Comparing Children's Crosshair and Finger Interactions in Handheld Augmented Reality: Relationships Between Usability and Child Development

Iulian Radu
School of Interactive Computing
Georgia Institute of Technology
Atlanta, GA, USA 30332
iulian@gatech.edu

Blair MacIntyre
School of Interactive Computing
Georgia Institute of Technology
Atlanta, GA, USA 30332
blair@cc.gatech.edu

Stella Lourenco
Department of Psychology
Emory University
Atlanta, GA, USA 30322
stella.lourenco@emory.edu

ABSTRACT
Augmented reality technology is a unique medium that can be helpful for young children's entertainment and education, but in order to achieve the benefits of this technology, augmented reality experiences need to be appropriately designed for young children's developing physical and cognitive skills. In the present study we investigated how 5-10 year-old children react to typical handheld augmented reality interaction techniques such as crosshair selection and finger selection, in AR environments that require them to change perspective or not. Our analysis shows significant impacts of age upon AR performance, with young children having slower selection times, more tracking losses, and taking longer to recover tracking. Significant differences were also found between AR interaction technique conditions, with finger selection being faster than crosshair selection, and interactions which required changes in perspective taking longer, generating more tracking losses, and more errors in selection. Furthermore, by analyzing children's performance in relation to metrics of physical and cognitive development, we identified correlations between AR interaction techniques performance and developmental tests of spatial relations, block construction and visuomotor precision. Gender differences were analyzed but no significant effects were detected.

Author Keywords
Augmented reality, children, interaction techniques, mobile, design guidelines, usability, developmental psychology.

ACM Classification Keywords
H.5.1 Multimedia Information Systems: Artificial, augmented, and virtual realities.

IDC '16, June 21-24, 2016, Manchester, United Kingdom

DOI: http://dx.doi.org/10.1145/2930674.2930726

INTRODUCTION
Augmented reality (AR) is a technology that enables computer-generated imagery to be superimposed on a view of the real world. AR has been shown to have measurable benefits to children's education over traditional approaches, with regards to increasing students' understanding of spatial and conceptual knowledge, improving long term memory retention, enabling collaboration around virtual content and increasing motivation. [1,2,3,4]. To achieve these benefits, augmented reality experiences need to be appropriately designed for young children. However, in the augmented reality design community there is currently a lack of systematic understanding of how to design AR experiences that are usable by children of specific ages, despite the fact that a variety of usability problems exist when young children attempt to engage with this technology [5].

In the present research we are interested in developing an understanding of AR usability for children, by specifically investigating the abilities of young children (aged 5-10) to use two common handheld-AR selection techniques: *crosshair selection* (where the selection point is a crosshair at a fixed location on the screen and the finger can be tapped on side on-screen buttons to trigger a selection) and *finger tap selection* (where the selection point is the position on the screen the finger is tapped).

We are interested in augmented-reality using handheld devices such as phones and tablets (*handheld-AR*) since we believe that this will be the dominant method for delivering AR experiences to large amount of young children for the foreseeable future. The results of this research will therefore focus on handheld-AR; however, it is likely that some of the results of this research will be generalizable to other kinds of physically engaging technologies where children interact by manipulating handheld displays.

BACKGROUND AND RELATED WORK
In this section, we review related work in the domains of interaction technique studies for children, augmented reality interactions, cognitive and physical development, and gender differences.

Children and Interaction Techniques

Many studies have investigated children's performance on interaction techniques for PC based applications, whereby an increase in the complexity of mouse- and joystick-based interaction gestures leads to decreased performance (as measured by slower task completion, and higher occurrence of errors) [6,7,8,9,10,11]. Age differences have been observed in children's interactions with various handheld devices such as pen and touchscreen input, and studies show that pointing accuracy increases with age and target size [12,13,14,15,16,17]. Compared to younger children, older children and adults are quicker, less jagged and more accurate [8,11,13]. These studies correlate with motor development literature that suggests children's motor skills continue to develop into early teens. These effects transfer to handheld touchscreen devices, as children also have trouble with more complex gestures such as double-tapping, when compared to single touching, and they show a high degree of unintended touching [8,12,18,19] .

Interaction in Augmented Reality

In an augmented-reality experience such as a mobile game, the user can interact with the game using a variety of techniques. The "interaction technique" allows the user to cause actions in the game, such as selecting which item they wish to act upon. As an example, the user could collect an item by touching on the mobile device screen with one's finger, or could aim with a crosshair in the middle of the screen. The relationship between children's age and their performance on these kinds of interactions has not been empirically evaluated, thus it is unclear how each interaction compares in terms of usability issues, or what underlying developmental skills might impact children's performance on these interactions.

A developmental framework for understanding children's usability issues in augmented reality has been proposed in [5], by hypothesizing relationships between augmented reality usability problems and the developmental skills which children might be required to use. Examples of usability problems experienced by children include inability to properly hold AR devices, inaccuracy in interacting with virtual objects, and an inability to remember spatial locations of virtual items, among other things [5,20,21,22]. When children are engaged in an augmented reality experience, they are performing a variety of activities which might not be strongly employed in exposure to typical technologies such as computers and touchscreen devices. For instance, an AR experience might require children to move around the game space in order to observe virtual content from different angles, and they may be required to hold and manipulate a smartphone while looking at augmented objects, and/or might be required to understand and remember the configuration of the virtual objects around the physical space, as well as the relationships between their body and virtual objects. Such activities require different cognitive and physical skills, and when young children lack these skills, they should have difficulties using AR interfaces.

Underlying Developmental Skills

Developmental psychology tells us that kindergarten and elementary-school children have certain capabilities and limitations, which are different than adults and which develop over time as children mature into adults [23,24]. These developing abilities mediate children's ability to use digital technology, and researchers have stressed the importance of considering developmental abilities when designing technology for children [25,26,27,28].

