

Exploring Expressive Augmented Reality: The FingAR Puppet System for Social Pretend Play

Zhen Bai, Alan F. Blackwell, George Coulouris

Computer Laboratory, University of Cambridge
15 JJ Thomson Avenue, Cambridge, CB3 0FD, United Kingdom
zhen.bai, alan.blackwell, george.coulouris@cl.cam.ac.uk

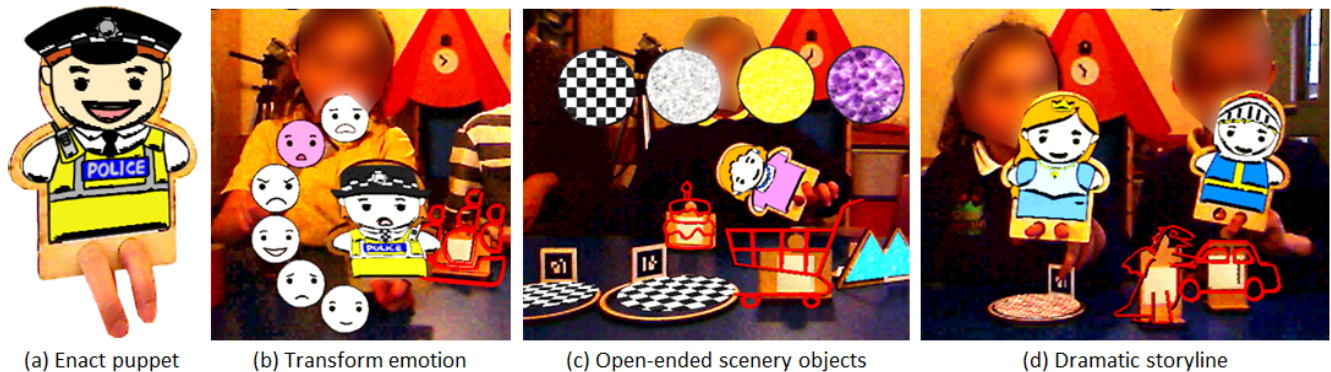


Figure 1. FingAR Puppet system overview.

ABSTRACT

We present “FingAR Puppet”, an Augmented Reality (AR) system enhancing social pretend play by young children. Unlike goal-oriented AR systems that augment reality with informative instructions, FingAR Puppet helps children associate expressive interpretations with immediate reality. Empirical results show that FingAR Puppet promotes reasoning about emotional states, communication and divergent thinking during social pretend play for children 4-6 years old. We suggest that this study opens an interesting space for future AR systems to support complex cognitive and social development in early childhood. We also identify broader implications from using theories of cognitive development to guide the design of tangible and augmented interactions.

Author Keywords

Augmented Reality; Tangible User Interface; Pretend Play; Children.

ACM Classification Keywords

H.5.1. Information Interfaces and Presentation: Multimedia Information Systems Artificial, augmented, and virtual realities

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
CHI 2015, April 18–23, 2015, Seoul, Republic of Korea.

Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-3145-6/15/04 ...\$15.00.
<http://dx.doi.org/10.1145/2702123.2702250>

INTRODUCTION

In this paper we describe an application of the Augmented Reality (AR) paradigm that explores the capacity for AR to support complex cognitive skills. Many early research and current commercial AR applications have emphasised the value of AR for guidance and information finding in domains such as maintenance, medicine, navigation and more [37]. However, in addition to such goal-oriented tasks, AR technologies may support more complex cognitive skills. Two specific examples are divergent thinking [30], where it is necessary to think about multiple possible courses of action, and theory of mind [24], where it is necessary to think about other people’s states of mind. Both of these cognitive skills are essential in early childhood development.

The developmental stages through which children shift from applying knowledge in situations with high perceptual similarity, to more general expressive and imaginative cognitive tasks, underpin essential skills in everyday life [4]. We suggest that the potential for AR to superimpose interactive virtual information onto immediate reality [1] extends beyond typical guidance and information finding, to offer support for more complex cognitive skills. This may be especially beneficial for children who have difficulty moving beyond interpretations based on high perceptual similarity. In particular, our earlier work indicates that AR can encourage children with autism to use objects more symbolically in play [2], which is closely linked with theory of mind and divergent thinking.

In this new study we developed the FingAR Puppet system to promote theory of mind and divergent thinking through social pretend play. The system employs a magic mirror

metaphor to portray an imaginary AR world in which users can see themselves and their physical environment extended with augmented cues and content that bridge from immediate perception to imaginary play. These augmentations support symbolic cognition (e.g. pretending a banana is a telephone) and social cognition (e.g. constructing joint pretense among multiple players) [23, 35]. These social pretend play skills enable children to simulate and respond to acted emotions, beliefs and desires, communicate interpretation of roles, props and events [16, 17], and flexibly use play materials in response to make-believe themes [28]. These cognitive and social abilities are critical preparation for later school life [5].

We present results from an evaluation of the system in extended use by children 4-6 years old. We find that FingAR Puppet effectively promotes reasoning about emotional states, communication and divergent thinking during social pretend play. These findings open an interesting space for future AR systems to support complex cognitive and social development in early childhood. They also suggest the potential for AR to be used in a wider range of expressive and interpretive tasks. Finally, we discuss the usability implications of the magic mirror metaphor, as an approach to AR design that offers particular advantages for physical object manipulation and social interaction.

RELATED WORK

There are several interactive systems intended to assist child development in storytelling, which shares elements with social pretend play. Although various physical objects are involved in these systems, they either remain literal without any symbolic transformation (e.g. Storymat [31], Augmented Knight's Castle [10], video puppetry [3]) or simply function as tangible entities to help children interact with virtual objects more efficiently (e.g. ShadowStory [18], Puppettime [22], wIzQuebes [38]). There are other interactive systems aiming to promote children's emotion understanding. For example, StoryFaces [32] is a touch-pad based storytelling system that allows the child to watch pre-made stories involving virtual characters with his/her face, and create their own stories. It lacks support for reasoning about emotional states from other's perspectives, and reciprocal social interaction through improvisational role enactment. PUPPET [21] is a VR environment with autonomous characters with different goals and emotional states. While findings show that PUPPET help children aged 7-9 to understand characters' emotional states, further exploration is yet needed for younger children.

