J Autism Dev Disord (2014) 44:2486–2498
DOI 10.1007/s10803-014-2118-6

ORIGINAL PAPER

Training Facial Expression Production in Children on the Autism Spectrum

Iris Gordon · Matthew D. Pierce · Marian S. Bartlett · James W. Tanaka

Published online: 29 April 2014
© Springer Science+Business Media New York 2014

Abstract Children with autism spectrum disorder (ASD) show deficits in their ability to produce facial expressions. In this study, a group of children with ASD and IQ-matched, typically developing (TD) children were trained to produce "happy" and "angry" expressions with the FaceMaze computer game. FaceMaze uses an automated computer recognition system that analyzes the child's facial expression in real time. Before and after playing the Angry and Happy versions of FaceMaze, children posed "happy" and "angry" expressions. Naïve raters judged the post-FaceMaze "happy" and "angry" expressions of the ASD group as higher in quality than their pre-FaceMaze productions. Moreover, the post-game expressions of the ASD group were rated as equal in quality as the expressions of the TD group.

Keywords Autism · Autism spectrum disorder · Facial expression · Expression production · Intervention · Social communication

Introduction

Facial expressions provide a window into an individual's private world by revealing information about a person's momentary affective state. Whereas facial expressions help initiate, modify and regulate patterns of social interaction (Barbu et al. 2001; Boyatzis et al. 1993), they are also subject to scrutiny in social situations (Ekman 1993; Izard and Malatesta 1987; Fridlund 1994). As a consequence, producing facial expressions that violate or are inconsistent with social expectations, or are ambiguous or difficult to interpret may hinder effective inter-personal communication. For example, if a friend receives a job promotion in our place, we might feign an expression of joy and elation to hide our true feelings of jealousy and disappointment that would offend our companion.

During social interactions, a person's internal emotion and the display of the outward facial expression are not always congruent. An emotion can be experienced internally without its externalization as a facial expression or body gesture (Campos 1985; Camras et al. 1998; Hiatt et al. 1979). Conversely, an externalized emotional display, such as a facial expression, can be expressed in the presence of an incongruent emotion, such as in cases of deception (Ekman and Friesen 1975; Ekman et al. 1991) or sympathy (Miller and Eisenberg 1988). This dissociation implies that facial expressions are not only the physiological consequences of an internal emotional state (i.e. spontaneous productions), but can also be a consciously controlled social display that are monitored and manipulated in order to meet external (social) demands (i.e. voluntary displays). Furthermore, unlike spontaneous facial expressions that are produced automatically, voluntary facial expressions are under a person's conscious control and can be initiated and regulated according to one's goals and intentions. In order to be produced efficiently, voluntary expressions rely on an individual's "expression concept", that is, the individual's internal representation of that expression.

I. Gordon (✉) · M. D. Pierce · J. W. Tanaka
Department of Psychology, University of Victoria,
P.O. Box 3050, Victoria, BC V8W 3P5, Canada
e-mail: igordon@uvic.ca

J. W. Tanaka
e-mail: jtanaka@uvic.ca

M. S. Bartlett
University of California at San Diego, San Diego, CA, USA

Facial Expressions in Autism

Autism spectrum disorder (ASD) is a pervasive developmental disorder typified by deficits in social interaction, communication, and restricted or repetitive behaviours (American Psychiatric Association 2000). One deficit in social communication is flat (lack of) or disorganized (ambiguous) facial expression production (Lord et al. 2000). Although Kanner (1943, 1968) first described the lack of social and emotional responsiveness of children with ASD, the social deficit in the perception and production of facial expressions was not identified. Langdell (1981) tested voluntary facial expression production of children with ASD, and children with non-specific developmental disorders (i.e. pervasive developmental disorder—not otherwise specified, or PDDNOS). Children were asked to produce happy and sad expressions while their productions were photographed and rated by naïve raters. Findings showed that ASD children's expressions were rated lower in quality when compared to the expressions of the PDDNOS children. In attempt to control for potential linguistic confounds, a second experiment relying strictly on non-verbal cuing was conducted, in which children were told to mimic a model's happy and sad facial expressions with and without visual feedback (mirror), while pictures of the children's faces were taken (Langdell 1981). Interestingly, the quality of the ASD productions was on par with the productions of the PDDNOS children when visual feedback through the mirror was provided. However, when the mirror was not available, the ASD productions were rated as lower in quality than the PDDNOS group. Langdell (1981) concluded that the expression production deficits seen in ASD was not due to a motor deficit, but rather from an inability to perceive and integrate the different components of facial expressions.

Several subsequent studies of emotional expressivity in ASD have extended findings to include other facial expressions while controlling for IQ. Macdonald et al. (1989) assessed high-functioning ASD adult's ability in recognizing and producing facial expressions in comparison to age- and IQ-matched TD control participants. TD and ASD participants were photographed while producing the facial expressions of happy, angry, fear, sad, and neutral in response to short vignettes and emotion-labels. Naïve raters who were asked to rate and label each photograph rated the ASD productions as lower in quality, and mislabeled the ASD expressions more frequently than the TD expressions. A higher proportion of mislabeling errors were found for the negative emotions (angry, fear, sad) whereas no group differences were found for the positive expressions of happy (Macdonald et al. 1989).

Loveland et al. (1994) quantified the extent to which facial affect was disordered in ASD by making the distinction between mimicked and posed expressions. Participants with ASD and Down's syndrome were rigorously matched on several IQ measures in order to remove any potential confound of intelligence. The mimicked expressions of happy, angry, sad, surprise and neutral facial were modeled by researchers, whereas posed expressions were prompted by the emotion label. Video-recordings of participants' faces were obtained and edited before being presented to judges who first labeled the expression, and then rated the expression for its overall quality. Consistent with the previous studies with photographs, ASD participants' posed expressions were rated as lower in quality than their mimicked expressions, and were qualitatively more bizarre and mechanical in production. When comparing across the two groups, ASD participant's facial expressions were rated as lower in quality than the DS group (Loveland et al. 1994). Findings from this study not only replicated those of previous research by demonstrating the flat and disorganized affect associated with ASD, yet also illuminated the conditions under which this deficit was observed. Whereas mimicry of facial expressions in ASD was relatively intact when compared to developmentally disordered peers, the expression quality of facial expressions posed without a visual example demonstrated marked deficits.

Interestingly, children with ASD show marked deficits in spontaneous mimicry (McIntosh et al. 2006; Rogers et al. 2003; Williams et al. 2001), however research has shown that these individuals retain their ability to mimic others when explicitly prompted (McIntosh et al. 2006; Rogers et al. 2003). Thus, by comparing the quality of facial displays that are mimicked (i.e. relying on an external model) to those that are posed (i.e. relying on the expression concept) it is possible to determine the cognitive source of the facial expression deficit in ASD. Results from previous experiments show that ASD individuals perform on par with typically developing children when they are required to mimic a facial expression in the presences of an external model. However, individuals with ASD show significant deficits in the quality of their expressions when asked to pose an expression in response to a verbal label. These results indicate that the characteristic flat or disorganized affect that typifies the autism does not result from an inability to activate or manipulate facial muscles, or from an inability to mirror the motor movements of a physical model. Rather, for children with ASD, there is a disconnection between the mental representation of an emotion and its production through a facial expression.