Children's AR usability issues are potentially explainable by psychology areas such as motor skills (bimanual coordination, hand-eye coordination, fine motor skills, gross motor skills, endurance), spatial cognition (spatial memory, spatial perception, spatial visualization), attention (selective attention, divided attention, executive control), logical thinking (abstract vs. concrete thinking) and memory (capacity and operations). These developmental factors are all undergoing development in young children, thus have the potential to influence how children react to AR applications. In the present research, we intend to measure children's developmental skills by using existing psychology metrics, and we intend to correlate these metrics to children's performance on specific AR designs. Through these correlations, we will identify differences between children's ability to use AR at different ages. Furthermore, by understanding the developmental factors and their potential connection to usability issues, AR designers can use this knowledge to explain or predict user reactions to existing applications, and to design applications that are either usable for specific age groups, or design applications that aim to develop specific skills.

Gender Differences

Gender differences in spatial skills are significant for teens and adults, with males exhibiting higher performance for visuospatial skills [29,30], and females exhibiting higher performance for spatial location memory [31]. However, gender differences in spatial development do not appear before age 10 [31,32]. It is not expected that gender will influence young children's understanding of AR space. Gender differences do not appear significant for young children's motor performance. Children's physical development is similar for both genders during the years of early childhood, and studies of young children's motor performance do not detect differences between genders in young children [33,34]. Studies of children's touchscreen performance also do not find significant effects of gender in young children [12].

DESIGN

In this research study we were interested in the following questions:

1. What kinds of performance differences are experienced by children when using different

kinds of augmented reality interaction techniques? We compared 4 kinds of interaction conditions, which differed among the factors of selection type (finger or crosshair) and motion requirement (perspective change was required or not required).

2. What is the relationship between children's age and their performance in using different AR interaction techniques? And, how does performance on interactive techniques relate to children's cognitive and physical development?
3. How does the children's gender relate to their performance on different AR interaction techniques?

Tasks and Game Design

To investigate the research questions, we designed a within-subjects experiment, whereby each participant played an augmented-reality game exposing them to different interaction technique conditions. Children's performance was measured while playing the game using a Motorola Atrix HD smartphone. After playing the game, children's skills were measured through several developmental tests.

The game is structured as a typical tabletop augmented reality game, where a three-dimensional virtual world appears on top of the "gameboard" paper once it is viewed through a smartphone camera (Figure 1). In our experimental setup, the game was placed on a table without chairs, and children had space to move around the table at their leisure. The height of the gameboard was adjusted to be at the level of each player's stomach, such that each player could comfortably observe the gameboard through the smartphone.

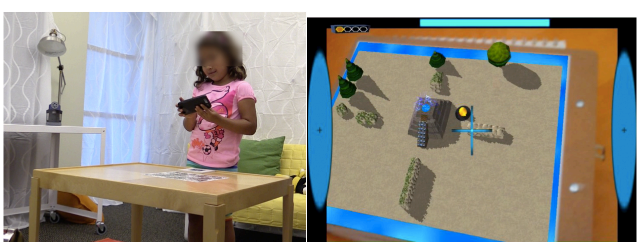

Figure 1. Child playing the augmented reality game, in a level where lemons were collected through Crosshair Selection.

In the game, the player took the role of a wizard who must collect a set of magical lemons in order to create items for their pets. The game was composed of a tutorial plus 4 levels. Each level of the game was associated with a different AR interaction technique condition, and was composed of a "lemon collection" phase and a "mini game" phase. During lemon collection, the player had to gather a set of 16 lemons. Once all lemons were collected, they magically transformed into a play item and the mini game started. In the mini game, the child played with their pet and the newly-created item (Figure 2). During the "lemon collection" phase, all children played the augmented reality game while standing or moving around the table. During the "mini game" phase at the end of each level, the game did not involve any augmented reality and did not require children to look at the gameboard; thus children were asked to sit during this phase. The mini game was a requirement of our game design in order to give children a rest period from standing, while at the same time offering entertainment and agency. During the data analysis, only data from the lemon collection phases was analyzed.

Figure 2. The game characters and magical objects (left) and the mini-game associated with one object (right).

Prior to the set of 4 gameplay levels, children were exposed to a tutorial phase. During our pilot testing we determined that children were performing poorly because they did not have previous exposure to the technology, so a tutorial was developed to familiarize children with augmented reality technology. During the tutorial, children had to move around the table and visualize the 3D game world from different angles, and they were exposed to all 4 different interaction technique conditions which were involved in the different levels of the game.

Interaction Technique Factors

We tested four variations of augmented-reality interaction technique conditions (Figure 3). These variations were driven by two factors: Selection Type and Movement Difficulty.

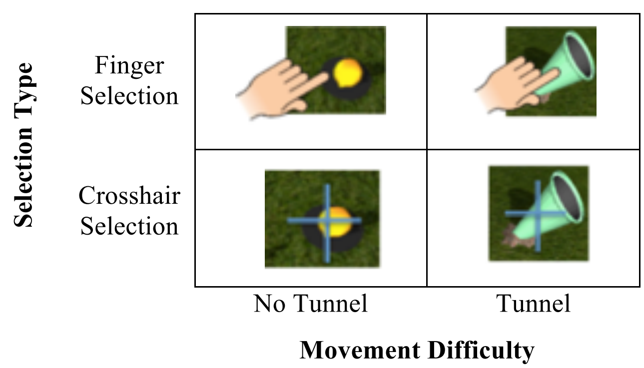

Figure 3. The interaction conditions tested.

Selection Type has two levels: Finger Selection or Crosshair Selection. When using Finger Selection, players must touch their finger to the screen position where a target lemon appears. In Crosshair Selection, the players have a crosshair in the center of their screen, and they select a target lemon by touching one of the screen side buttons once the lemon is overlapping the crosshair (Figure 1 right). Each selection type has an accuracy distance threshold, which was empirically determined. In the Finger Selection mode, the lemon surface can be 60 pixels away from the screen touch point; this corresponded to 4.55mm in our experimental smartphone. This threshold was selected to account for the problem of finger occlusion [35], and it corresponds to the size of an average 3-10-year old child's finger [36]. In Crosshair Selection, the lemon can be 35 pixels away from the center of the crosshair (2.7 mm). This distance was used to account for angular reorientation of the phone while children press the crosshair buttons, and was determined during our pilot study by analyzing how much children shake the phone in the last 1s prior to touching the crosshair buttons.