SYSTEM DESCRIPTION

The design of the FingAR Puppet system had two related goals: (1) to provide a complete social pretend play experience; and (2) to more specifically enhance children's emotion expression and understanding, joint pretense and divergent thinking in the context of this experience. We applied the magic mirror display metaphor [20] to support a reflected view of reality that superimposes imaginary objects and situations. The children using the system look into this magic mirror while interacting with physical objects alongside other players in a tabletop play environment (Fig.2).



Figure 2. Children interact with the FingAR Puppet system.

Design Rationale

The cognitive mechanism of pretend play is an early manifestation of theory of mind [16]. As an example of this relationship, consider these play scenarios:

Eileen pretends that the puppet is a policeman

The primary representation of reality (a puppet) is transformed via a symbolic relationship to support a secondary representation of pretense ("the puppet is a policeman"). In FingAR Puppet, this symbolic transformation is externalised through the augmentation of the play context with a pretense stimulus seen in the magic mirror.

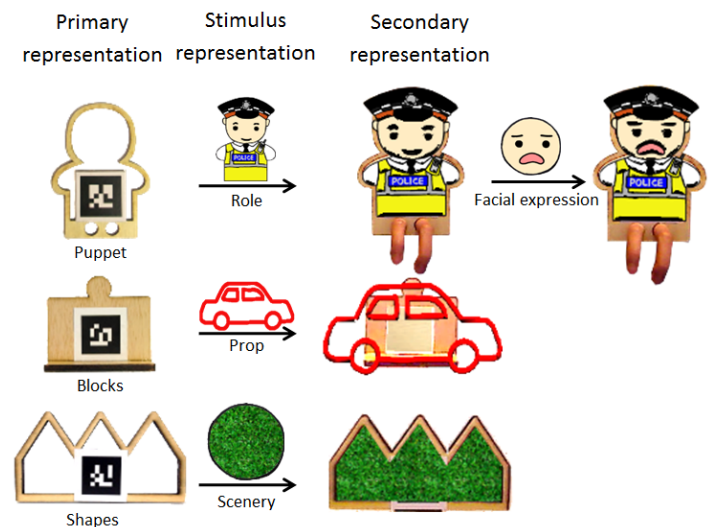


Figure 3. The FingAR Puppet design concept.

As shown in Fig.3, the primary representation of reality in the FingAR Puppet system contains physical objects: puppets, blocks and generic shapes. The stimulus representation corresponding to each type of physical referent are role (puppet), prop (block) and scenery (shapes). The design intention of these stimuli is that they should encourage the user to carry out symbolic transformation involving gradually increased mental effort. We included two types of stimuli: (1) roles and props to support specific symbolic transformations;

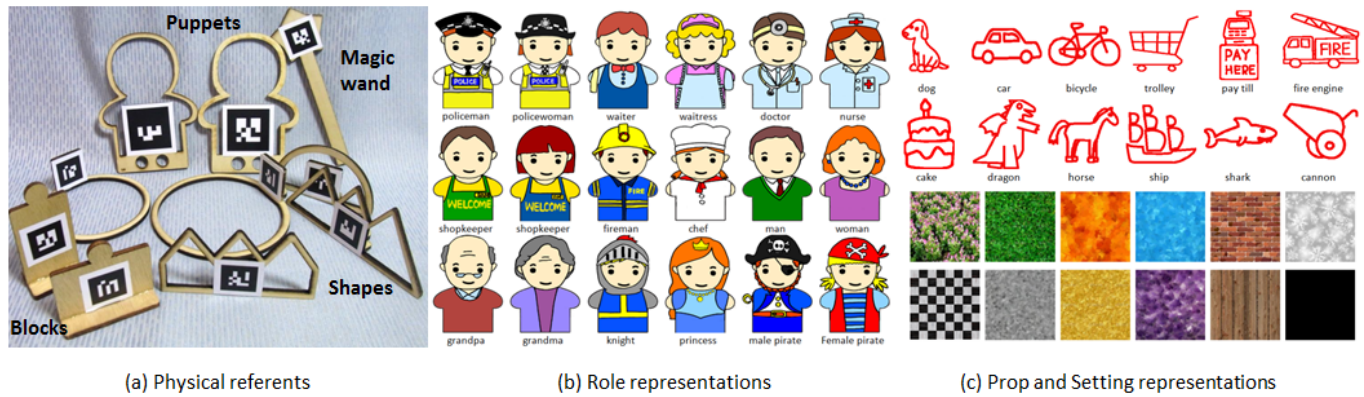


Figure 4. The collection of primary and stimulus representations.

(2) open-ended scenery to encourage divergent exploration [29], in which users construct their own meanings. Symbolic transformation is observed when the user assigns one of the stimulus representations to a physical referent. For example, when the user assigns the policeman role to the puppet, the secondary representation “the puppet is a policeman” was created. Similarly with emotion transformation, by assigning the “scared” facial expression to the policeman, the policeman’s emotional state would be transformed to “scared”.

The system supports two phases of activity: preparation and play. In the preparation phase, children set up roles, props and scenes. Each child has one puppet, several blocks, and several shapes. In the play phase, children make stories together using the AR elements they have chosen. During the play phase, children are also able to (1) change the facial expression of their puppets; (2) change role, prop or scenery whenever necessary; and (3) record a video of the play.

Interaction Design Details

Object Transformation

The collection of physical referents (puppet, block and shape) and stimulus representations (role, prop, and scenery) are shown in Fig.4.

Role/Puppet We chose to base the FingAR puppet concept on the finger-leg puppet, a familiar cheap toy for young children. Our main criterion was simplicity of operation – this form of puppet is more mobile and expressive than dolls, and easier to operate than other forms of puppet such as glove puppet, marionette and shadow puppet. The system includes 18 social and fictional based roles, all sharing the same pose and facial features (Fig.4 (b)). Based on common play themes of children’s social pretend play [33], the system supported several role-specific play themes including rescue (e.g. policeman and fireman), restaurant (e.g. chef and waiter), shopping (e.g. shopkeeper), hospital (e.g. doctor, nurse), and adventure (e.g. princess, knight, pirate). We maintained a balanced number of male and female roles [11].