The goal of the current experiment is to test whether the production of posed expressions can be strengthened through practice and training. Children with ASD, and age and IQ-matched TD children played the FaceMaze game. FaceMaze is an interactive, PacMan-like video game in which participants navigate through a maze of obstacles while collecting

Table 1 Average K-Bit-2 scores for ASD and TD participants

Group	Verbal IQ	Non-verbal IQ	Composite IQ
ASD	108.12 (5.48)	107.00 (4.50)	108.94 (5.24)
TD	112.06 (2.43)	112.29 (3.09)	114.59 (2.40)
t test	$p = 0.52$	$p = 0.40$	$p = 0.38$

Parentheses denote standard errors

"tokens". Obstacles in the maze are overcome by producing facial expressions as measured by the computer recognition emotion toolbox (CERT). CERT analyzes the child's facial expressions via the webcam and provides real-time feedback to the user with respect to the quality of their expression productions. Children were videotaped while posing "happy," "angry" and "surprise" expressions before and after playing the Happy and Angry versions of the "FaceMaze" game intervention. Naïve participants were then asked to rate the videos for expression quality. We hypothesized that if the FaceMaze strengthens the link between the conceptual and motor representation of "happy" and "angry" emotions, exclusively, the post-game facial expressions of children with ASD should be rated as higher in quality than their pre-game expressions.

Method Part 1: Stimulus Generation

Participants

Thirty children with ASD, aged 6–18 years ($M = 10.89$, $SD = 3.39$), were recruited from the Centre for Autism Research Training and Education (CARTE) database. All children were diagnosed with Autism Spectrum Disorder through the British Columbia Autism Assessment Network (BCAAN) using the Autism Diagnostic Observation Schedule (ADOS) and Autism Diagnostic Interview (ADI). Two participants did not complete the task, and another 11 participants produced unusable data (see stimulus generation), resulting in 17 participants aged 6–18 years ($M = 10.76$, $SD = 3.59$). Twenty-three typically developing (TD) aged 8–16 years ($M = 10.39$, $SD = 2.64$) were recruited from the Centre for Autism Research Training and Education (CARTE) database. Of these, 17 age- and IQ-matched controls ($M = 10.94$, $SD = 2.79$) were obtained (see Table 1). Participants were compensated with a $10 gift-card to Chapters, and a small toy for their time.

Materials

The Computer Expression Recognition Toolbox (CERT)

To implement our training program, we employed the Computer Emotion Recognition Toolbox (CERT) developed by Bartlett et al. (2005, 2006; Littlewort et al. 2011). The Computer Expression Recognition Toolbox (CERT) is a fully automated computer vision system that analyzes facial expressions in real-time, using video input (Bartlett et al. 2005, 2006; Donato et al. 1999; Littlewort et al. 2011). CERT automatically detects facial actions from the Facial Action Coding System (FACS). The program was trained to detect each facial action based on over 8000 FACS-coded images of voluntary and spontaneous expressions. The CERT program automatically detects frontal faces in the video stream and codes each frame with respect to the 20 major action units (AU) most closely associated with the seven basic emotions. (For information on the training of the CERT program, see Littlewort et al. (2011).) Detection accuracy for individual facial actions has been shown to be 90 % for voluntary expressions, and 80 % for spontaneous expressions that occur within the context of natural head movements and speech, as measured by a 2-alternative forced paradigm that approximates area under the ROC curve (Littlewort et al. 2011). In addition, estimates of expression intensity generated by CERT correlate with FACS' expert intensity codes (Bartlett et al. 2006). This system has been successfully employed in a range of studies of spontaneous expressions. (For a review, see Bartlett and Whitehill 2011.)

Using CERT, we designed the "FaceMaze" game in which a player navigates a pac-man-like figure through a series of corridors, and removes face obstacles by producing the appropriate happy or angry expressions (Cockburn et al. 2008; see Fig. 1). CERT detects the target expression via webcam input, rates the quality of the expression and provides real time feedback to the player. When a user enacts the correct corresponding facial expression, the "expression meter" (a red bar that depicts the length of time an expression is held) begins to fill. While CERT detects the correct facial expression, the expression meter continues to fill with the red bar until finally the obstacle is removed from the maze path. If CERT does not detect the correct expression, the meter will terminate and the obstacle remains. Only when CERT detects the correct expression will the expression meter resume its movement again. The expression meter serves as feedback for the player, informing the player when their facial expression is matching or not, and the disappearance of the obstacles serve as a reward for correct facial expression production. Due to CERT's accuracy in dynamic facial detection, the expression meter will not fill if the wrong facial expression is produced, thus encouraging the user to produce the expression prompted and not one that may be easier to produce for the player.

FaceMaze interprets "happy" as the upward inflection of the lip produced by zygomaticus major facial muscle. FaceMaze operationalizes "angry" as the tensing of the

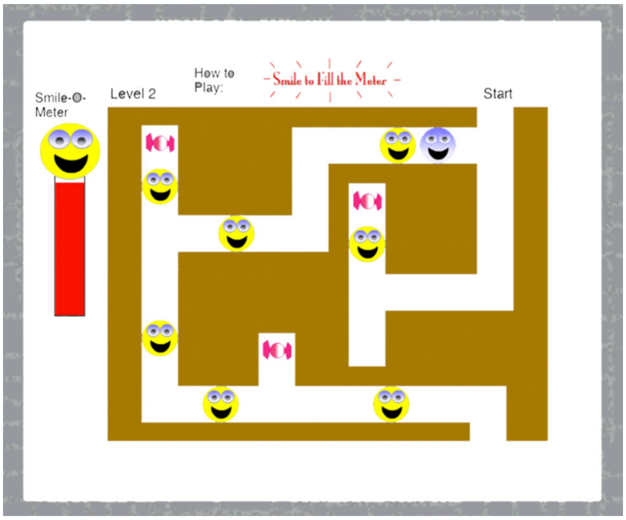

Fig. 1 The “Happy” level of FaceMaze. The player moves a *blue*, neutral pac-man like face throughout the maze, with the goal of collecting tokens (*pink* candy wrappers). In order to remove obstacles in their path, players must mimic the facial expression displayed by the obstacle. In HappyMaze, obstacles are other happy faces (*yellow*). When the player mimics the expression correctly, the *blue face* displays the expression and the smile-o-meter (*left*) fills (Color figure online)

corrugator supercilii that resulted in the visible furrowing of the brow detected by CERT, resulted in activation of the “anger detector” needed to successfully overcome barriers within the AngryMaze. Each stage was comprised of 3 separate game-levels consisting of a unique maze layout, eight face obstacles, and five “tokens”. Players were thus required to produce a total of 24 facial productions in order to complete each stage successfully.

Procedure

Prior to playing the Facemaze game, participants were administered the Kaufman Brief intelligence Test (2nd edition) (Kbit-2). The Kbit-2 yields both a verbal (crystalized) and non-verbal (fluid) intelligence score.

Pre- and Post-game Expression Production

Before and after playing the FaceMaze game, participants were asked to pose a “happy” face, an “angry” face and a “surprise” face. Participants were seated in front of the computer and were asked to look at a fixation cross. Children were prompted to “show me your best Happy face, Angry face and Surprise face” while their expressions were videotaped. The experimenter paused for 3 s between each prompt in order to allow the child to produce and maintain their expression. Videos of the happy, angry and surprised expressions were recorded at three time points: (1) before playing the happy version of FaceMaze, (2) after

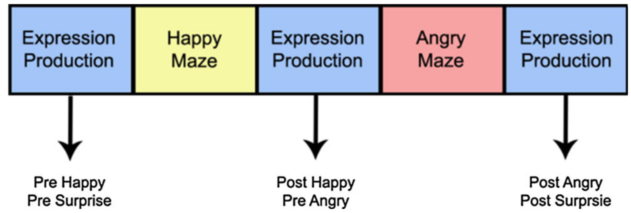

Fig. 2 Timeline of events used for obtaining videos of participant’s facial expressions. Videos were recorded during “expression production” phases as demarked

playing the happy version of FaceMaze (and before playing the angry version of FaceMaze and (3) after playing the angry version of FaceMaze (see Fig. 2). Videos were excluded from the final stimulus set if the participant’s face was not visible due to occlusion, or if the child moved out of screen. Furthermore, participants must have had all six videos in order to be included in the stimulus set. A total of 204 videos were included, 102 from the ASD group, and another 102 from the TD group.

