Research in Developmental Disabilities 36 (2015) 396–403

Contents lists available at ScienceDirect

Research in Developmental Disabilities

Augmented reality-based self-facial modeling to promote the emotional expression and social skills of adolescents with autism spectrum disorders

Chien-Hsu Chen [a,*], I-Jui Lee [a], Ling-Yi Lin [b]

[a] Ergonomics and Interaction Design Lab, Department of Industrial Design, National Cheng Kung University, No. 1 University Road, Tainan, Taiwan
[b] Department of Occupational Therapy, National Cheng Kung University, No. 1 University Road, Tainan, Taiwan

ARTICLE INFO

Article history:
Received 7 August 2014
Received in revised form 2 October 2014
Accepted 10 October 2014
Available online 8 November 2014

Keywords:
Augmented reality (AR)
Emotions
Self-facial modeling
Three-dimensional (3-D) facial expressions
3-D facial animation

ABSTRACT

Autism spectrum disorders (ASD) are characterized by a reduced ability to understand the emotions of other people; this ability involves recognizing facial expressions. This study assessed the possibility of enabling three adolescents with ASD to become aware of facial expressions observed in situations in a school setting simulated using augmented reality (AR) technology. The AR system provided three-dimensional (3-D) animations of six basic facial expressions overlaid on participant faces to facilitate practicing emotional judgments and social skills. Based on the multiple baseline design across subjects, the data indicated that AR intervention can improve the appropriate recognition and response to facial emotional expressions seen in the situational task.

1. Introduction

Autism spectrum disorders (ASD) are characterized by atypical patterns of behavior and impaired social communication (American Psychiatric Association, 2000; Krasny, Williams, Provencal, & Ozonoff, 2003). The challenge of social interaction for people with ASD is appropriately recognizing and understanding facial expressions that indicate emotions (Dawson, Webb, & McPartland, 2005; Ryan & Charragain, 2010; Williams, Gray, & Tonge, 2012). People with ASD have difficulty understanding the expressions and emotional states of other people and determining their intentions and thought processes, which results in an impaired ability to respond with appropriate expressions and to interact appropriately with their peers (Krasny et al., 2003). In addition, facial expression processing is atypical in people with ASD (Annaz, Karmiloff-Smith, Johnson, & Thomas, 2009); although some people with high-functioning autism (HFA) are relatively adept at social communication involving complex facial emotions, they have difficulty with nonverbal communication (Elder, Caterino, Chao, Shacknai, & De Simone, 2006). Emotion recognition is among the skills most crucial to social interaction and developing empathy (Baron-Cohen, 2002). Relevant studies have described empathy as a lens through which people comprehend emotional expressions and respond appropriately (Sucksmith, Allison, Baron-Cohen, Chakrabarti, & Hoekstra, 2013). However, people with ASD have deficits that include being unable to view events from the perspective of other people and to respond with appropriate expressions (Baron-Cohen & Belmonte, 2005; Baron-Cohen, Leslie, & Frith, 1985).

Research on emotional impairment in ASD has focused primarily on examining emotional recognition and understanding and on teaching facial expressions by labeling them on formatted photographs (Ashwin, Wheelwright, & Baron-Cohen, 2005;

* Corresponding author. Tel.: +886 6 2757575x54324.
E-mail addresses: chenhsu@mail.ncku.edu.tw (C.-H. Chen), P38981047@mail.ncku.edu.tw (I.-J. Lee), lingyi@mail.ncku.edu.tw (L.-Y. Lin).

http://dx.doi.org/10.1016/j.ridd.2014.10.015
0891-4222/

Begeer, Rieffe, Terwogt, & Stockmann, 2006; Ben Shalom et al., 2006; Castelli, 2005; Wang, Dapretto, Hariri, Sigman, & Bookheimer, 2004). For example, various facial expressions in photos and videos have been used to develop the communication skills of people with ASD, which enables them to focus on the specific visual representations and facial cues according to which the facial emotions of others can be determined (Blum-Dimaya, Reeve, Reeve, & Hoch, 2010). Current intervention systems used for people with ASD involve applying a third-person perspective to recognize and manipulate feelings based on the facial synthesis of 3-D characters; they support reusability of facial components, and have an avatar-user interaction model with real time responses (Kientz, Goodwin, Hayes, & Abowd, 2013), for example, online games that depict an imaginary world from a third-person perspective to represent the actions and statuses of other people. However, although the expressions of an avatar or cartoon character (Tseng & Do, 2010) facilitate learning using emotions, methods in which information is not presented from the perspective of the participants do not enable people with ASD to see the expressions on their own faces and thereby connect the expression with thoughts (Young & Posselt, 2012).