Movement Difficulty, the other factor influencing our interaction conditions, impacts whether the players need to change perspective during gameplay. It has two types: No Tunnels and Tunnels. In the No Tunnels condition, the player can see the targets from any angle and does not need to move their body. In the Tunnel condition, the targets are enclosed in virtual tunnels, thus the player must change their perspective in order to select the target. The targets in the Tunnel conditions were rotated such that, between each lemon, players were forced to change their angle relative to the gameboard by 45 degrees while remaining relatively in front of the table.

Each child was exposed to 4 different interaction technique conditions, varying on the conditions of Selection Type and Movement Difficulty (described above). The Selection Type conditions were randomized between players, while the Movement Difficulty conditions were not randomized (all players experienced No Tunnels before Tunnels). The game environment was randomized between conditions.

Physical and Cognitive Development Factors

In this study the children's *Gender* were both Male and Female, and their *Age* was between 5 and 10 years old. For analysis we measured participant ages in months, and also grouped age into three groups: 5-6, 7-8, and 9-10 year olds.

We measured participants' *Hand Size* because it should influence how easily children are able to manipulate the phone and perform various interaction techniques.

Children were exposed to three kinds of developmental tests, from the NEPSY II and Woodcock Johnson III batteries (Figure 4) [37,38]:

Spatial Relations Test – this test asks children to solve two-dimensional spatial puzzles. To receive a high score, children must be able to isolate shapes and perform mental rotation tasks.

Visuomotor Precision Test – in this activity, children are asked to follow a path while using a pencil. To receive a high score, children must employ visuomotor (hand-eye coordination) skills.

Block Construction Test – in this activity, children are shown a figure of a three-dimensional block structure and asked to build the structure using physical toy blocks. To receive a high score, children must employ visuospatial reasoning and perform fine-motor physical manipulation.

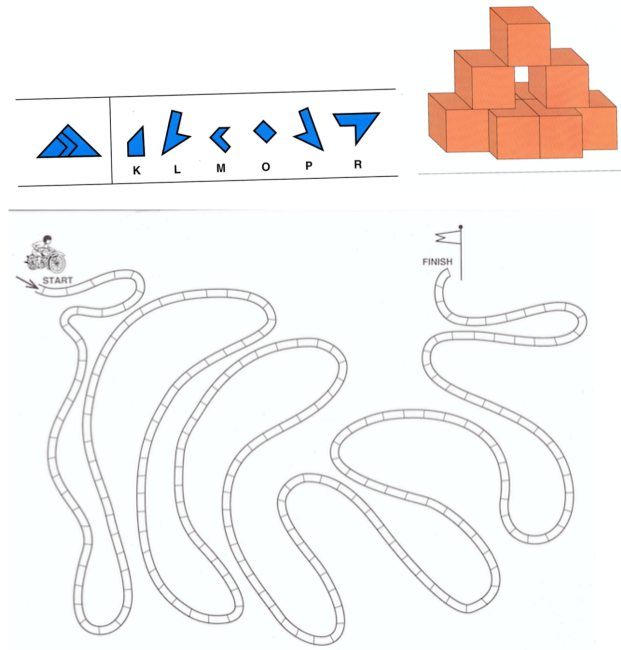

Figure 4. Developmental test samples, from Woodcock-Johnson III Spatial Relations (top left), NEPSY II Block Construction (top right) and Visuomotor Precision (bottom)

Performance Measures

To measure children's reaction to the augmented reality experience we used performance measures collected through the software, and subjective self-reported measures. These measures were calculated for each of the 4 game conditions, by averaging the 16 lemons collected by a participant within each condition.

Task Time: The average amount of gameplay time a child spent in order to collect each lemon. This metric does not include time spent during tracking loss.

Number of Tracking Losses: The average number of times that the child lost tracking. Tracking loss occurs when the phone camera stops seeing the gameboard image.

Time to Recover Tracking: The average amount of time that a player needed in order to exit tracking loss and resume playing. This amount of time was not included in the Task Time measure, and was calculated only when Number of Tracking Losses was non-zero.

Number of Selection Errors: The average number of invalid selections performed while attempting to collect a lemon. In the Finger Selection condition, an invalid selection occurs when the player touches the screen outside the target area. In the Crosshair Selection this occurs when the player clicks the selection button, but the crosshair is pointing away from the target area.

Subjective Measures

At the end of each level we asked children to report how they felt, by asking if they were Comfortable, if the level was Easy, and if they had Fun while playing. The questions were randomized between levels, and delivered using a modified rating scale (Figure 5) based on the Smiley-o-Meter [39].

Figure 5. Subjective experience questionnaire item.

Participants and Data Collection

Children were recruited from the Emory University psychology department's child subject pool. A convenience-sampling method was used, whereby we contacted families about participating in a study that investigates smartphone-based games for children.

After collecting the data, we performed outlier removal. During some trials, we observed participants stopping their gameplay due to extraneous events, such as stopping to clear runny noses, interrupting the gameplay to say something to the experimenter, etc. These interruptions would impact a participant's target collection trials, thus for the purposes of data analysis we excluded trials in which the completion time was beyond 2.5 standard deviations of a child's average times within the same experimental condition. This accounted for 3% of all trials. Furthermore, 1 child was removed from the data analysis because their age-standardized test scores were beyond 2.5 standard deviations past the mean of our sample.

The resulting dataset consists of 37 children of both genders across three age groups, whose demographics are shown in Table 1.

Age Group	Number of Children
5-6 years	17
7-8 years	11
9-10 years	9

Gender	Number of Children
Female	23
Male	14

Table 1. The demographics of our sample.

RESULTS

In this section we present our analysis and results. The results will be discussed in relation to each other and overall context in the following section.