To ensure the child’s success, participants played the “happy” version of FaceMaze first and then the “angry” version. In both conditions, participants were given a practice trial before playing the game, in which they were required to navigate toward an obstacle (an emotive icon depicting either a happy or angry facial expression), produce the corresponding facial expression, and acquire a token. After becoming acquainted with the rules of the game, participants were required to play for 4 min or 3 levels, whichever was completed first.

Part 2: Stimulus Rating

Participants

Forty-six naïve undergraduate participants from the University of Victoria participated in the video rating. Twenty-two participants (5 male), aged 18–22 ($M = 20.04$, $SD = 1.19$) rated the videos obtained from the ASD children, and 24 participants (6 male), aged 18–32 ($M = 22.63$, $SD = 3.21$) rated the TD children’s videos. All participants had normal or corrected-to-normal vision. Participants received course credit as compensation for their time.

Stimuli

Each video clip was presented on a computer screen situated 1 m away from participants, resulting in a retinal image of 16.51 × 10.16 cm on a white screen, creating a visual angle of 44.7° in the horizontal plane and 27.64 in the vertical plane.

A subset of 204 video clips obtained from the stimulus generation portion was used. 102 video clips were taken from those generated by the ASD group, and another 102 video clips were taken from the TD group. The ASD group's videos were further divided into categories based on the differential completion rates of the AngryMaze. As a result, the videos of eight participants (48 videos) from the 3-level condition, four participants (24 videos) from the 2-level condition, and five participants (30 videos) from the one-level condition were used, as well as the corresponding happy and surprise video clips.

Procedure

Participants were seated in front of the computer and full consent was obtained before the experiment start. Participants were told that they would be rating a series of video clips based on their subjective observations, and were explained the rating scales, while underscoring that each video clip had to be rated on all 6 scales, including any ratings of "0" (not at all). A practice video was presented and participants rated the video; if there were no other questions, participants proceeded to the experiment phase.

In the experiment phase, participants were required to rate a total of 102 videos, with each video comprising one trial. Video clips were divided into two blocks of 51 videos, with half the pre- and half the post-training videos randomly presented in each block. Presentation of the blocks was then counterbalanced across participants.

At the beginning of each trial, a screen reading "get ready" was presented for 2 s, followed by presentation of the video clip, succeeded by a screen reading "Please rate the video now. When you are finished, please press 'spacebar' to continue…". Thus, participants could view and rate the videos at their own pace. After the participants completed their task, they were debriefed and thanked for their time.

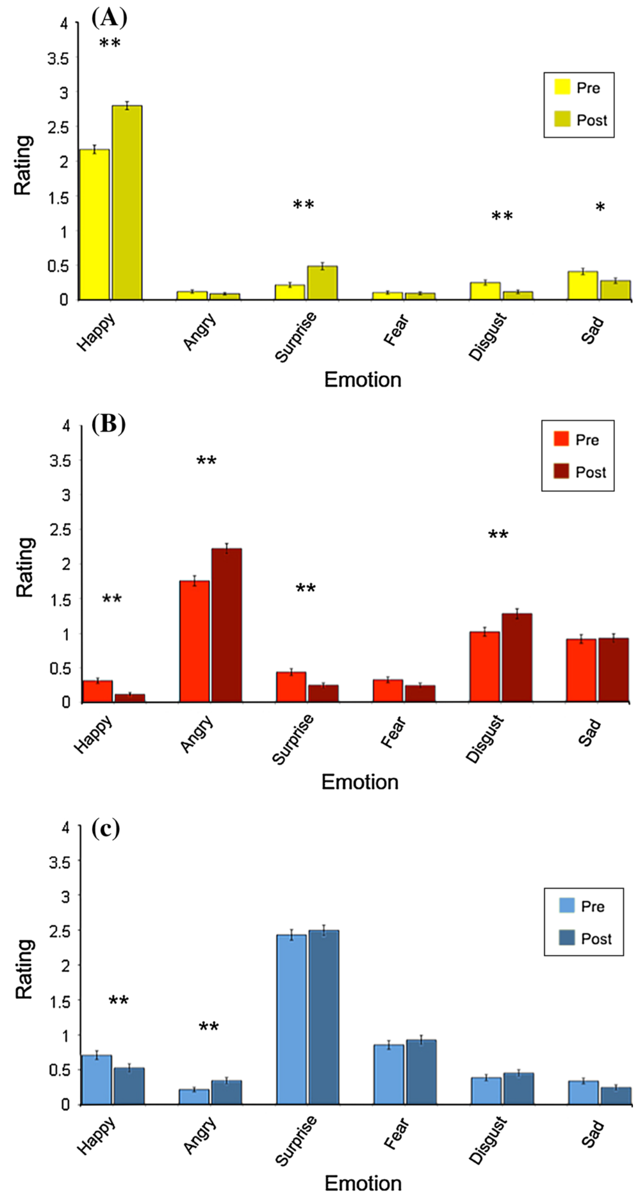

Fig. 3 *Bar-graph* of expression quality ratings for the **a** Happy expression, **b** Angry expression, and **c** Surprise expression, before and after training. *Error bars* denote Standard Error of the mean. *Asterisk* represents a significant difference at $p < 0.05$. *Asterisk* represents a significant difference at $p < 0.05$. *Double asterisk* represents significance at $p < 0.01$

Results

Ratings of the FaceMaze Videos of ASD Children

HappyMaze

In order to determine the effects of playing HappyMaze on facial expression production, expression quality ratings were subjected to a 2 (time: pre × post) × 6 (emotion: happy, angry, surprise, fear, disgust, sad) repeated-measures ANOVA. All tests used the Greenhouse–Geisser adjustments, and all post hoc comparisons were Bonferonni corrected. A significant main effect of Time was found, $F(1, 407) = 30.45, p < 0.001, \eta_p^2 = 0.07$, such that post-HappyMaze expression quality ratings ($M = 0.65$, $SE = 0.02$) were significantly higher than those of pre-HappyMaze expression quality ratings ($M = 0.55$, $SE = 0.01$). The main effect of Emotion was also significant, $F(2.72, 1{,}108.84) = 898.24$, $p < 0.001$, $\eta_p^2 = 0.69$, with expression quality ratings of happy ($M = 2.48$, $SE = 0.05$) significantly higher than those of angry ($M = 0.11$, $SE = 0.02$), surprise ($M = 0.35$, $SE = 0.03$), fear ($M = 0.10$, $SE = 0.02$), disgust ($M = 0.19$, $SE = 0.02$), and sad ($M = 0.35$, $SE = 0.03$). Furthermore, ratings of surprise and sad were also significantly higher than ratings of angry, fear, disgust, with ratings of disgust being significantly higher than ratings of angry and fear. In order of magnitude, expression quality ratings of happy were the largest, followed by surprise, fear and sad, and finally angry and disgust. Critically, a significant Time × Emotion interaction was observed, F (3.55, 1,445.67) = 36.74, $p < 0.001$, $\eta_p^2 = 0.08$.

Post-hoc comparisons revealed a significant increase in expression quality ratings of happy post-HappyMaze ($M = 2.80$, SE = 0 0.06) when compared to pre-HappyMaze productions ($M = 2.11$, SE = 0.06), t (407) = −5.39, $p < 0.001$. In addition, a significant increase in expression quality ratings of surprise post-HappyMaze ($M = 0.49$, SE = 0.05) as compared to pre-HappyMaze productions ($M = 0.22$, SE = 0.03) was also observed, t (407) = −4.27, $p < 0.001$. Furthermore, a significant decrease in expression quality ratings of disgust post-HappyMaze ($M = 0.12$, SE = 0.04), when compared to pre-HappyMaze productions ($M = 0.25$, SE = 0.03), was also significant, t (407) = 3.31, $p = 0.001$. Finally, a significant decrease in expression quality ratings of sad post-HappyMaze ($M = 0.28$, SE = 0.04), when compared to pre-HappyMaze productions ($M = 0.41$, SE = 0.04), was also observed, t (407) = 2.42, $p < 0.05$ (see Fig. 3).