In addition, video self-modeling (VSM) has been used for social skills training and involves participants watching a video of a person modeling a desired behavior and then imitating the behavior of the person in the video (Axe & Evans, 2012). However, using VSM as an intervention strategy for individuals with ASD does not enable immediate feedback on the facial states of the participants during a scenario. These systems simply record the events occurring during the scenario and the physical behaviors imitated by the participants; therefore, participants have difficulty obtaining self-facial expression instruction. People with ASD experience difficulty in accessing self-facial expression treatment because training scenarios and real-time mood simulations in which people can pretend to feel various emotions are unavailable. Thus, emerging technologies such as augmented reality (AR) can be applied to teach learners to explore material from various perspectives (Asai, Kobayashi, & Kondo, 2005). Because these technologies have the potential to stimulate the senses of the user, they may be particularly useful for teaching subject matter that learners have difficulty experiencing in the real world (Chien, Chen, & Jeng, 2010; Shelton & Hedley, 2002) and for facilitating social interaction. In addition, unlike traditional learning content that provides only static texts and facial images to describe an emotional expression, the AR instructional model can be used to present the core learning content directly to participants with ASD and assist them in exploring self-facial expression. Therefore, we created an AR application that can be used to increase emotional expression recognition and social skills.

2. Methods

2.1. Participants

The three adolescent participants with ASD (Zhu, Lin, and Lai) were recruited through the Autism Association in Taiwan. The inclusion criteria specified that participants must have: (1) been clinically diagnosed with ASD, (2) no other specific disabilities, (3) a full-scale Intelligence Quotient (IQ) of more than 85, and (4) been clinically evaluated according to DSM-IV-TR criteria. All participants were fluent in Mandarin Chinese or Taiwanese and had no delays in cognitive development. The participants' sensory abilities were within the normal range. However, they all had poor social and communication skills and rarely understood how to respond with appropriate facial expressions to other people's emotions. The participants ranged in age from 10 to 13 years ($n = 3$; mean age: 12 years, 2 months). The mean (SD) full-scale IQ, verbal IQ, and performance IQ scores were 101 (9.07), 100 (8.73), and 101 (1.73), respectively. The participants' intelligence, sensory abilities, and social and communication skills were based on multiple information sources, such as parental interviews, teachers' reports, verbal IQ scores (Wechsler Intelligence Scale for Children), and levels of functional language and social adaptation levels (based on clinical observations or behavior and adaptation scales). The male:female ratio was 2:1. All participants had a medical disability identification card issued by medical institutions in Taiwan and had received counseling in special education schools and institutes in Taiwan. The participants' vision and hearing were normal. In addition, the National Cheng Kung University Hospital internal review board gave ethical clearance for the study (B-BR-103-028-T). Parental consent forms were obtained before the participants were enrolled in the study. All participants signed a youth consent form.

2.2. Developing the augmented-reality-based self-facial modeling learning system

The 3-D facial models of virtual characters were designed to fit the heads of all the participants, and six facial expressions that communicated basic emotions (happiness, sadness, fear, disgust, surprise, and anger) were developed according to the Facial Action Coding System (FACS) for this study (Hamm, Kohler, Gur, & Verma, 2011). To create the 3-D head model, we used frontal and side-view pictures of the faces of each participant to generate facial skin on the models using Facial Studio 3.2 (Di-O-Matic, 2014), which provides the user with more than 500 controls over the head-creation process, and animated the models using 3ds Max® 2012 (Autodesk, San Rafael, CA). Three 3-D head avatars (2 males and 1 female) were created and used in this study. After we built the models, we used the Unity game engine, which enables customized modeling, rigging, and animation and can be integrated with 3-D modeling and animation software pipelines, such as 3ds Max®, if the pipelines can support exporting in standard formats, e.g., FBX and OBJ sequences. Finally, the AR system consolidation was built using Qualcomm® AR (QCAR) in Unity with a Vuforia™ platform (Qualcomm, Inc., San Diego, CA), which enables rapid and accurate natural feature tracking of textured planar objects. Developing an AR environment with this software was easy