Prop/Block We provided wood blocks in a generic rectangular shape as the physical referent for props. The grip on top of the block was meant for the child to hold when selecting and moving the prop, in order to prevent marker occlusion. There

were 12 virtual objects as the target for prop transformation (Fig.4 (c) top). These were rendered in a wireframe style and related to the specific themes of the pre-defined roles, such as fire engine with the rescue theme, cake with the dining theme, dragon with the adventure theme, etc. The child could flip the orientation of the prop by flipping the associated wood block. The wireframe style was partly motivated by our intention that in future, users should be able to draw their own target props while maintaining visual consistency with the existing props.

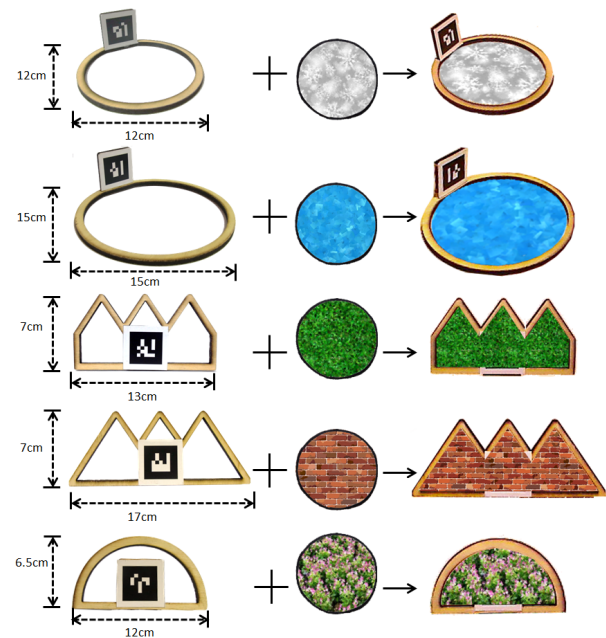


Figure 5. Scenery transformation.

Scenery/Shape We chose 12 open-ended materials to help children create meanings for physical referents (Fig.4 (c) bottom). Some materials reflected natural scenes (e.g. flower, grass, water, snow, wood and stone) or built environment scenes (e.g. brick, tile and fire). Others had more ambiguous meanings, since familiarity with stimuli affects divergent thinking processes [29]. We wished to encourage both (1) creation of different pretense interpretations using familiar

materials; and (2) creation of more novel symbolic interpretations using unfamiliar materials. Materials could be associated with three types of wood shape: circle, semi-circle and sawtooth. Slight variations in size and shape were used to avoid possession conflict between players. Example combinations of shape and material are shown in Fig.5.

Transformation Action The tangible nature of the physical referent allows children to externalize their mental process of object transformation through direct manipulation [34], as demonstrated in Fig.6. To complete an object transformation, the child points the physical referent toward the target role, prop or scenery in the ‘mirror’. The corresponding virtual object is then attached to or combined with the physical referent in the augmented view.

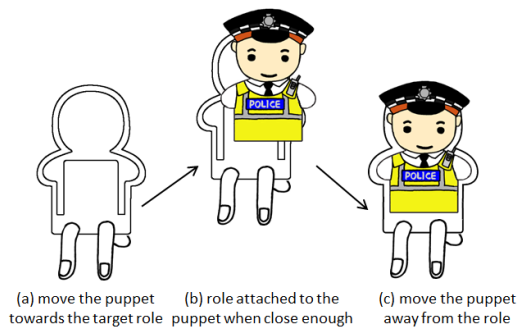


Figure 6. The action of role transformation

Emotion Transformation

Facial expression visual form We designed six facial expressions: five basic emotions (fear, surprise, happiness, anger, sadness) [9] and a neutral expression. The corresponding visual effects are seen in Fig.7.

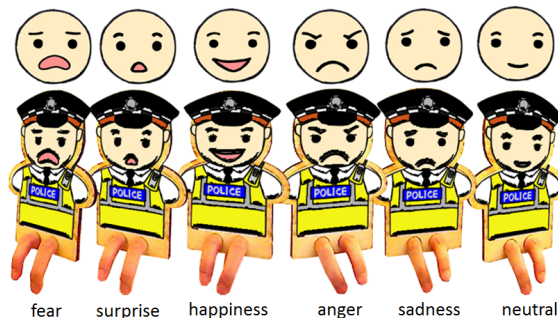


Figure 7. The collection of facial expressions

Facial expression switch There are two important usability considerations for facial expression switching. First, the function should be easy to access, motivating children to explore this feature often in addition to other play objects. Second, the switch should be rapid to support fluency of story telling. We investigated several different expression switching methods, eventually adopting a wand-based approach (Fig.8). The user moves the wand towards the puppet to trigger an expression-switching mode. Alternative expressions are presented as a radial menu around the puppet. The user

points the wand at the desired expression, then moves the wand away from the puppet to dismiss the radial menu.

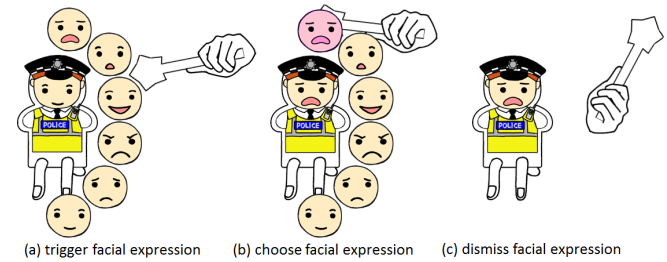


Figure 8. Action flow of wand-based switch

SYSTEM IMPLEMENTATION

We developed the FingAR Puppet system based on several open source libraries including the Microsoft XNA Game Studio 4.0 (system framework), GoblinXNA 4.1 (AR registration and rendering), ALVAR2.0 (marker tracking) and Emgu CV2.4 (image processing). We created the 3D role models from 2D images because: (1) a 2D image is easy to generate and it potentially enables user authoring in the future; and (2) the shadow details produced by the 3D model enrich the visual illustration of the role and facial expression. We implemented a “rise up” 3D effect similar to baking cookies which involved three major steps (Fig.9): (1) generate a distance transform image [26] of the original 2D image; (2) create the 3D mesh model based on the distance transform image; (3) map the original image as texture to the 3D mesh.