Effects of Interaction Technique Type

Our first research question investigates the differences between various interaction technique conditions. For each dependent measure of performance we used a repeated-measures ANOVA, analyzing the within-subjects factor of Interaction Type. This factor represented a combination of two sub-factors: Selection Type (Finger vs. Crosshair) and Movement Difficulty (No Tunnel vs. Tunnel). Table 2 shows the average values for each performance metric in relation to the interaction technique conditions.

	Task Time (s)	Tracking Errors	Time to Recover Tracking (s)	Selection Errors
Finger Selection	3.69 (sd=1.92)	0.16 (sd=0.18)	1.75 (sd=1.52)	1.12 (sd=1.02)
Crosshair Selection	5.20 (sd=3.04)	0.15 (sd=0.19)	1.58 (sd=1.63)	1.21 (sd=1.48)
No Tunnels	2.87 (sd=1.75)	0.06 (sd=0.09)	1.24 (sd=1.79)	0.84 (sd=0.93)
Tunnels	6.01 (sd=3.50)	0.24 (sd=0.29)	2.10 (sd=1.63)	1.50 (sd=1.50)

Table 2. The averages of AR performance metrics, under different interaction conditions.

Selection Type: Finger vs. Crosshair

We found a significant effect of Selection Type on Task Time. Selecting targets by touching with the finger was significantly faster than selecting with the crosshair ($F(1,34)=18.55$, $p<0.001$), with an average 1.2s (29%) longer per selection. There were no significant differences between finger vs. crosshair selection in terms of number of tracking losses ($F(1,34)=0.20$, $p=0.66$), time required to recover tracking ($F(1,21)=0.42$, $p=0.52$), or number of touch errors ($F(1,34)=0.014$, $p=0.91$). Analysis of the between-subject factor Age Group yielded no significant interaction effects with the factor of Selection Type ($F(2,34)=0.99$, $p=0.38$).

Movement Difficulty: No Tunnels vs. Tunnels

The tunnel conditions were significantly worse than non-tunnel conditions on most performance measures. The tunnel conditions required significantly more time to complete compared to the no tunnel conditions ($F(1,35)=35.375$, $p<0.001$), with an average 3.14s (52%) longer per selection. The tunnel conditions caused children to have significantly more tracking losses ($F(1,34)=13.23$, $p<0.0.001$), with an average 0.18 (75%) more tracking losses per selection. Finally, in the tunnel conditions children had significantly more touch errors than in the non-tunnel conditions ($F(1,34)=10.21$, $p<0.01$), with an average difference of 0.66 (44%) more touch errors per selection. The analysis did not find significant differences between time to recover tracking loss in Movement Difficulty conditions ($F(1,18)=0.71$, $p=0.41$). Analysis of the between-subject factor Age Group yielded no significant interaction effects with the factor of Movement Difficulty ($F(2,34)=0.26$, $p=0.77$).

There was no significant interaction detected between the factors of Movement Difficulty and Selection Type ($F(1,34)=0.01$, $p=0.94$).

Effects of Age

Our second research question investigates the relationship between children's age and their performance on the dependent measures. In order to investigate this question, we first conducted simple linear regression analysis between the factors of Age (in months) and each performance measure. We followed with a repeated-measures ANOVA, with Interaction Type as the within-subjects factor, and Age Group as a between-subject factor. Finally, significant between-factor effects were followed-up with post-hoc Tukey tests to account for multiple comparisons. Table 3 shows the average values for each performance metric in relation to age groups.

	Task Time (s)	Tracking Losses	Time to Recover Tracking (s)	Selection Errors
Age 5-6	5.74 (sd=2.91)	0.22 (sd=0.21)	2.61 (sd=1.16)	1.52 (sd=1.38)
Age 7-8	3.45 (sd=0.87)	0.11 (sd=0.15)	1.00 (sd=0.89)	0.88 (sd=0.90)
Age 9-10	3.21 (sd=0.51)	0.06 (sd=0.05)	0.69 (sd=0.60)	0.85 (sd=0.52)
Overall	4.45 (sd=2.35)	0.15 (sd=0.17)	1.66 (sd=1.30)	1.17 (sd=1.11)

***Table 3.* The averages of AR performance metrics *for different age groups*.**

Age vs. Task Time

A simple linear regression showed a significant effect of Age while predicting Task Time ($F(1,35)=10.83$, $p<0.01$) indicating that as age increases, there is a decrease in time required to make a correct selection. The repeated-measures ANOVA was significant for the between-subjects factor Age Group ($F(2,34)=6.23$, $p<0.01$). Post-hoc analysis indicated that 5-6 years-old children were on average 2.29s (40%) significantly slower than compared to the 7-8 years-old children ($p<0.05$), and on average 2.53s (44%) significantly slower compared the 9-10 years-old children ($p<0.05$); however, the 7-8 years-old children were not significantly different than the 9-10 years-old children ($p=0.97$).

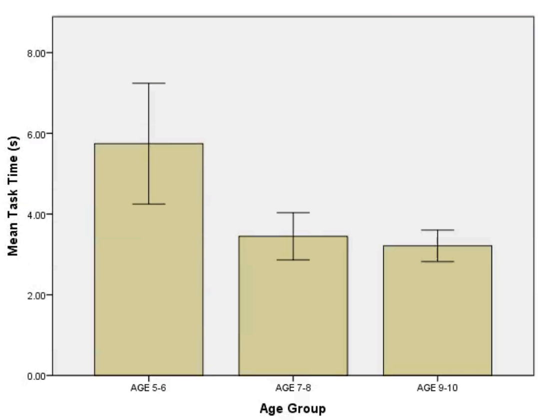

Figure 6. Mean task time for each age group.

Age vs. Number of Tracking Losses

A simple linear regression showed a significant effect of Age while predicting Number of Tracking Losses ($F(1,35)=6.44$, $p<0.05$) indicating that as age increases, number of tracking losses decreases. The repeated-measures ANOVA was significant for the between-subjects factor Age Group ($F(2,34)=3.30$, $p<0.05$); however, post-hoc analysis did not yield statistically significant differences between any age group pairs.