AngryMaze

In order to determine the effects of playing AngryMaze on facial expression production, expression quality ratings were subjected to a 2 (time: pre × post) × 6 (emotion: happy, angry, surprise, fear, disgust, sad) repeated-measures ANOVA. All tests used the Greenhouse–Geisser adjustments, and all post hoc comparisons were Bonferonni corrected. A significant main effect of Time, F (1, 407) = 4.60, $p < 0.05$, $\eta_p^2 = 0.01$, was observed, such that post-AngryMaze expression quality ratings ($M = 0.84$, SE = 0.02) were significantly higher than those of pre-AngryMaze expression quality ratings ($M = 0.80$, SE = 0.02). Furthermore, a significant main effect of Emotion was also observed, F (3.54, 1,442.04) = 221.86, $p < 0.001$, $\eta_p^2 = 0.35$, with expression quality ratings of angry ($M = 1.99$, SE = 0.06) significantly higher than those of happy ($M = 0.22$, SE = 0.03), surprise ($M = 0.34$, SE = 0.03), fear ($M = 0.29$, SE = 0.03), disgust ($M = 1.15$, SE = 0.05) and sad ($M = 0.93$, SE = 0.05). Furthermore, expression quality ratings of disgust were significantly higher than those of happy, surprise, fear and a sad, with those of sad significantly higher than those of happy, surprise and fear. In order of magnitude, expression quality ratings of angry were the largest, followed by disgust, then sad, then surprise, and finally those of happy, and fear. Finally, a significant Time × Emotion interaction was found, F (3.72, 1,514.17) = 14.67, $p < 0.001$, $\eta_p^2 = 0.04$.

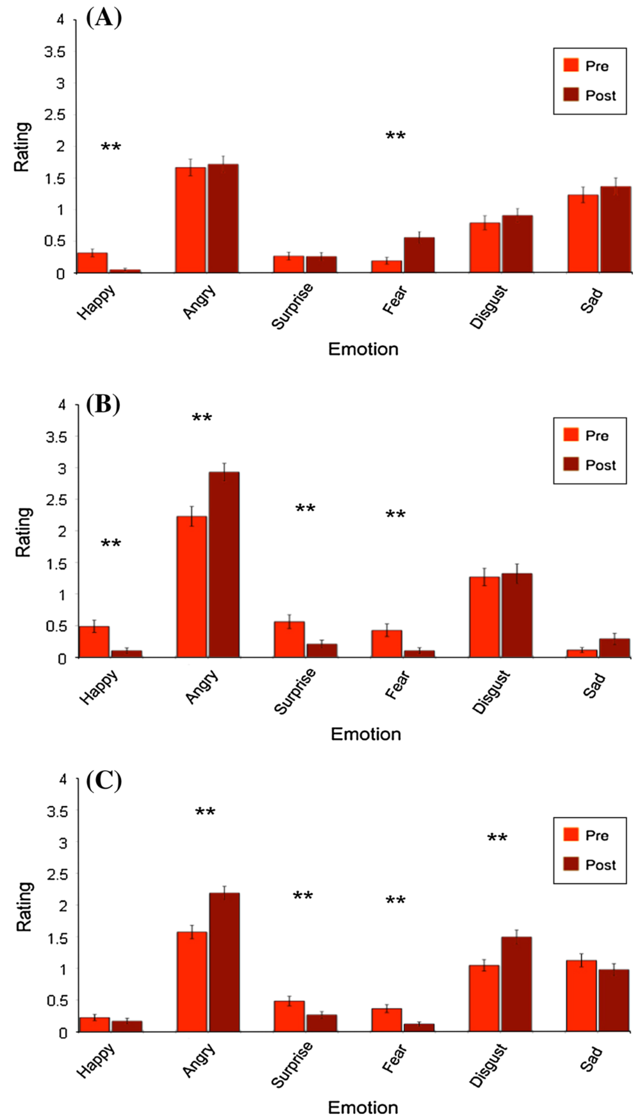

Fig. 4 *Bar-graph* of expression quality ratings for the **a** Level 1 Group, **b** Level 2 Group, and **c** Level 3 Group, before and after training. *Error bars* denote Standard Error of the mean. *Asterisk* represents a significant difference at $p < 0.05$. *Double asterisk* represents significance at $p < 0.01$

Post-hoc analysis revealed a significant increase in expression quality ratings of angry post-AngryMaze ($M = 2.22$, SE = 0.07) when compared to pre-AngryMaze ratings ($M = 1.75$, SE = 0.07), t (407) = −5.41, $p < 0.001$. Furthermore, a significant decrease in post-AngryMaze expression quality ratings of happy ($M = 0.12$, SE = 0.03), and surprise ($M = 0.25$, SE = 0.03) was observed, when compared to pre-AngryMaze ratings (happy: $M = 0.31$, SE = 0.04, surprise: $M = 0.44$, SE = 0.05), t (407) = 5.31, $p < 0.001$, and t (407) = 3.58, $p < 0.001$, respectively. Finally, a significant increase in post-AngryMaze expression quality ratings of disgust

($M = 1.28$, SE $= 0.07$) when compared to pre-AngryMaze expression quality ratings ($M = 1.02$, SE $= 0.06$) was also found, t (407) $= -3.16$, $p < 0.001$ (see Fig. 3).

In light of the different completion rates for the AngryMaze condition, data was subsequently divided based on the number of game-levels completed, resulting in three groups of participants who either completed one level of the AngryMaze, two levels of AngryMaze, or three levels of AngryMaze.

Subsequent post hoc comparisons were carried out for each of group level of completion, separately. For the Level 1 group, a significant increase in post-AngryMaze expression quality ratings of fear ($M = 0.56$, SE $= 0.09$), when compared to pre-AngryMaze expression quality ratings ($M = 0.19$, SE $= 0.05$), was found, t (119) $= -4.26$, $p < 0.001$. Furthermore, post-AngryMaze expression quality ratings of happy ($M = 0.05$, SE $= 0.02$) were significantly smaller than pre-AngryMaze expression quality ratings (happy: $M = 0.32$, SE $= 0.06$), t (119) $= 6.34$, $p < 0.001$ (see Fig. 4). Importantly, no differences were found between pre- and post-training productions for the target expression of angry, report non-significant t value here.

For the Level 2 group, a significant increase in post-AngryMaze expression quality ratings of angry ($M = 2.93$, SE $= 0.14$) when compared to pre-AngryMaze expression quality ratings ($M = 2.23$, SE $= 0.16$) was found, t (95) $= -4.15$, $p < 0.001$. Furthermore, significant decreases in post-AngryMaze expression quality ratings of happy ($M = 0.10$, SE $= 0.05$), t (95) $= 4.15$, $p < 0.001$, surprise ($M = 0.21$, SE $= 0.06$), t (95) $= 2.97$, $p < 0.001$, and fear ($M = 0.10$, SE $= 0.05$), t (95) $= 3.31$, $p < 0.001$, were found, when compared with pre-AngryMaze expression quality ratings (happy: $M = 0.49$, SE $= 0.10$, surprise: $M = 0.56$, SE $= 0.11$, fear: $M = 0.43$, SE $= 0.10$; see Fig. 4).

For the Level 3, a significant increase in post-AngryMaze expression quality ratings of angry ($M = 2.19$, SE $= 0.11$) was found when compared to pre-AngryMaze expression quality ratings ($M = 1.57$, SE $= 0.11$), t (191) $= -4.82$, $p < 0.001$. Additionally, an increase in post-AngryMaze expression quality ratings of disgust ($M = 1.49$, SE $= 0.11$) when compared to pre-AngryMaze expression quality ratings ($M = 1.05$, SE $= 0.09$) was found, t (191) $= -3.40$, $p = 0.001$. Lastly, significant decreases in post-AngryMaze expression quality ratings of surprise ($M = 0.27$, SE $= 0.05$), and fear ($M = 0.12$, SE $= 0.03$), were observed, t (87) $= -2.97$, $p < 0.005$, and t (87) $= -2.97$, $p < 0.005$, respectively, when compared to pre-AngryMaze expression quality ratings (surprise: $M = 0.48$, SE $= 0.08$, fear: $M = 0.36$, SE $= 0.06$; see Fig. 4).