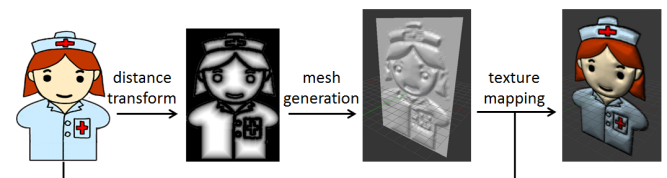


Figure 9. Implementation pipeline of the “rise up” 3D effect.

EXPERIMENT DESIGN

We designed an experiment to evaluate the effectiveness of the FingAR Puppet system in promoting reasoning about emotional states, verbal communication of shared pretense and divergent thinking. There are three null hypotheses in respect of the first three aspects:

H_{0A} : There is no significant difference in the frequency of emotional state expression between conditions with facial expression switching disabled and enabled.

H_{0B} : There is no significant difference in the frequency of causal elaboration of emotional state between conditions with facial expression switching disabled and enabled.

H_{0C} : There is no significant difference in the percentage of explicit verbal communication per object transformation between scenery selection period and role and prop selection periods.

Method

Each pair of participants interacted with the FingAR Puppet system in two conditions: facial expression switching disabled (Face_Switch_OFF) and enabled (Face_Switch_ON). This within-subjects design helped to avoid individual differences on social pretend play behaviours. The condition order was counterbalanced in order to eliminate learning and novelty effects. In both conditions participants were asked to make stories with the system for 15 minutes.

Participants

Fourteen participants from 48 to 74 months ($M = 63.00$, $SD = 8.79$) were recruited from a local primary school, eight girls and six boys. All parents of the participants signed a consent form and provided basic information about their child. Teachers paired participants into seven groups based on familiarity with each other (three mixed-gender groups, two girl-only groups and one boy-only group). Three participants used computer devices on a daily basis, and all others at least once a week. The verbal mental age of participants was from 58 to 77 months ($M = 67.14$, $SD = 7.23$) based on the British Picture Vocabulary Scale, 3rd edition. We did not record or select participants' ethnic background.

Apparatus and Data Collection

The study took place in a common room between the nursery and reception classes during normal school hours. The setup of the AR system included: (1) a Macintosh Laptop; (2) a 24-inch monitor; (3) a Logitech Webcam Pro 9000; (4) a 55*110*50cm table and (5) play materials (Fig.10 (a)). We used a video camera to record play behaviour during each session (Fig.10 (b)). Video footage from days two and three was transcribed into discrete play acts (by editing with Camtasia Studio), identifying all individual actions and speech acts.

Procedure

We conducted the experiment on three consecutive days with the support of three teaching assistants. On the first day, the experimenter taught all participants, one group at a time, how to use the AR system and let them explore the system freely for about 10 minutes. On the second day, the participants were asked to carry out play in pairs for 15 minutes in either the Face_Switch_OFF or Face_Switch_ON condition. At the beginning of each session, the experimenter gave identical instructions: "you are going to make a story together and you can play for 15 minutes." During the session, the experimenter and the teacher provided minimal prompts only when the participants were obviously not engaged in the play or had difficulty developing play ideas independently. After 15 minutes, the experimenter waited for the current play episode to finish and asked the participants to stop. The experimenter then went through a questionnaire with each participant. This procedure was repeated with the other experimental condition on the third day.

Measures

We adopted several measures from existing literature in early childhood development to design our coding scheme for emotional state [36], causal elaboration [7], and verbal communication on object transformation [12].



(a) interaction area

(b) recording area

Figure 10. The experiment apparatus.

Emotional state Indicators of emotional states include: (1) verbal terms (e.g. happy, sad, angry, etc.); (2) behavioural terms (e.g. cry, kiss, hug, etc.); (3) tone of voice (e.g. angry voice). We counted the total number of emotional states that occurred within the play acts of each participant. If the same emotional state occurred more than once in the same play act, it was only counted once. If there was more than one different emotional state in the same play act, each one was counted independently. We divided the total number by the play time (minutes) in both conditions to calculate the frequency of emotional state occurrence. In order to make a rigid comparison between the two conditions, we excluded the facial expression switching action in the Face_Switch_ON condition from emotional state occurrence. The experimenter made the initial coding. We then randomly selected a 4-minute clip from each video (23% of the total video footage) and invited an independent rater who was not aware of the hypotheses to code the emotional state occurrence. The inter-subject agreement is highly satisfactory (Cohen's kappa = 0.98).

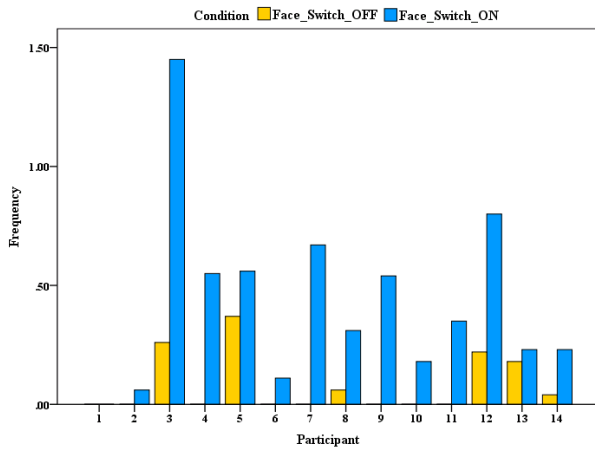
Causal elaboration of emotion Causal elaboration was identified as a pair of verbal or gestural play acts that revealed the cause and effect relationship of certain emotional states. Within the pair, at least one of the play acts contains emotional state. For example:

- (1) child A said: "I'm *angry* now" [effect]
child B said: "because there are no customers" [cause]
- (2) child A said to child B: "happy birthday" [cause]
child B made his puppet *kiss* child A's puppet and said: "thank you" [effect]