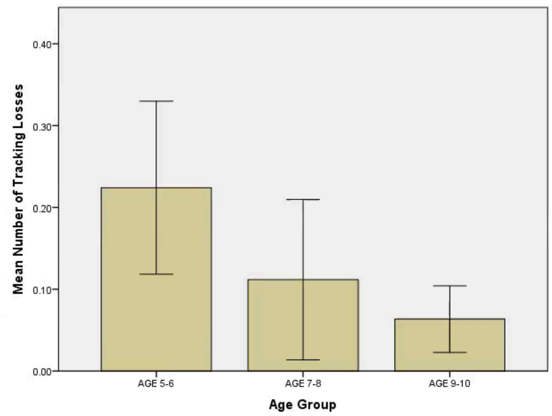

Figure 7. Mean number of tracking losses for each age group.

Age vs. Time to Recover Tracking

A simple linear regression showed a significant effect of Age while predicting Time to Recover Tracking ($F(1,31)=20.57$, $p<0.001$), indicating that as age increases, time to recover tracking decreases. The between-subjects factor ANOVA was significant for Age Group ($F(2,30)=11.40$, $p<0.001$). Post-hoc analysis indicated that the 5-6 years-old children were on average 1.61s (62%) significantly slower at recovering tracking than the 7-8 years-old children ($p<0.005$), and on average 1.92s (74%) significantly slower than the 9-10 years-old children ($p<0.005$); however, the 7-8 years-old children were not significantly different than the 9-10 years-old children ($p=0.78$).

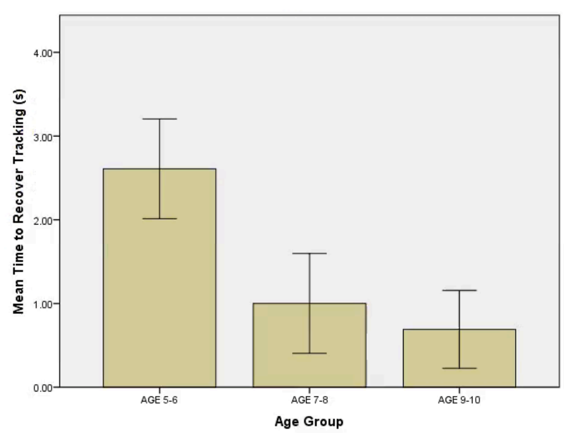

Figure 8. Mean time to recover from tracking loss, for each age group.

Age vs. Number of Selection Errors
A simple linear regression showed no significant effect of Age while predicting Number of Selection Errors ($F(1,35)=3.77$, $p=0.06$). The repeated-measures ANOVA was not significant for the between-subjects factor Age Group ($F(2,34)=1.67$, $p=0.20$).

Effect of Developmental Skills
Our second question also investigates the relationship between developmental test scores and the children's performance on the four different AR interaction technique types. We hypothesized that better developmental test scores would correspond to better AR performance, thus we performed one-tailed Pearson correlations to determine relationships between developmental tests and AR performance. Age is significantly related to all the developmental tests. Therefore, in order to control for the effects of age prior to correlational analysis, we used linear regression to remove the effect of age from the AR performance metrics, and we used age-standardized scores for our participants' developmental tests.

Spatial Relations
For the AR interaction technique condition of Finger Selection in levels involving No Tunnels, there were significant inverse correlations between the Spatial Relations test and children's Task Times ($r=-0.37$, $p<0.05$) and number of Tracking Losses ($r=-0.41$, $p<0.01$).

For the AR interaction technique condition of Crosshair Selection in levels involving No Tunnels, there were significant inverse correlations between the Spatial Relations test and children's number of Selection Errors ($r=-0.319$, $p<0.05$), number of Tracking Losses ($r=-0.36$, $p<0.05$).

Visuomotor Precision
For the AR interaction technique condition of Finger Selection in levels involving No Tunnels, there were significant inverse correlations between the Visuomotor test and children's number of Selection Errors ($r=-0.35$, $p<0.05$), and number of Tracking Losses ($r=-0.29$, $p<0.05$). For Finger Selection in levels involving Tunnels, there were inverse significant correlations between this test and the number of Selection Errors ($r=-0.32$, $p<0.05$).

For the AR interaction technique condition of Crosshair Selection in levels involving No Tunnels, there were significant inverse correlations between the Visuomotor test and children's number of Selection Errors ($r=-0.42$, $p<0.01$), and number of Tracking Losses ($r=-0.31$, $p<0.05$).

Block Construction
For the AR interaction technique condition of Finger Selection in levels involving No Tunnels there were significant inverse correlations between the Block Construction test and children's number of Selection Errors ($r=-0.29$, $p<0.05$). For Finger Selection in levels involving Tunnels, there also were inverse significant correlations between this test and the children's number of Selection Errors ($r=-0.38$, $p<0.05$).

Hand Length
After controlling for age, there were no significant correlations between this test and performance metrics on any AR interaction types.

Subjective Measures
The average results from each of the three subjective measures is reported in Table 4 in relation to each interaction type.

We performed a repeated measures ANOVA on reported degree of Fun, with the within-subjects factor of Interaction Technique condition, and a between-subjects factor of Age Group. No significant differences were detected between any factor levels.

We performed a repeated measures ANOVA on reported degree of Ease of Use, with the within-subjects factor of Interaction Technique condition, and a between-subjects factor of Age Group. Significant differences were found between the different Selection Types, with Finger Selection being ranked on average 22% more easy to use than Crosshair Selection. No significant differences were detected between age groups.

We performed a repeated measures ANOVA on reported degree of Comfort, with the within-subjects factor of Interaction Technique condition, and a between-subjects factor of Age Group. Significant differences were found between the different Selection Types, with Finger Selection being ranked on average as 5% more comfortable than Crosshair Selection. No significant differences were detected between age groups.

