their instructor to show them the correct placement.

Table 2: Mean and standard deviation values of training times for User Study II.

Mode	Difficulty	Training Time (seconds)
AR		153.1 (49.1)
	Easy	133.1 (35.7)
	Hard	173.1 (53.4)
Desktop		132.1 (36.6)
	Easy	128.3 (38.2)
	Hard	135.9 (35.9)
In-Person		102.3 (40.9)
	Easy	86.0 (39.1)
	Hard	118.5 (37.2)

5.3 Discussion

Overall, we believe the design of AR-based tools for people with ASD requires different design thinking than that is for the neurotypical children. While they were not able to complete the tangram puzzles without the hints in any instruction mode, the time they took to complete the puzzles with instructions was noticeably high for the AR mode. The issues of the marker can be resolved with more practice. We also noticed the need for encouragement while learning is essential. Future designers of such interventions need to provide positive encouragements while using digital instruction modes such as AR or VR.

Most importantly, for them any instruction mode without the presence of their instructor will not work well as they require attention and support regularly. It will be interesting to explore how remote presence of the instructor or caregivers can be incorporated in the AR or VR-based modes to enable more effective training without the physical presence of the instructor.

6 CONCLUSION AND FUTURE WORK

In this work we investigated the use of AR-based systems to impart procedural task training to young children (User Study I) and how similar systems can be used for people with autism (User Study II). Previous research indicates the promise of using virtual and augmented realities in order to increase engagement and improve learning outcomes. To compare different modes of instruction, we designed a tangram building exercise with varying difficulty levels (easy and hard) and imparted instructions in three modes to the subjects - AR-based, Desktop-based, and in-person training.

The first user study comprised of 12 children in the age group of 9 - 12 years. We observed that Desktop-based instruction was the least favored mode of learning among the participants. Interestingly, the training time and solve time were the longest in this case as well. In-person training was the most favored among the participants, while the AR-mode showed better performances that the Desktop-based mode. AR-based and in-person instructions also resulted in more tangram pieces being correctly placed by the participants, as compared to Desktop-based instruction. With this study, we established how specialized populations with little or no experience with computers respond to new instruction modes, particularly using AR. Our study indicates the potential of exploring AR technology to achieve better learning outcomes and learner engagement among less privileged population who has limited prior exposure to technology.

In future, we would like to explore using brain-sensing technologies how different instruction modes create different cognitive loads for children and how an adaptive intelligent tutoring system can be developed with low-cost solutions for less privileged children. Here we noticed that children performed reasonably well using the AR-based mode. Due to the cost considerations, we created an affordable AR-based solution. We would like to compare this tool with other advanced AR devices such as HoloLens.

The second user study was conducted with five people with autism using the same instruction modes. We noticed that here AR-based mode performed worse than other modes, however, similar to the first user study in-person mode performed the best. While AR did not cause any adverse effect we noticed that marker-based AR possesses usability challenges for people with autism, particularly when the stimulus-response compatibility is compromised. We also noticed that participants with autism could not complete any of the tangram puzzles without instructions, which indicates the learning was not established. In future, we would like to explore alternative AR-based solutions that does not rely on markers and evaluate real-world impact using other procedural tasks.

REFERENCES

Zhen Bai, Alan F Blackwell, and George Coulouris. 2012. Making Pretense Visible and Graspable: An augmented reality approach to promote pretend play. In *Mixed and Augmented Reality (ISMAR), 2012 IEEE International Symposium on*. IEEE, 267–268.

Zhen Bai, Alan F Blackwell, and George Coulouris. 2013. Through the looking glass: Pretend play for children with autism. In *Mixed and Augmented Reality (ISMAR), 2013 IEEE International Symposium on*. IEEE, 49–58.

Beatriz Bernardo, Patrícia Alves-Oliveira, Maria Graça Santos, Francisco S Melo, and Ana Paiva. 2016. An Interactive Tangram Game for Children with Autism. In *International Conference on Intelligent Virtual Agents*. Springer, 500–504.

Mark Billinghurst. 2002. Augmented reality in education. *New horizons for learning* 12, 5 (2002).

Gerry Bohning and Jody Kosack Althouse. 1997. Using tangrams to teach geometry to young children. *Early childhood education journal* 24, 4 (1997), 239–242.

Emily C Bouck, Rajiv Satsangi, Teresa Taber Doughty, and William T Courtney. 2014. Virtual and concrete manipulatives: A comparison of approaches for solving mathematics problems for students with autism spectrum disorder. *Journal of Autism and developmental disorders* 44, 1 (2014), 180–193.