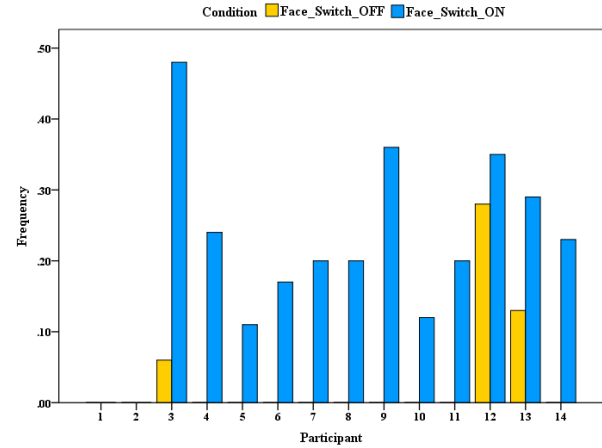
The causal elaboration could also be within one play act. For example:

- (1) child A said "I'm *sad* [effect] because my dog went missing [cause]"

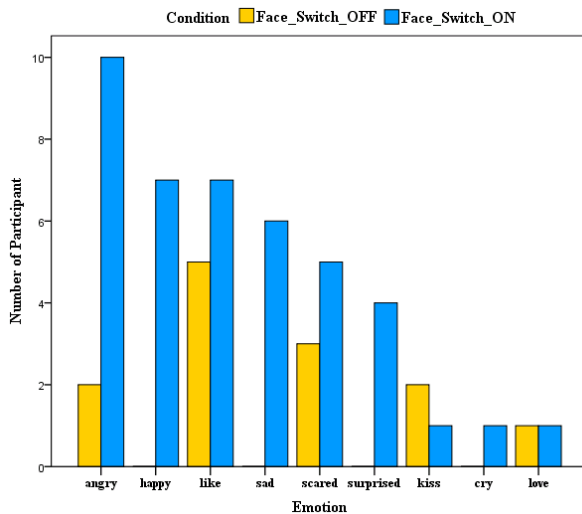
For each play act containing emotional state, we (1) checked if there was any causal elaboration within this play act; (2) checked all play acts 10 seconds before and after. We counted the total number of causal elaborations and divided by the play time of each group to calculate the frequency of cause elaboration. The experimenter made the initial coding, with two independent raters as before. The results demonstrated a high inter-subject agreement (average Cohen's kappa = 0.92).



(a) Frequency of emotional state for each participant

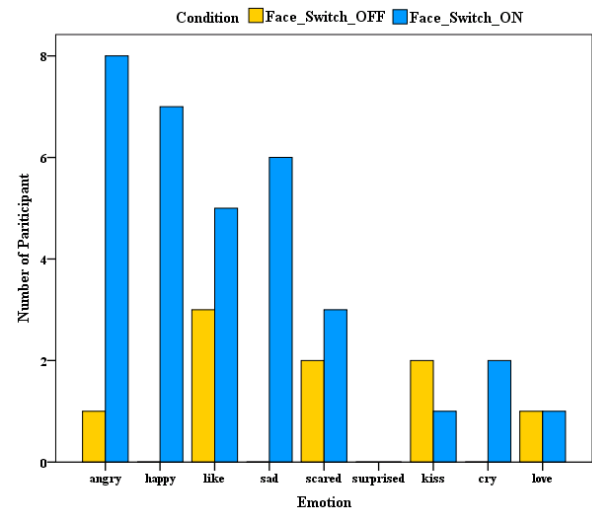


(a) Frequency of causal elaboration for each participant



(b) Number of participants using each emotion term

Figure 11. Results of emotional states.



(b) Number of participants explaining each emotion term

Figure 12. Results of causal elaboration of emotion.

Verbal communication on object transformation We calculated the percentage of object transformations with explicit verbal communications during the role and prop selection period and scenery selection period. The coding steps were: (1) extracted the selection zones, which began with entering role/prop/scenery selection view and ended when exiting the selection view; (2) counted the total number of selection actions within all selection zones, and (3) counted the total number of explicit verbal communications that clarify the target of object transformation (e.g. “I’m a knight”, “grass”); (4) calculated the percentage of selection actions associated with explicit verbal transformation communications.

RESULTS

Emotional States

The frequency of emotional state occurrence per participant increased in the Face_Switch_ON condition ($Mean = 0.43$, $SD = 0.37$) relative to the Face_Switch_OFF condition ($Mean =$

0.08 , $SD = 0.12$). The frequency of occurrence for each participant is illustrated in Fig.11(a). The trend of individual difference is statistically significant according to the paired Wilcoxon signed-rank test ($Z = -3.18$, $p < 0.01$). The order of participants follows the group order (e.g. participant 1 (P1) and participant 2 (P2) were in group 1). All participants produced play acts involving emotional state more frequently in the Face_Switch_ON condition except for P1, who didn’t generate any emotional state related play behaviour in either condition. We summarized the number of participants using each emotion term (Fig.11(b)). Besides the five non-neutral facial expressions provided by the system, other emotional terms used by the participants included like, kiss, cry and love.

Causal Elaboration of Emotion

The frequency of causal elaborations per participant was higher in the Face_Switch_ON condition ($Mean = 0.21$, $SD = 0.13$) than the Face_Switch_OFF condition ($Mean = 0.03$, $SD = 0.07$). The difference is statistically significant according to the paired Wilcoxon signed-rank test ($Z = -3.06$, $p < 0.01$).

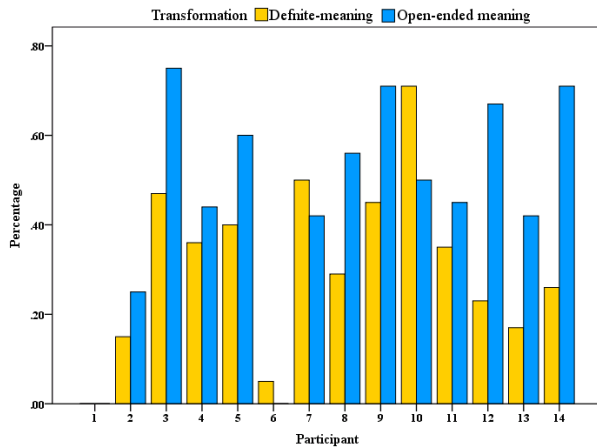


Figure 13. Percentage of verbal communication on transformation per participant.