	Fun	Easy	Comfortable
Finger Selection	4.26 (sd=0.87)	4.34 (sd=0.66)	3.97 (sd=0.85)
Crosshair Selection	4.04 (sd=1.05)	3.93 (sd=1.03)	3.76 (sd=1.05)
No Tunnels	4.20 (sd=0.98)	4.22 (sd=0.82)	3.93 (sd=0.99)
Tunnels	4.09 (sd=1.01)	4.05 (sd=0.85)	3.80 (sd=1.03)

Table 4. The averages of participant subjective ratings, under different interaction conditions.

Effects of Gender
Our third research question is investigating gender differences in AR performance. A linear regression was attempted to predict AR performance based on Gender, while accounting for the factor of Age. No significant correlations were detected between Gender and any of the AR performance metrics or developmental test scores.

DISCUSSION

In the following section we integrate the experimental results and discuss general implications for the design of augmented reality applications for children.

Selection Types

Overall, our data indicates that the Finger Selection interaction was preferable over the Crosshair interaction. When using the Finger Selection interaction, children completed the task significantly faster, and they reported significantly higher Ease-of-use and higher Comfort levels; no other significant performance or self-report differences were found in comparison to the Crosshair interaction. However, our developmental test measures indicate that different cognitive- and physical-skills are underlying these two interaction types.

Users behave differently when an AR interface is designed with different kinds of selection types. The type of interaction mechanic causes users to grip the device differently, thus requiring users to employ different cognitive and physical skills.

The Finger Selection interaction requires children to hold the device steady primarily with one hand, while the other hand is moved such that the finger touches the screen at the appropriate target location. The analysis of children's AR performance in relation to their developmental test scores, indicated that accuracy for Finger Selection levels was significantly correlated to visuomotor skills as well as block-construction skills. This data indicates that this kind of interaction relies on the user's hand-eye coordination, as well as their ability to manipulate objects with their hands.

In contrast, the Crosshair Selection interaction requires children to aim the crosshair by precisely reorienting the device (likely by using both hands). Furthermore, in order to reorient the device toward the target, the user must have an understanding of the spatial relationships between the device and the physical gameboard. The analysis of children's developmental tests showed that accuracy on this interaction technique is correlated to visuomotor skills and spatial relations skills. This indicates that, in contrast to the Finger Selection interaction, the Crosshair Selection interaction relies more on spatial understanding and less on children's ability to manipulate objects with their hands individually.

While these different developmental skills did not appear to create significant differences between performance on the two selection conditions (except in the case of performance time), designers should be aware of these underlying skills when creating AR applications for children. During the informal observations of participants being trained in the Crosshair Selection conditions, we often have observed children pointing the screen away from the target object instead of towards it, potentially due to the fact that they did not appropriately understand the spatial relationship between the gameboard and the device. After exposure to the training level, this ceased to be a significant effect. However, this issue may indicate that Crosshair Selection interactions are problematic for children who are younger and/or less developed in spatial skills.

The form factor of the device is another issue to consider when designing AR interaction techniques. In our experiment we used a smartphone, which children could hold with both hands or a single hand. Even though children needed to hold this device in their hands at all times while playing the game, it was suitable for both interaction technique types. The Finger Selection interaction is preferable in this case, since children find it faster and more easy to use. Using this interaction can become problematic in AR applications for larger devices, such as tablets or larger smartphones. A device may be too large or too heavy for a child to hold while one hand is touching on the screen. In such a case, the Crosshair-based interaction is more suitable as it allows children to grasp the device with both hands while interacting with the application.

Dealing with Tracking Technology

Our augmented-reality game was based on Vuforia technology, which uses the phone's camera to track a paper-based printed image (the "gameboard"). The 3D game appears on the printed image when parts of the image are visible through the phone's camera, and it disappears when the image is no longer within the camera view. For our experimental smartphone (Atrix HD), the camera is placed on the top-left of the back side of the phone, as is the case with most smartphones.

Tracking loss for our participants occurred whenever the child would point the camera away from the printed image (either while standing still or moving around the gamboard), or when they covered the camera with their fingers. In the experimental condition where there were No Tunnels, children typically did not change perspective, therefore tracking losses were not caused by movement around the gameboard. Our data indicate that the number of tracking losses was significantly inversely-correlated with children's scores on spatial relations skills and visuomotor skills. In order to avoid tracking losses, children need to be aware of where the gameboard is in relation to the device, and they need good hand-eye coordination to recover in case the phone is moving away from the board or if their finger is moving in the way of the camera. Thus, children who are good at spatial and visuomotor skills encounter less tracking losses.

When playing Tunnel levels, which required walking and changes in perspective, children had a significantly higher number of tracking losses. Walking and changing perspective created more opportunities for children to aim the device away from the gameboard, and/or to put their finger in the way of the camera, thus leading to tracking loss. Our analysis did not detect any significant correlations between developmental tests and the number of tracking losses in Tunnel levels, possibly due to the fact there are more uncontrolled factors causing tracking loss, such as children's height or differing movements around the gameboard.

In order to recover AR tracking, children needed to point the phone camera at the printed gameboard image. The time it took children to recover from tracking loss was not significantly different between any of the game levels, thus it is not strongly influenced by the style of selection (Finger or Crosshair) or by the movement difficulty (Tunnel or No Tunnel). Recovering tracking appears to be a general process independent of the type of interaction in the AR application.

AR experiences can be designed to minimize the number of tracking losses experienced by players. In our AR game, losing tracking paused the gameplay, thus did not create any negative effect on the child's gameplay. However, we did notice that children became frustrated if they lost tracking frequently. Thus it is preferable if the AR experience implements features to avoid tracking loss. The AR application can be designed such that players are encouraged to be looking at the gameboard while moving (for example, if the AR application depicts a phenomenon that is interesting while being watched from changing perspectives, like a virtual prism). Furthermore, the AR technology can detect how much of the printed image is visible within the camera, therefore it can display a warning if the child is playing too close to the border of the gameboard, or if the child's finger is starting to occlude the camera.