Surprise

In order to determine the effects of playing HappyMaze and AngryMaze on facial expression production, expression quality ratings were subjected to a 2 (time: pre × post) × 6 (emotion: happy, angry, surprise, fear, disgust, sad) repeated-measures ANOVA. All tests used the Greenhouse–Geisser adjustments, and all post hoc comparisons were Bonferonni corrected. For the control Surprise expression, the main effect of Emotion was significant, F (3.28, 1,335.92) $= 295.36$, $p < 0.001$, $\eta_p^2 = 0.42$, with expression quality ratings of surprise ($M = 2.46$, SE $= 0.07$) significantly higher than those of happy ($M = 0.62$, SE $= 0.05$), angry ($M = 0.28$, SE $= 0.03$), fear ($M = 0.89$, SE $= 0.06$), disgust ($M = 0.42$, SE $= 0.03$), and sad ($M = 0.29$, SE $= 0.03$). Furthermore, ratings of fear were significantly higher than those of happy, angry, disgust and sad, whereas ratings of happy and disgust were significantly higher than those of anger and sad. In order of magnitude, expression quality ratings of surprise were the highest, followed by fear, happy, disgust, and finally angry and sad. No significant main effect of Time, F (1, 407) $= 0.20$, $p = 0.65$, $\eta_p^2 = 0.00$, was observed. A significant interaction effect of Time × Emotion, F (4.40, 1,790.12) $= 3.87$, $p < 0.005$, $\eta_p^2 = 0.01$, was found (see Fig. 3).

Post-hoc analysis revealed a significant decrease in expression quality ratings of happy post-gameplay ($M = 0.53$, SE $= 0.06$) when compared to pre-gameplay ratings ($M = 0.71$, SE $= 0.06$), t (407) $= 3.53$, $p < 0.001$. Furthermore, an increase in expression quality ratings of angry post-gameplay ($M = 0.71$, SE $= 0.06$) when compared to pre-gameplay ratings ($M = 0.22$, SE $= 0.03$) was also found, t (407) $= -2.67$, $p < 0.005$.

In summary, for the ASD group, naïve observers rated post-HappyMaze productions of Happy expressions higher in expression quality than pre-HappyMaze productions. Furthermore, a decrease in expression quality ratings of disgust and sad, and increase in expression quality ratings of surprise were also observed. Similarly, naïve observers rated post-AngryMaze productions of Angry higher in expression quality when compared to pre-AngryMaze productions. Furthermore, an increase in expression quality ratings of disgust, and decrease in expression quality ratings of happy and surprise was also found. With respect to different levels of completion, significant increases in expression quality ratings of angry were found for both the Level 2 and 3 groups. Significant decreases in expression quality ratings were also found, with expression quality ratings of happy decreasing for the Level 1 group, a decrease in expression quality ratings of happy, surprise, and fear for the Level 2 group, and a decrease in expression quality ratings of surprise and fear for the Level 3 group. Lastly, significant increases in expression quality ratings of fear for the Level 1 group, and disgust for the Level 3 group were also observed.

Ratings of the FaceMaze Videos for TD Children

HappyMaze

In order to determine the effects of playing HappyMaze on facial expression production in TD children, expression quality ratings were subjected to a 2 (time; pre × post) × 6 (emotion; happy, angry, surprise, fear, disgust, sad) repeated-measures ANOVA. All tests used the Greenhouse–Geisser adjustments, and all post hoc comparisons were Bonferonni corrected. A significant main effect of Time was found, F (1, 407) = 6.82, $p < 0.01$, $\eta_p^2 = 0.02$, with expression quality ratings post-HappyMaze ($M = 0.54$, SE = 0.01) significantly smaller than those of pre-HappyMaze ($M = 0.59$, SE = 0.01). Furthermore, a main effect of Rating was also found, F (2.20, 895.67) = 2,077.74, $p < 0.001$, $\eta_p^2 = 0.84$, with expression quality ratings of happy ($M = 2.8$, SE = 0.05) reliably higher than those of surprise ($M = 0.29$, SE = 0.03), anger ($M = 0.04$, SE = 0.01), fear ($M = 0.08$, SE = 0.01), disgust ($M = 0.07$, SE = 0.02), and sad ($M = 0.10$, SE = 0.02). Furthermore, ratings of surprise were also significantly higher than those of anger, fear, disgust and sad, and ratings of sad were also higher than those of angry. In order of magnitude, ratings of happy were the largest, followed by surprise, then sad, and the remaining angry, fear and disgust. Finally, a significant Time × Rating interaction, F (2.64, 1,075.80) = 4.91, $p < 0.005$, $\eta_p^2 = 0.01$, was also found.

Post-hoc analysis revealed a significant difference decrease in post-HappyMaze expression quality ratings of surprise ($M = 0.19$, SE = 0.03) when compared to pre-HappyMaze expression quality ratings ($M = 0.40$, SE = 0.04), t (407) = 4.29, $p < 0.001$ (see Fig. 5).

AngryMaze

In order to determine the effects of playing AngryMaze on facial expression production in TD children, expression quality ratings were subjected to a 2 (time; pre × post) × 6 (emotion; happy, angry, surprise, fear, disgust, sad) repeated-measures ANOVA. All tests used the Greenhouse–Geisser adjustments, and all post hoc comparisons were Bonferonni corrected. For the Angry expression, a significant main effect of Rating, F (3.33, 1,356.69) = 241.41, $p < 0.001$, $\eta_p^2 = 0.37$, was found,

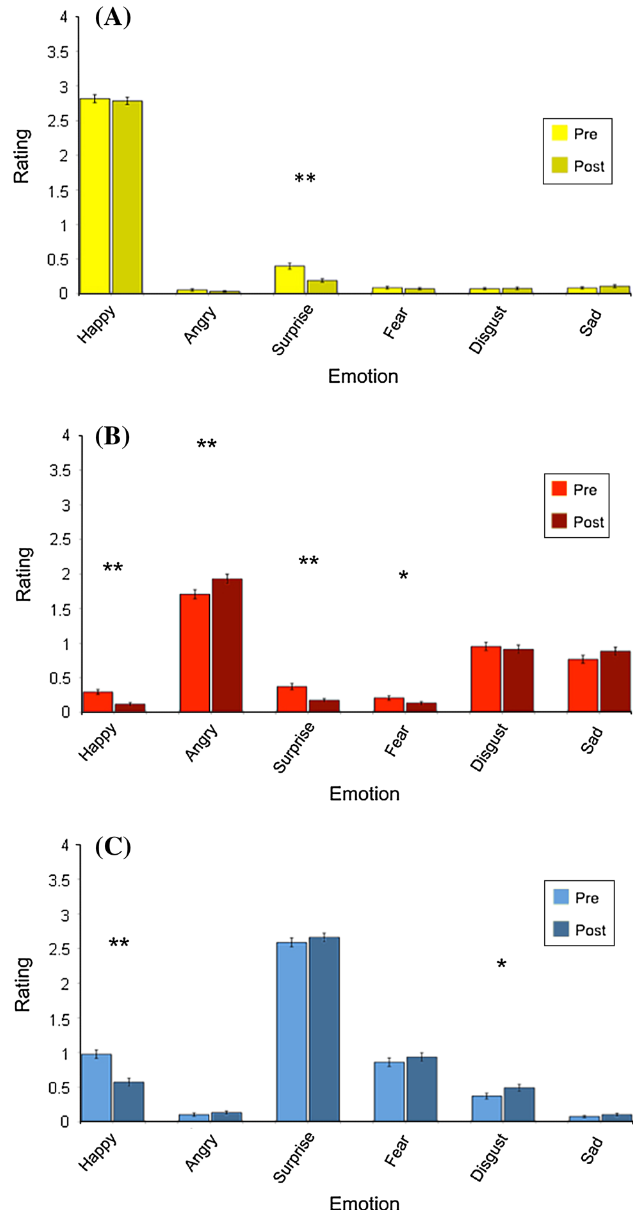

Fig. 5 *Bar-graph* of expression quality ratings for the **a** Happy expression, **b** Angry expression, and **c** Surprise expression, before and after training. *Error bars* denote Standard Error of the mean. *Asterisk* represents a significant difference at $p < 0.05$. *Double asterisk* represents significance at $p < 0.01$