Fig.12(a) illustrates the frequency of causal elaboration generated by each participant. There were only three participants that made some causal elaboration in the Face_Switch_OFF condition while there were twelve participants that produced causal elaboration in the Face_Switch_ON condition. We summarized the number of participants that produced play acts involving causal elaboration of different emotion terms (Fig.12(b)). It followed a similar trend as the number of participants that produced play acts involving emotional states, except no participant managed to produce any causal elaborations relating to the emotion term “surprised”. The implications of the results will be explained in the discussion session.

Verbal Communication on Transformation

Participants made more verbal communications on object transformation when choosing from open-ended representations in the scenery selection period ($Mean = 0.46$, $SD = 0.24$) than the definite-meaning in the role and prop selection period ($Mean = 0.31$, $SD = 0.19$). The difference is statistically significant according to the paired Wilcoxon signed-ranks test ($Z = -2.45$, $p < 0.05$) (Fig.13). Participants carried out diverse scenery transformations. Table 1 shows representative examples of scenery transformation based on verbal reports.

Child Questionnaire

The FingAR Puppet system offered a positive user experience – the majority of participants rated the play experience as ‘brilliant’ in both conditions (10/14 in Face_Switch_ON and 8/14 in Face_Switch_OFF condition). When asked which play variant was more fun, all participants named the condition with facial expression switching enabled. Some representative explanations were: “it can be sad, happy, that helps to create the story”, “when the character sees a shark I can change it to shocked”, “when it’s just one face, when you’re angry you cannot change the face”. Participants also confirmed that the interaction structure helped create stories: “pick character, second part (prop) whatever you want, next you choose floor, setting”, “when change people, you make different story ... like change the setting”.

Table 1. Examples of scenery transformation.

| | Circle | Semi-circle | Sawtooth |
|-------------|--|----------------|-----------------------|
| Green | grass, flower | grass | grass, park* |
| Red | fire | fire | fire |
| Green-red | flower | | flower |
| Blue | swimming pool, water, sea, blue sky, big lake (with two circles combined)* | | blue sky, water |
| White | ice rink, snow, water | snow, mountain | snow, ice-berg castle |
| Brick | castle | | |
| Black-white | racing flag, floor, cafe*, shield*, nurse house* | | |
| Rock | | | |
| Golden | pancake, gold water* | | |
| Purple | | diamond | |
| Wood | | path | |
| Black | black hole* | | |

* Symbolic ideas verbally explained in the latter play session

DISCUSSION

Emotion Expression and Understanding

These results demonstrate that the FingAR Puppet system effectively encourages children to express and understand emotion in a social play context (rejecting null hypotheses H_{0A} and H_{0B}). When the facial expression switching was enabled, users not only produced more play acts relating to emotional states, but also explained the cause and effect relationship of emotion more frequently, providing opportunities for them to reason about and respond to emotion change. Overall, participants were more likely to be emotionally aware and expressive during social pretend play when given the ability to switch the facial expression of the puppet.

The emotion terms most often used in the play explanations (Fig.11(b)) can be classified as desire-based (angry, happy, like, sad), or belief-based (scared and surprised) [27]. Understanding desire-based emotions occurs as early as two years old, while the ability to interpret people’s beliefs is typically not achieved until about four years old. Our results confirm that participants were more likely to express desire-based than belief-based emotions, which corroborates the existing literature. No participants gave explanations of “surprise” confirming previous findings that children take longer to understand the complex belief-based nature of surprise [19].

Verbal Communication of Transformation

Results show that participants verbally explained their object transformation decisions during role, prop and scenery selection. Participants made more verbal explanations in relation to the open-ended representations of the scenery selection period, rather than the definite-meaning representations of the role and prop selection period. This supported our hypothesis (rejecting null hypotheses H_{0C}) that participants tend to assure both themselves and their playmates about the transformation they made by verbally clarifying the target object

of the transformation when associating an open-ended representation to the physical referent.

Diverse Scenery Transformation

The result that participants generated diverse imaginary representations associated with the shapes in the scenery selection period corroborate previous research findings that open-ended tasks are effective in promoting divergent thinking [29]. We observed that participants tended to generate more imaginary representations with familiar than less familiar scenery materials. For example many participants created grass, water and flowers by assigning green, blue and green-red materials to shapes respectively. It was, however, rare to see participants creating imaginary representations beyond these obvious associations. Furthermore, participants tended to generate more novel imaginary representations with less familiar materials. For example, they created diamonds and black holes using purple and black materials respectively. They were also more likely to interpret the same AR scenery objects differently with less familiar materials. For example, by assigning the black and white material to the circle referent, participants generated different imaginary representations such as racing flag, floor, cafe, shield, and nurse house. The above observations were well explained by the theory that familiar stimuli facilitate fluency of divergent thinking while unfamiliar stimuli support originality of divergent thinking. The observations also corroborate with recent findings that allowing users to customize play scenes enriches creativity and communication in shared play activities, from a study investigating creative play between co-located users via video mediated communication [15].

Usability

Usability is a challenge in AR systems for children due to their limited cognitive and motor abilities [25]. Nevertheless, we observed that all our participants were able to interact with the FingAR Puppet system in terms of object manipulation, selection and facial expression change.

Object Manipulation

All participants manipulated the physical puppet properly by putting their fingers through the two holes on the bottom of the puppet, and acted with the puppet by moving the “legs” around. Only one participant tried to make the puppet stand on its own. For props and scenes, participants held them either with the attached grip or the body. The latter sometimes caused occlusion of the marker.

Selection

All participants understood the point-to-select mechanism for role, prop and scenery, and used the wand to choose control buttons. However, we noticed two issues with selection in the preparation phase. The first was that participants occasionally mismatched object types, for example using the prop representation object to choose scenery. Although this was infrequent, a simplified selection mechanism could be explored in future. The second issue was that some younger participants reached out towards the screen when choosing role, prop or scenery items, rather than using the world coordinate system “reflected” by the magic mirror (Fig.14).