Perspective, Accuracy and Occlusion

In levels involving No Tunnels, the game created three-dimensional spheres (the lemons) on the gameboard. From the player's default perspective in front of the gameboard, these targets were always visible and not occluded by any other game structures. In levels involving Tunnels, the lemons were encased in three-dimensional tunnels, which required participants to change perspective in order to see the lemons inside the tunnels. Compared to levels with No Tunnels, the levels involving Tunnels led to significantly longer task completion time, more selection errors and higher number of tracking losses. These kinds of effects are expected to occur in other AR experiences where players are required to change perspective around the gameboard.

In order to collect the lemons inside the tunnels, most participants changed their perspective by walking around the gameboard, and some participants bent their body while standing still. Longer task completion times in the Tunnel levels were likely caused by the fact that players had to reorient their body in order to aim at the lemons. Lower accuracy rates might also be caused by the fact that, when children's bodies are bent, it becomes more difficult to aim at a target. Furthermore, as discussed in the previous section, it is likely that moving around the gameboard also contributed highly to the number of tracking losses encountered by children in the Tunnel conditions, because tracking would be lost as children moved.

The three-dimensional structure of the tunnels was also a factor which created issues for selection accuracy. The optimal angle to collect a lemon is to look at it while being aligned with the entrance of the tunnel – this way it appears as a full sphere on the smartphone screen; however, if a player did not look from the entrance of the tunnel, the tunnel walls could occlude the lemons within, thus yielding a smaller selection area.

These effects will occur in any AR experience where players are required to change their perspective around three-dimensional content. The higher degree of inaccuracy in these experiences can be a positive factor since it adds challenge to the experience. However, if desired there are methods for making the experience easier, such as by placing guidance arrows to indicate how the player should change their perspective in order to interact with game items, or by changing the interaction technique such that targeting is more automatic once a part of the target is visible.

Rehabilitation and Skill Learning

In order to interact with an augmented reality experience, children are required to employ different developmental skills. Our analysis detected that the spatial relations, visuomotor precision, and block construction tests were correlated to different kinds of AR performance, whereby higher ability scores correlated to better performance. It is possible that through repeated exposure to augmented reality experiences, children can further develop these solicited skills. Furthermore, it is possible that augmented reality interactions can be designed to rehabilitate children who are lacking in specific skills.

Studying AR Interaction Techniques for Children

Performing interaction technique studies with children in augmented reality has posed specific challenges. Studies on non-AR touchscreen interactions have been done as highly controlled tasks, where each trial presents children with on-screen targets that are simple geometric shapes of constant size. In such 2D applications, the experimenter can control the on-screen target size, and can also require an initial finger position in relation to the target, in order to standardize the results between participants. In contrast, in augmented reality, a target is a 3D virtual object that is anchored to a physical object (the gameboard) instead of on the screen. Therefore, the target's on-screen size and position change in response to the child's distance and orientation to the gameboard; thus the size and position of targets cannot be directly controlled unless the child's movement is controlled. These factors create differences between trials within- and between- children, especially in levels where large amount of movement is required (such as in our levels involving tunnels). This likely leads to less statistical power to detect differences between experimental conditions, and creates difficulties in performing more precise studies such as Fitts' Law investigations.

CONCLUSION

In the present study we investigated how children 5-10 years old react to typical augmented-reality interaction techniques such as crosshair selection and finger selection, in AR environments that require them to change perspective or not. Our analysis showed significant impact of age upon performance, as well as differences between AR interaction technique conditions in terms of performance time, accuracy, and tracking losses. By analyzing children's performance in relation to metrics of physical and cognitive development, we also identified underlying developmental factors that influence the usability of different AR interaction techniques.

SELECTION AND PARTICIPATION OF CHILDREN

38 children aged 5-10 were recruited from the Emory University psychology department's child subject pool. A convenience-sampling method was used, whereby we contacted families and informed them of the opportunity to participate in a study that investigates smartphone-based games for children. Upon arrival at our study lab, parents and children were informed about all the study procedures and asked if they are comfortable participating in the study. This information was provided orally and through written IRB-approved consent forms. 100% of families agreed to participate; voluntary participation was recorded through signatures on the study's IRB-approved consent forms.

ACKNOWLEDGEMENTS

We wish to thank the children and parents who participated in this study; Emory University's Child Study Center and Spatial Cognition Lab; Sahithi Bonala, Andrea Lau, Katherine Guzdial, Simina Avram, and the hard-working undergraduate students who helped with this study; as well as the anonymous reviewers of this paper.