Yiyu Cai, Noel KH Chia, Daniel Thalmann, Norman KN Kee, Jianmin Zheng, and Nadia M Thalmann. 2013. Design and development of a virtual dolphinarium for children with autism. *IEEE transactions on neural systems and rehabilitation engineering* 21, 2 (2013), 208–217.

Sandra Chan, Yiyu Cai, Andrew Lu, Nay Zin Tun, Lihui Huang, and Indhumathi Chandrasekaran. 2016. Virtual Reality Enhanced Pink Dolphin Game for Children with ASD. In *Proceedings of the 3rd Asia-Europe Symposium on Simulation & Serious Gaming (VRCAI '16)*. ACM, New York, NY, USA, 215–218. https://doi.org/10.1145/3014033.3014039

C.-Y. Chou, C.-C. Yang, and Z.-H. Chen. 2013. Applying augmented reality to assisting children in solving Tangram puzzle. *Proceedings of the 21st International Conference on Computers in Education, ICCE 2013* (2013), 139–144.

Arindam Dey, Mark Billinghurst, Robert W. Lindeman, and J. Edward Swan. 2018. A Systematic Review of 10 Years of Augmented Reality Usability Studies: 2005 to 2014. *Frontiers in Robotics and AI* 5 (2018), 37. https://doi.org/10.3389/frobt.2018.00037

Arindam Dey, Alex Chatburn, and Mark Billinghurst. 2019. Exploration of an EEG-Based Cognitively Adaptive Training System in Virtual Reality. In *Virtual Reality (VR), 2011 IEEE International Conference on*. IEEE, 1–7.

Edgar A Doll. 1953. The measurement of social competence: A manual for the Vineland Social Maturity Scale. (1953).

Matt Dunleavy, Chris Dede, and Rebecca Mitchell. 2009. Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. *Journal of science Education and Technology* 18, 1 (2009), 7–22.

Andreas Dünser, Karin Steinbügl, Hannes Kaufmann, and Judith Glück. 2006. Virtual and augmented reality as spatial ability training tools. In *Proceedings of the 7th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction: design centered HCI*. ACM, 125–132.

Lizbeth Escobedo, Monica Tentori, Eduardo Quintana, Jesus Favela, and Daniel Garcia-Rosas. 2014. Using augmented reality to help children with autism stay focused. *IEEE Pervasive Computing* 13, 1 (2014), 38–46.

M Frutos-Pascual, B García-Zapirain, and A Méndez-Zorrilla. 2012. Improvement in cognitive therapies aimed at the elderly using a mixed-reality tool based on tangram game. In *Computer Applications for Graphics, Grid Computing, and Industrial Environment*. Springer, 68–75.

Daniel H Geschwind and Pat Levitt. 2007. Autism spectrum disorders: developmental disconnection syndromes. *Current opinion in neurobiology* 17, 1 (2007), 103–111.

Steven J Henderson and Steven K Feiner. 2011. Augmented reality in the psychomotor phase of a procedural task. In *Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on*. IEEE, 191–200.

Y-H Hung, C-H Chen, and S-W Huang. 2017. Applying augmented reality to enhance learning: a study of different teaching materials. *Journal of Computer Assisted Learning* 33, 3 (2017), 252–266.

NS Lakshmiprabha, Alexandre Santos, Dimitar Mladenov, and Olga Beltramello. 2014. [Poster] An augmented and virtual reality system for training autistic children. In *Mixed and Augmented Reality (ISMAR), 2014 IEEE International Symposium on*. IEEE, 277–278.

Kangdon Lee. 2012. Augmented reality in education and training. *TechTrends* 56, 2 (2012), 13–21.

Chien-Yu Lin, Hua-Chen Chai, Jui-ying Wang, Chien-Jung Chen, Yu-Hung Liu, Ching-Wen Chen, Cheng-Wei Lin, and Yu-Mei Huang. 2016. Augmented reality in educational activities for children with disabilities. *Displays* 42 (2016), 51–54.

AJ Malin. 1969. Malin's intelligence scale for Indian children. *Nagpur (India): Child Guidance Center* (1969).

Katerina Malisova, Katerina Mania, S Moysiadou, and Elias Kourkoutas. 2015. Tangram quests: a tablet adventure game assisting children with high functioning autism to enhance their communicative and social skills. *EDULEARN15 Proceedings* (2015), 481–487.