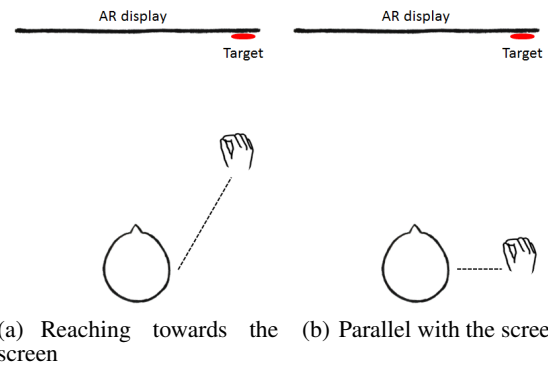


Figure 14. Illustration of participant's point-to-select behaviours using the screen as the coordinate system.

Facial Expression Change

Participants understood how to use the wand to change facial expressions. Nevertheless, this was more difficult than other actions in the FingAR Puppet system. This can be explained by the demands of bimanual manipulation. According to the kinematic chain model [13], the non-dominant hand often sets a spatial frame of reference within which the dominant hand orients fine temporal and spatial movements (e.g. in handwriting or sewing). Since participants held the puppet with their dominant hand, this meant that they had to position the wand with their non-dominant hand.

Study Limitations

There are three limitations to be aware of when interpreting these results. First, although participants produced play acts with emotional states more frequently in the Face.Switch.ON condition, these emotional states were mostly desire-based, meaning that the FingAR Puppet system may be more limited in social pretend play for complex emotions. Second, the study only examined children interacting in pairs. The limited size of the interactive space and the side-by-side seating arrangement using the AR magic mirror might become an issue with more objects or additional players. Third, this was a short-term experiment with a small sample. Although the familiarisation session on the first day was meant to eliminate learning and novelty effects, a longer-term study would be required to explore more complex emotional states and creative symbolic transformations.

Future Work

As discussed, we hope to extend our explorations into longer-term development of social-emotional knowledge, and more complex emotions such as surprise, embarrassment and pride that are especially relevant to older age groups. We would like to examine potential benefits of the FingAR Puppet system for children with autism, who often lack imaginative and social interaction skills. We would also like to observe children's different social symbolic play behaviours between using the FingAR Puppet system and in an equivalent natural play setting without computer assistance, in order to further our knowledge of the strengths and weaknesses of AR technologies in play contexts.

CONCLUSION

We presented the theoretical grounding, design and evaluation of the FingAR Puppet system, which is intended to provide an enhanced social pretend play experience for preschool children. Observations show that children are highly engaged with the FingAR puppet system. Experiment results confirm that (1) AR technology can effectively externalize the cognitive process of pretense by enabling children to transform physical referents to pretense ideas; (2) facial expression switching can encourage children to express and understand emotion in stories; and (3) open-ended representations encourage children to create diverse symbolic transformations, and to more actively communicate such ideas with a playmate as joint pretense.

This study opens an interesting space for future AR systems to support users to augment the real world with their expressive interpretation, in addition to perceiving informative augmentations that help users to obtain reality-based knowledge. Such expressive interpretations are intended to be open-ended and interpersonal, thus in turn may help reinforce key cognitive processes that shape peoples thinking pattern. Compared with technologies isolated from the immediate reality, AR helps bridge experience between computer-assisted and real life scenarios. This is especially beneficial for young children who are more capable to apply skills to situations with high perceptual similarity.

Beyond the immediate application encouraging childhood development of social pretend play, this study also points to broader implications for tangible and augmented interaction, introducing a cognitive development perspective to earlier work. The spread of ubiquitous and pervasive computing means that novel interactive systems are rapidly being deployed throughout the home and school environments in which children spend their early childhood years. New technologies have always modified the behaviours of the physical world, and past evidence is that cognitive development in early childhood is easily able to accommodate these environmental changes. However, at the same time, the embedding of digital infrastructure into our physical environment is extending the range of information media in our culture. Media literacy is increasingly delivered, not only by image, text, sound and screen, but through the novel behaviours of augmented “smart” objects.

The cognitive development perspective introduced in this paper has drawn attention to ways in which children in early-years education develop core cognitive skills in their interpretation and appropriation of augmented tangible objects. This is complementary to other ways in which the HCI community has extended its understanding of ubiquitous and pervasive interaction through turns to the social, and to embodiment, when theorising digitally augmented environments [6]. Furthermore, cognitive development is a complementary perspective to the more established computational analyses of distributed cognition and representation, by which we understand adult literacy and usability of tangible interaction [8].

During the study, children jointly produced diverse imaginary interpretations of their physical environment, encouraged by

the open-ended representations of FingAR, apparently exercising the cognitive skills of theory of mind and divergent thinking. These skills lead to social and creative competence in adulthood. Deficits in these skills also have key roles in some developmental disorders, leading to inability to engage in social pretend play during childhood, and subsequent disadvantages in adult life. Although we have previously explored the use of augmented mirrors as a therapeutic aid for children with those disorders [2], in this study we have explored the ways in which new styles of literacy and competence are likely to be developed by typical children within future ubiquitous and pervasive computing environments. Our theoretically-grounded exploration of children’s experience, building on prior work in developmental neuroscience and early-years education, can help to inform future understanding of the changing patterns of cognition within a culture of physically pervasive digital media [14].

ACKNOWLEDGEMENTS

We would like to thank Cambridge Overseas Trust for sponsoring this research project. We would also like to thank all psychology experts, school teachers and participants families for taking part in the study and providing insightful feedback.