REFERENCES

1. Radu, Iulian. "Augmented reality in education: a meta-review and cross-media analysis." *Personal and Ubiquitous Computing* 18, no. 6 (2014): 1533-1543.
2. Vincenzi D. A. , Valimont B. , Macchiarella N., Opalenik C., Gangadharan S. N., and Majoros A. E. (2003). "The Effectiveness of Cognitive Elaboration Using Augmented Reality as a Training and Learning Paradigm," Annual Meeting of the Human Factors and Ergonomics Society, Denver, CO, USA, pp. 2054-2058.
3. Lindgren R. and Moshell J. M. (2011). "Supporting Children's Learning with Body-Based Metaphors in a Mixed Reality Environment," Simulation, pp. 177-180
4. Tang, A., C. Owen, et al. (2003). "Comparative effectiveness of augmented reality in object assembly." Proceedings of the conference on Human factors in computing systems - CHI '03: 73.
5. Radu, I. and B. MacIntyre (2012). Using children's developmental psychology to guide augmented-reality design and usability. Mixed and Augmented Reality (ISMAR), 2012 IEEE International Symposium on, IEEE.
6. Inkpen, K. (2001). "Drag-and-Drop versus Point-and-Click Mouse Interaction Styles for Children." Drag-Drop versus Point Click methods.
7. Inkpen, K., D. Dearman, et al. (2006). "Left‚Äêhanded scrolling for pen‚Äêbased devices." International Journal of Human‚ÄêComputer Interaction 21(1): 91-108.
8. Oehl, M., C. Sutter, et al. (2007). "Considerations on efficient touch interfaces‚Äìhow display size influences the performance in an applied pointing task." Human interface and the management of information. Methods, techniques and tools in information design: 136-143.
9. Hourcade, J. P. (2006). Learning from preschool children's pointing sub-movements. Interaction Design and Children.
10. Hourcade, J. P. (2008). "Interaction design and children." Foundations and Trends in Human-Computer Interaction 1: 277-392.
11. Joiner, R., D. Messer, et al. (1998). "It is best to point for young children: A comparison of children's pointing and dragging." Computers in Human Behavior 14(3): 513-529.
12. Vatavu, Radu-Daniel, Gabriel Cramariuc, and Doina Maria Schipor. "Touch interaction for children aged 3 to 6 years: Experimental findings and relationship to motor skills." International Journal of Human-Computer Studies 74 (2015): 54-76.
13. Anthony, L., Brown, Q., Nias, J., Tate, B., Mohan, S., 2012. Interaction and recognition challenges in interpreting children's touch and gesture input on mobile devices. In: Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces, ITS '12, ACM, New York, NY, USA, pp. 225–234.
14. Anthony, L., Brown, Q., Tate, B., Nias, J., Brewer, R., Irwin, G., 2013a. Designing smarter touch-based interfaces for educational contexts. Pers. Ubiquitous Comput. 18(6) (November), pp. 1471–1483.
15. Arif, A.S., Sylla, C., 2013. A comparative evaluation of touch and pen gestures for adult and child users. In: Proceedings of the 12th International Conference on Interaction Design and Children, IDC '13, ACM, New York, NY, USA, pp. 392–395.
16. Brewer, R., Anthony, L., Brown, Q., Irwin, G., Nias, J., Tate, B., 2013. Using gamification to motivate children to complete empirical studies in lab environments. In: Proceedings of the 12th International Conference on Interaction Design and Children. IDC '13. ACM, New York, NY, USA, pp. 388–391
17. McKnight, L. and Cassidy, B., 2012. Children's interaction with mobile touch-screen devices: experiences and guidelines for design. Social and Organizational Impacts of Emerging Mobile Devices: Evaluating Use: Evaluating Use, p.72.
18. McKnight, L. and D. Fitton (2010). Touch-screen technology for children: giving the right instructions and getting the right responses. Proceedings of the 9th International Conference on Interaction Design and Children, ACM.
19. Romeo, G., S. Edwards, et al. (2003). "Touching the screen: Issues related to the use of touchscreen technology in early childhood education." British Journal of Educational Technology 34(3): 329-339.

20. Nischelwitzer, A., F.-J. Lenz, et al. (2007). "Some Aspects of the Development of Low-Cost Augmented Reality Learning Environments as Examples." Universal Access in Human-Computer Interaction.: 728-737.
21. Hornecker, E. and A. Dünser (2009). "Of pages and paddles: Children's expectations and mistaken interactions with physical-digital tools." Interacting with Computers 21(1-2): 95-107.
22. Dunleavy, M., C. Dede, et al. (2009). "Affordances and Limitations of Immersive Participatory Augmented Reality Simulations for Teaching and Learning." Journal of Science Education and Technology 18(1): 7-22.
23. Piaget, J. (1970). "Science of education and the psychology of the child. Trans. D. Coltman."
24. Rosser, R. A. (1994). Cognitive development: Psychological and biological perspectives, Allyn and Bacon Boston.
25. Bruckman, A. and A. Bandlow (2002). Human-computer interaction for kids. The human-computer interaction handbook, L. Erlbaum Associates Inc.
26. Wyeth, P. a. P., H.C. (2003). "Using developmental theories to inform the design of technology for children." Proceedings of the 2003 conference on Interaction design and children: 93-100.
27. Gelderblom, H. (2009). Designing technology for young children: guidelines grounded in a literature investigation on child development and children's technology. Doctor of Philosophy, University of South Africa.
28. Bekker, T. and A. N. Antle (2011). "Developmentally situated design (DSD): making theoretical knowledge accessible to designers of children's technology." Proceedings of the 2011 annual conference on Human factors in computing systems: 2531--2540.
29. Crucian, G. P. and S. A. Berenbaum (1998). "Sex differences in right hemisphere tasks." Brain and cognition 36(3): 377-389.
30. Geary, D. C. and M. C. DeSoto (2011). "Sex Differences in Spatial Abilities Among Adults from the United States and China." Evolution and Cognition 7(2): 172-177.
31. Voyer, D., A. Postma, et al. (2007). "Gender differences in object location memory: A meta-analysis." Psychonomic Bulletin & Review 14(1): 23-38.
32. Johnson, E. S. and A. C. Meade (1987). "Developmental patterns of spatial ability: An early sex difference." Child Development: 725-740.
33. Kerr, R. (1975). "Movement control and maturation in elementary-grade children." Perceptual and Motor Skills 41(1): 151-154.
34. Vasta, R., K. G. Regan, et al. (1980). "Sex differences in pattern copying: Spatial cues or motor skills?" Child Development: 932-934.
35. Boring, S., Ledo, D., Chen, X.A., Marquardt, N., Tang, A., Greenberg, S., 2012. The fat thumb: using the thumb's contact size for single-handed mobile interaction. In: Proceedings of the 14th International Conference on Human–Computer Interaction with Mobile Devices and Services. MobileHCI '12. ACM, New York, NY, USA, pp. 39–48.
36. Hohendorff, B., Weidermann, C., Burkhart, K.J., Rommens, P.M., Prommersberger, K.J. and Konerding, M.A., 2010. Lengths, girths, and diameters of children's fingers from 3 to 10 years of age. Annals of Anatomy-Anatomischer Anzeiger, 192(3), pp.156-161.
37. Kemp, S.L., Korkman, M. and Kirk, U., 2001. Essentials of NEPSY assessment (Vol. 6). John Wiley & Sons.
38. Woodcock, R.W., 1990. Theoretical foundations of the WJ-R measures of cognitive ability. Journal of Psychoeducational Assessment, 8(3), pp.231-258.
39. Read, J., MacFarlane, S. and Casey, C., 2001. Measuring the usability of text input methods for children. In People and Computers XV—Interaction without Frontiers (pp. 559-572). Springer London.