with ratings of angry ($M = 1.82$, SE = 0.06) reliably higher than those of happy ($M = 0.21$, SE = 0.02), surprise ($M = 0.27$, SE = 0.03), fear ($M = 0.17$, SE = 0.02), disgust ($M = 0.94$, SE = 0.05) and sad ($M = 0.83$, SE = 0.05). Furthermore, ratings of disgust and sad were also significantly higher than those of happy, surprise and fear, with ratings of fear lower than those of surprise and happy. In order of magnitude, ratings of angry were the highest, followed by disgust and sad, then surprise and happy, and finally fear. No reliable main effect of Time

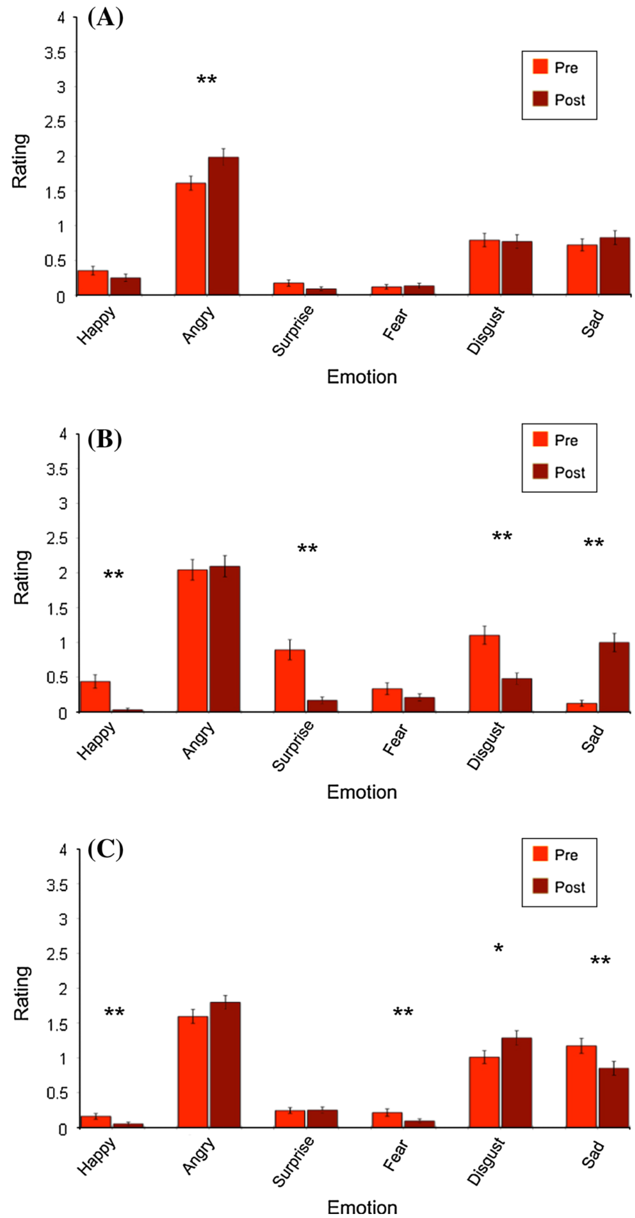

Fig. 6 *Bar-graph* of expression quality ratings for the **a** Level 1 group, **b** Level 2 group and **c** Level 3 group, before and after training. *Error bars* denote Standard Error of the mean. *Asterisk* represents a significant difference at $p < 0.05$. *Double asterisk* represents significance at $p < 0.01$

was observed, F (1, 407) = , $p = 1.56$, $\eta_p^2 = 0.21$, however a significant Time × Rating interaction was found, F (3.77, 1,532.56) = 7.86, $p < 0.00$, $\eta_p^2 = 0.02$.

Post-hoc analysis revealed a reliable increase in expression quality ratings of angry post-AngryMaze ($M = 1.93$, SE = 1.40) when compared to pre-AngryMaze expression quality ratings ($M = 1.71$, SE = 1.33), t (407) = −3.02, $p < 0.005$. Furthermore, a reliable decrease in post-AngryMaze expression quality ratings of happy ($M = 0.12$, SE = 0.02), surprise ($M = 0.17$, SE = 0.02) and fear ($M = 0.13$, SE = 0.02) was observed, t (407) = 4.38, $p < 0.001$, t (407) = 4.24, $p < 0.001$, and t (407) = 2.19, $p < 0.05$, respectively, when compared to pre-AngryMaze expression quality ratings (happy: $M = 0.29$, SE = 0.04, surprise: $M = 0.37$, SE = 0.04, fear: $M = 0.21$, SE = 0.03; see Fig. 5).

Similar to the ASD group, TD children completed the AngryMaze with varying degrees of success, resulting in three groups of participants who either completed 1 level of the AngryMaze, 2 levels of the AngryMaze, or 3 levels of AngryMaze. As a result, subsequent post hoc comparisons were carried out for each of group level of completion, separately.

For the Level 1 group, a reliable increase in expression quality ratings of angry post-AngryMaze ($M = 1.99$, SE = 0.12) when compared to pre-AngryMaze ratings ($M = 1.61$, SE = 0.10) was observed, t (143) = −3.33, $p < 0.005$ (see Fig. 6).

For the Level 2 group, a significant decrease in expression quality ratings of happy ($M = 0.03$, SE = 0.02), surprise ($M = 0.16$, SE = 0.05), and disgust ($M = 0.48$, SE = 0.08) was observed, t (95) = 4.06, $p < 0.001$, t (95) = 4.86, $p < 0.001$, and t (95) = 5.30, $p < 0.001$, respectively, when compared to pre-AngryMaze expressions quality ratings (happy: $M = 0.44$, SE = 0.10, surprise: $M = 0.90$, SE = 0.14, disgust: $M = 1.10$, SE = 0.13). Furthermore, a reliable increase in expression quality ratings of sad were found post-AngryMaze ($M = 1.00$, SE = 0.13), when compared to pre-AngryMaze expression quality ratings ($M = 0.13$, SE = 0.04), t (95) = −6.40, $p < 0.001$ (see Fig. 6).

For the Level 3 group, a reliable decrease in post-AngryMaze expression quality ratings of happy ($M = 0.05$, SE = 0.03), fear ($M = 0.10$, SE = 0.03), and sad ($M = 0.85$, SE = 1.00), was observed, when compared to pre-AngryMaze ratings (happy: $M = 0.16$, SE = 0.04, fear: $M = 0.21$, SE = 0.05, sad: $M = 1.17$, SE = 0.11), t (167) = 2.32, $p < 0.05$, t (167) = 2.26, $p < 0.05$, and t (167) = 3.27, $p < 0.005$, respectively. Furthermore, a significant increase in expression quality ratings of disgust ($M = 1.29$, SE = 0.10) post-AngryMaze, was found, t (167) = −2.35, $p < 0.05$, when compared to pre-AngryMaze ratings ($M = 1.01$, SE = 0.09; see Fig. 6).