Éric Marchand, Fabien Spindler, and François Chaumette. 2005. ViSP for visual servoing: a generic software platform with a wide class of robot control skills. *IEEE Robotics & Automation Magazine* 12, 4 (2005), 40–52.

Michael R Marner, Andrew Irlitti, and Bruce H Thomas. 2013. Improving procedural task performance with Augmented Reality annotations. In *Mixed and Augmented Reality (ISMAR), 2013 IEEE International Symposium on*. IEEE, 39–48.

Don D McMahon, David F Cihak, Rachel E Wright, and Sherry Mee Bell. 2016. Augmented reality for teaching science vocabulary to postsecondary education students with intellectual disabilities and autism. *Journal of Research on Technology in Education* 48, 1 (2016), 38–56.

Reema Muneer, Tanushree Saxena, and Prathibha Karanth. 2015. Virtual Reality Games as an Intervention for Children: A Pilot Study. *Disability, CBR & Inclusive Development* 26, 3 (2015), 77–96.

Nigel Newbutt, Connie Sung, Hung Jen Kuo, and Michael J Leahy. 2016a. The potential of virtual reality technologies to support people with an autism condition: A case study of acceptance, presence and negative effects. *Annual Review of Cyber Therapy and Telemedicine (ARCTT)* 14 (2016), 149–154.

Nigel Newbutt, Connie Sung, Hung-Jen Kuo, Michael J Leahy, Chien-Chun Lin, and Boyang Tong. 2016b. Brief report: A pilot study of the use of a virtual reality headset in autism populations. *Journal of autism and developmental disorders* 46, 9 (2016), 3166–3176.

Rincon M Nubia, Garay R Fabian, Rodriguez A Wilson, and Perez B Wilmer. 2015. Development of a mobile application in augmented reality to improve the communication field of autistic children at a Neurorehabilitar Clinic. In *Engineering Applications-International Congress on Engineering (WEA), 2015 Workshop on*. IEEE, 1–6.

Masataka Ohta. 1987. Cognitive disorders of infantile autism: A study employing the WISC, spatial relationship conceptualization, and gesture imitations. *Journal of Autism and Developmental Disorders* 17, 1 (1987), 45–62.

Edwin Olson. 2011. AprilTag: A robust and flexible visual fiducial system. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*. IEEE, 3400–3407.

Zhigeng Pan, Adrian David Cheok, Hongwei Yang, Jiejie Zhu, and Jiaoying Shi. 2006. Virtual reality and mixed reality for virtual learning environments. *Computers & graphics* 30, 1 (2006), 20–28.

Jacqui Rodgers, Sarah Wigham, Helen McConachie, Mark Freeston, Emma Honey, and Jeremy R Parr. 2016. Development of the anxiety scale for children with autism spectrum disorder (ASC-ASD). *Autism Research* 9, 11 (2016), 1205–1215.

Barnabas Takacs. 2005. Special education and rehabilitation: teaching and healing with interactive graphics. *IEEE Computer Graphics and Applications* 25, 5 (2005), 40–48.

Monica Tentori, Lizbeth Escobedo, and Gabriela Balderas. 2015. A smart environment for children with autism. *IEEE Pervasive Computing* 14, 2 (2015), 42–50.

Alessandro Trivilini, Licia Sbattella, and Roberto Tedesco. 2011. Multimodal Interaction for Users with Autism in a 3D Educational Environment. In *IADIS International Conference on Interfaces and Human Computer Interaction 2011*. 1–9.

Annie Joyce Vullamparthi, Sarat Chandra Babu Nelaturu, Dakshayani D Mallaya, and S Chandrasekhar. 2013. Assistive learning for children with autism using augmented reality. In *Technology for Education (T4E), 2013 IEEE Fifth International Conference on*. IEEE, 43–46.

Michelle Wang and Denise Reid. 2013. Using the virtual reality-cognitive rehabilitation approach to improve contextual processing in children with autism. *The Scientific World Journal* 2013 (2013).

Hsin-Kai Wu, Silvia Wen-Yu Lee, Hsin-Yi Chang, and Jyh-Chong Liang. 2013. Current status, opportunities and challenges of augmented reality in education. *Computers & education* 62 (2013), 41–49.

Begoña García Zapirain, Amaia Méndez Zorrilla, and Sabin Larrañaga. 2010. Psychostimulation for elderly people using puzzle game. In *Games Innovations Conference (ICE-GIC), 2010 International IEEE Consumer Electronics Society's*. IEEE, 1–8.