REFERENCES

1. R. T. Azuma et al. A survey of augmented reality. *Presence*, 6(4):355–385, 1997.
2. Z. Bai, A. F. Blackwell, and G. Coulouris. Through the looking glass: Pretend play for children with autism. In *Mixed and Augmented Reality (ISMAR), 2013 IEEE International Symposium on*, pages 49–58. IEEE, 2013.
3. C. Barnes, D. E. Jacobs, J. Sanders, D. B. Goldman, S. Rusinkiewicz, A. Finkelstein, and M. Agrawala. Video puppetry: a performative interface for cutout animation. In *ACM Transactions on Graphics (TOG)*, volume 27, page 124. ACM, 2008.
4. A. L. Brown and M. J. Kane. Preschool children can learn to transfer: Learning to learn and learning from example. *Cognitive Psychology*, 20(4):493–523, 1988.
5. K. Coolahan, J. Fantuzzo, J. Mendez, and P. McDermott. Preschool peer interactions and readiness to learn: Relationships between classroom peer play and learning behaviors and conduct. *Journal of Educational Psychology*, 92(3):458, 2000.
6. P. Dourish. *Where the action is: the foundations of embodied interaction*. MIT press, 2004.
7. J. Dunn, J. Brown, C. Slomkowski, C. Tesla, and L. Youngblade. Young children’s understanding of other people’s feelings and beliefs: Individual differences and their antecedents. *Child development*, 62(6):1352–1366, 1991.
8. D. Edge, A. Blackwell, and L. Dubuc. The physical world as an abstract interface. *Contemporary Ergonomics*, pages 224–228, 2006.
9. P. Ekman. An argument for basic emotions. *Cognition & Emotion*, 6(3-4):169–200, 1992.

10. W. Farr, N. Yuill, and S. Hinske. An augmented toy and social interaction in children with autism. *International Journal of Arts and Technology*, 5(2):104–125, 2012.
11. G. G. Fein. Pretend play in childhood: An integrative review. *Child development*, pages 1095–1118, 1981.
12. C. Garvey and R. Berndt. The organization of pretend play. 1975.
13. Y. Guiard. Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *Journal of motor behavior*, 19(4):486–517, 1987.
14. N. K. Hayles. *How we think: Digital media and contemporary technogenesis*. University of Chicago Press, 2012.
15. S. E. Hunter, P. Maes, A. Tang, K. M. Inkpen, and S. M. Hesse. Waazam!: supporting creative play at a distance in customized video environments. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 1197–1206. ACM, 2014.
16. A. M. Leslie. Pretense and representation: The origins of "theory of mind." *Psychological review*, 94(4):412, 1987.
17. A. Lillard. Pretend play as twin earth: A social-cognitive analysis. *Developmental Review*, 21(4):495–531, 2001.
18. F. Lu, F. Tian, Y. Jiang, X. Cao, W. Luo, G. Li, X. Zhang, G. Dai, and H. Wang. Shadowstory: creative and collaborative digital storytelling inspired by cultural heritage. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1919–1928. ACM, 2011.
19. R. MacLaren and D. Olson. Trick or treat: Children's understanding of surprise. *Cognitive development*, 8(1):27–46, 1993.
20. P. Maes, T. Darrell, B. Blumberg, and A. Pentland. The alive system: Wireless, full-body interaction with autonomous agents. *Multimedia Systems*, 5(2):105–112, 1997.
21. P. Marshall, Y. Rogers, and M. Scaife. Puppet: a virtual environment for children to act and direct interactive narratives. In *2nd international workshop on narrative and interactive learning environments*, pages 8–15, 2002.
22. M. Nitsche and S. Nayak. Cell phone puppets: turning mobile phones into performing objects. In *Entertainment Computing-ICEC 2012*, pages 363–372. Springer, 2012.
23. J. Piaget. *Play, dreams and imitation in childhood*, volume 25. Routledge, 2013.
24. D. Premack and G. Woodruff. Does the chimpanzee have a theory of mind? *Behavioral and brain sciences*, 1(04):515–526, 1978.
25. I. Radu and B. MacIntyre. Using children's developmental psychology to guide augmented-reality design and usability. In *Mixed and Augmented Reality (ISMAR), 2012 IEEE International Symposium on*, pages 227–236. IEEE, 2012.
26. A. Rosenfeld and J. L. Pfaltz. Sequential operations in digital picture processing. *Journal of the ACM (JACM)*, 13(4):471–494, 1966.
27. T. Ruffman and T. R. Keenan. The belief-based emotion of surprise: The case for a lag in understanding relative to false belief. *Developmental Psychology*, 32(1):40, 1996.
28. M. A. Runco. *Divergent thinking*. Ablex Publishing, 1991.
29. M. A. Runco and R. S. Albert. The reliability and validity of ideational originality in the divergent thinking of academically gifted and nongifted children. *Educational and Psychological Measurement*, 45(3):483–501, 1985.
30. S. W. Russ. Play, creativity, and adaptive functioning: Implications for play interventions. *Journal of Clinical Child Psychology*, 27(4):469–480, 1998.
31. K. Ryokai and J. Cassell. Computer support for children's collaborative fantasy play and storytelling. In *Proceedings of the 1999 conference on Computer support for collaborative learning*, page 63. International Society of the Learning Sciences, 1999.
32. K. Ryokai, H. Raffle, and R. Kowalski. Storyfaces: pretend-play with ebooks to support social-emotional storytelling. In *Proceedings of the 11th International Conference on Interaction Design and Children*, pages 125–133. ACM, 2012.
33. R. K. Sawyer. Levels of analysis in pretend play discourse: Metacommunication in conversational routines. *PLAY AND CULTURE STUDIES*, 5:137–158, 2003.
34. B. Shneiderman. Direct manipulation: a step beyond programming languages. *Sparks of Innovation in Human-Computer Interaction*, 17, 1993.
35. S. Smilansky. *The effects of sociodramatic play on disadvantaged preschool children*. Wiley New York, 1968.
36. H. Tager-Flusberg and K. Sullivan. Attributing mental states to story characters: A comparison of narratives produced by autistic and mentally retarded individuals. *Applied Psycholinguistics*, 16(03):241–256, 1995.
37. F. Zhou, H. B.-L. Duh, and M. Billinghurst. Trends in augmented reality tracking, interaction and display: A review of ten years of ismar. In *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 193–202. IEEE Computer Society, 2008.
38. Z. Zhou, A. D. Cheok, J. Tedjokusumo, and G. S. Omer. wizquestm—a novel tangible interface for interactive storytelling in mixed reality. *Int J Virtual Real*, 7(4):9–15, 2008.