Surprise

For the control Surprise expression, analysis revealed a significant main effect of Rating, F (3.18, 1,293.66) = 526.66, $p < 0.001$, $\eta_p^2 = 0.56$, such that expression quality ratings of surprise ($M = 2.62$, SE = 0.05) were significantly higher than those of happy ($M = 0.77$, SE = 0.05), angry ($M = 0.12$, SE = 0.02), fear ($M = 0.89$, SE = 0.05), disgust ($M = 0.43$, SE = 0.04), and sad ($M = 0.09$, SE = 0.01). Furthermore, ratings of happy and fear were significantly higher than those of angry, disgust and sad, with ratings of disgust reliably higher than those of angry and sad. In order of magnitude, expression quality ratings of surprise were the largest, followed by those of fear and happy, then disgust, and finally angry and sad. No significant main effect of Time was found, F (1, 407) = 0.55, $p < 0.46$, $\eta_p^2 = 0.00$, however a reliable interaction effect of Time × Rating was observed, F (3.54, 1,441.48) = 11.39, $p < 0.05$, $\eta_p^2 = 0.03$.

Subsequent post hoc analysis revealed a significant decrease in expression quality ratings of happy post-game-play ($M = 0.57$, SE = 0.06) when compared to pre-game-play expression quality ratings ($M = 0.98$, SE = 0.06). Furthermore, a reliable increase in expression quality ratings of disgust was found post-game-play ($M = 0.49$, SE = 0.05) when compared to pre-game-play expression quality ratings ($M = 0.37$, SE = 0.04; see Fig. 5).

In summary, for the TD group, no changes in happy expression quality ratings were found for Happy expressions post-HappyMaze when compared to pre-HappyMaze productions. Furthermore, a decrease in expression quality ratings of surprise was also observed. In contrast, naïve observers rated post-AngryMaze productions of Angry higher in expression quality when compared to pre-AngryMaze productions. Furthermore, decreases in expression quality ratings of happy, surprise and fear were also found. With respect to different levels of completion, significant increases in expression quality ratings of angry were found for the Level 1 group. Furthermore, significant decreases in expression quality ratings were also found, with expression quality ratings of happy, surprise, and disgust decreasing for the Level 2 group, and a decrease in expression quality ratings of happy, fear, and sad for the Level 3 group. Lastly, significant increases in expression quality ratings of sad for the Level 2 group, and disgust for the Level 3 group were also observed.

Comparing the Ratings of the FaceMaze Videos of ASD Children and TD Children

In order to determine the efficacy of FaceMaze in enhancing facial expression production, a series of t test were carried out between the ASD and TD group's target expression quality ratings. A reliable difference in pre-HappyMaze expression quality ratings of happy between the ASD ($M = 2.17$, SE = 0.06) and TD ($M = 2.82$, SD = 0.06) groups was found, t (407) = 7.74, $p < 0.001$, however no difference between ASD ($M = 2.80$, SD = 0.06) and TD ($M = 2.78$, SD = 0.05) post-Happymaze happy expression quality ratings was observed (see Fig. 7). Furthermore, whereas no reliable difference was found for pre-AngryMaze angry expression quality ratings between the ASD ($M = 1.75$, SD = 0.07) and TD ($M = 1.71$, SD = 0.07) groups, a reliable difference was found in post-AngryMaze angry expression quality ratings between the ASD ($M = 2.22$, SD = 0.07) and TD ($M = 1.92$, SD = 0.07) groups, t (407) = −2.91, $p < 0.005$ (see Fig. 7). Lastly, no reliable difference in pre- or post-FaceMaze expression quality ratings of surprise between the ASD (pre: $M = 2.43$, SD = 0.07, post: $M = 2.49$, SD = 0.08) and TD group (pre: $M = 2.58$, SD = 0.06, post: $M = 2.66$, SD = 0.06), were found.

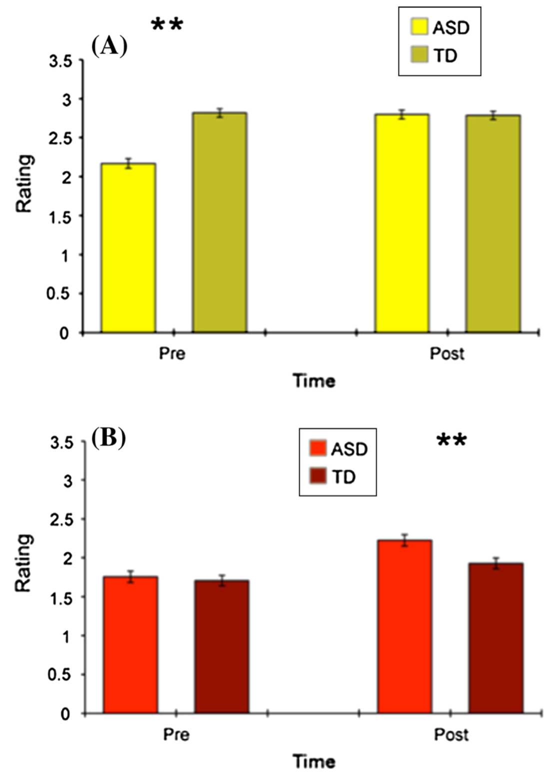

Fig. 7 *Bar-graph* of happy expression quality ratings for the **a** HappyMaze condition, and **b** AngryMaze condition, before and after training, for both ASD and TD groups. *Error bars* denote Standard Error of the mean. *Double asterisk* represents a significant difference at $p < 0.005$

Discussion

The goal of the current experiment was to determine the efficacy of the FaceMaze game in altering facial expression production in children with ASD, who are typified by disorders in social communication such as flat or disorganized affect. Findings revealed that FaceMaze was effective in increasing the perceptibility of target facial expressions in children with ASD by first enhancing facial expression fidelity, and second, by attenuating competing emotion displays. More importantly, the control Surprise expression showed no changes in target expression quality ratings, underscoring that the increases in the Happy and

Angry expressions' perceptibility resulted from directed training and not merely from participants activating facial muscles. Furthermore, the decrease in expression quality ratings of happy for both the Angry and Surprise expressions also underscores a conceptual differentiation between the positive Happy display, and the negative Angry display or neutral Surprise display, substantiating the efficacy of FaceMaze training in targeting the expression concept.

Modulation of competing expressions was also observed in the ASD group for both the HappyMaze and AngryMaze conditions. With respect to the Happy facial expression, the small yet significant increase in ASD participant's expression quality ratings of surprise may reflect the tone in which facial expressions were elicited, and not necessarily their quality. According to the circumplex model of affect, the emotion of Surprise has a neutral affect-valence, but is considered high on arousal, as opposed to Happy, which is high on positive valence and almost neutral on arousal (Russell and Barrett 1999). Since our study required participants to rate facial expressions on expression quality, it is likely that changes in expression intensity were captured by co-occurring arousal changes in the surprise expression rating.

TD participant's pre-training facial expressions showed typical developmental trends, with highly readable Happy facial expressions, and ambiguous Angry facial displays (Lewis et al. 1987; Odom and Lemond 1972). No changes were observed in TD participants' Happy facial expressions post-training, however this is not unexpected given previous research showing that TD children's performance of voluntary happy displays are comparable to that of adult's (Lewis et al. 1987). Importantly, whereas the ASD group lagged behind the TD group in the quality of their Happy expressions before training, the quality of their happy expressions were comparable to their TD peers after training.

With respect to the Angry expression, enhancements of both ASD and TD participant's Angry expressions were not surprising given previous research demonstrating a marked deficit in the perception of both TD and ASD children's negative displays (Lewis et al. 1987; Macdonald et al. 1989; Odom and Lemond 1972). Increases in expression quality ratings of disgust in the ASD group are consistent with previous research demonstrating similarity between disgust and angry facial displays in both their production (Dailey et al. 2002; Smith and Scott 1997; Susskind et al. 2007), and perception (Aviezer et al. 2008; Bullock and Russell 1984; Ekman and Friesen 1975; Widen and Russel 2010). In studies categorizing facial expressions based on perceptual quality, facial expressions of Disgust have been categorized as Angry displays in 33 % of trials using the facial expression exclusively (Widen and Russel 2010), and in as much as 87 % of trials when body posture and environmental context are also taken into account (Aviezer et al. 2008). With respect to their psychometric properties, facial expressions of Anger and Disgust are similarly classified as both negative in valence and high on arousal according to the circumplex model of affect, with Anger being only slightly greater in arousal (Russell and Barrett 1999). Results from the current study help clarify the relationship between Angry and Disgust facial expressions by providing evidence for the expression quality of disgust in Angry facial displays.

FaceMaze presents a novel and engaging approach to enhance the quality of voluntary facial expressions for typically developing children and children on the autism spectrum. Beyond its entertainment value, the game is a promising intervention tool to improve the quality of facial expressions produced by individuals with ASD. From an embodied cognition perspective, the integration of perceptual and motor facets not only provides a more naturalistic process, but would also allow for an enhancement of all cognitive aspects of emotion expression, such as expression recognition (Atkinson and Adolphs 2005; Deriso et al. 2012; Goldman and Sripada 2005; Niedenthal et al. 2001; Oberman et al. 2007). Furthermore, the quality of voluntary and spontaneous expressions are correlated (Berenbaum and Rotter 1992) suggesting that the two processes may rely on similar cognitive mechanisms (Winkielman et al. 2009a, b). Further research is warranted to investigate whether the directed training in voluntary expression production in ASD participants facilitates the production spontaneous expressions in naturalistic settings. Regardless, individuals on the autism spectrum will benefit from FaceMaze training that has the potential to improve their voluntary facial expressions, and as a consequence enhance their everyday social interactions.

Although other treatment programs targeting facial expression production have shown positive results, they are labor-intensive, requiring one-on-one tutoring with human therapists over the course of several days (Charlop et al. 2010; DeQuiznio et al. 2007; Gena et al. 1996; Stewart and Singh 1995). Personalized training in facial expressions is a "tiring procedure for therapists to use and difficult to use with consistency" (Gena et al. 1996, p. 547). With respect to ASD, these treatments may also be more difficult to implement as a result of co-morbid social anxiety. In contrast, FaceMaze is a promising, cost-effective training program in facial expression production that is engaging for the child and that can be conducted in safe, familiar setting.

Acknowledgments Support for this work was provided by NIH grant NIMH-RC1 MH088633, and the Temporal Dynamics of Learning Center NSF grant SMA-1041755 to James W. Tanaka and Marian S. Bartlett, and the National Science and Engineering Research Councils of Canada to James W. Tanaka. Further thanks goes to the participating families and the Center for Autism Research Technology and Education. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. This article has been prepared from a doctoral dissertation.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental health disorders* (5th ed.). Washington DC: Author.

Atkinson, A., & Adolphs, R. (2005). Visual emotion perception: Mechanisms and processes. In L. Feldman-Barret, P. M. Niedenthal, & P. Winkielman (Eds.), *Emotion and consciousness* (pp. 150–184). New York: Guilford Press.

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., et al. (2008). Angry, disgusted, or afraid? *Psychological Science, 19*(7), 724–732.

Barbu, S., Jouanjean, A., & Allès-Jardel, M. (2001). Behavioural determinants of friendships among preschoolers: An ethological and naturalistic approach. *International Review of Social Psychology, 2*, 75–92.

Bartlett, M. S., Littlewort, G., Frank, M. G., Lainscsek, C., Fasel, I., & Movellan, J. R. (2005). Recognizing facial expression: Machine learning and application to spontaneous behavior. In *IEEE international conference on computer vision and pattern recognition*, pp. 568–573.

Bartlett, M. S., Littlewort, G. C., Frank, M. G., Lainscsek, C., Fasel, I., & Movellan, J. R. (2006). Automatic recognition of facial actions in spontaneous expressions. *Journal of Multimedia, 1*(6), 22–35.

Bartlett, M., & Whitehill, J. (2011). Automated facial expression measurement: Recent applications to basic research in human behavior, learning, and education. In A. Calder, G. Rhodes, J. V. Haxby, & M. H. Johnson (Eds.), *Handbook of face perception*. Oxford: Oxford University Press.

Berenbaum, H., & Rotter, A. (1992). The relationship between spontaneous facial expressions of emotion and voluntary control of facial muscles. *Journal of Nonverbal Behavior, 16*, 179–190.

Boyatzis, C. J., Chazan, E., & Ting, C. Z. (1993). Preschool children's decoding of facial emotions. *Journal of Genetic Psychology, 154*, 375–382.

Bullock, M., & Russell, J. A. (1984). Preschool children's interpretation of facial expressions of emotion. *International Journal of Behavioral Development, 7*, 193–214.

Campos, J. J. (1985). *Current issues in the study of emotion and emotional development*. Paper presented at a symposium on cognition-emotion relations at the meetings of the American Psychological Association, Los Angeles, CA.

Camras, L. A., Oster, H., Campos, J., Campos, R., Ujiie, T., Miyake, K., et al. (1998). Production of emotional facial expressions in European American, Japanese, and Chinese infants. *Developmental Psychology, 34*(4), 616–628.

Charlop, M. H., Dennis, B., Carpenter, M. H., & Greenberg, A. L. (2010). Teaching socially expressive behaviors to children with autism through video modeling. *Education and Treatment of Children, 33*(3), 371–393.

Cockburn, J., Bartlett, M., Tanaka, J., Movellan, J., & Schultz, R. (2008). SmileMaze: A tutoring system in real-time facial expression perception and production in children with Autism Spectrum Disorder. In *Proceedings from the IEEE international conference on automatic face & gesture recognition* (peer-reviewed conference proceeding), pp. 978–986.

Dailey, M. N., Cottrell, G. W., Padgett, C., & Adolphs, R. (2002). EMPATH: A neural network that categorizes facial expressions. *Journal of Cognitive Neuroscience, 14*, 1158–1173.

DeQuiznio, J. A., Townshend, D. B., Sturmey, P., & Poulson, C. L. (2007). Generalized imitation of facial models by children with autism. *Journal of Applied Behavioral Analysis, 40*(4), 755–759.

Deriso, D. M., Susskind, J., Tanaka, J., Winkielman, P., Herrington, J., Schultz, R., et al. (2012). Exploring the facial expression perception-production link using real-time automated facial expression recognition. In *Computer Vision–ECCV 2012. Workshops and demonstrations* (pp. 270–279). Berlin, Heidelberg: Springer.

Donato, G. L., Bartlett, M. S., Hager, J. C., Ekman, P., & Sejnowski, T. J. (1999). Classifying facial actions. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 21*(10), 974–989.

Ekman, P. (1993). Facial expressions of emotion. *American Psychologist, 43*, 384–392.

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues*. Englewood Cliffs, NJ: Prentice-Hall.

Ekman, P., O'Sullivan, M., Frtesett, W. V., & Scherer, K. (1991). Face, voice, and body in detecting deceit. *Journal of Nonverbal Behavior, 15*, 125–135.

Fridlund, A. (1994). *Human facial expression: An evolutionary view*. New York: Academic Press.

Gena, A., Krantz, P. A., McClannahan, L. E., & Poulson, C. L. (1996). Training and generalization of affective behavior displayed by youth with autism. *Journal of Applied Behavior Analysis, 29*, 291–304.

Goldman, A. I., & Sripada, C. S. (2005). Simulationist models of face-based emotion recognition. *Cognition, 94*, 193–213.

Hiatt, S., Campos, J., & Emde, R. (1979). Facial patterning and infant emotional expression: Happiness, surprise, and fear. *Child Development, 50*, 1020–1035.

Izard, C. E., & Malatesta, C. (1987). Perspectives on emotional development: I. Differential emotions theory of early emotional development. In J. Osofsky (Ed.), *Handbook of infant development* (2nd ed., pp. 494–554). New York: Wiley.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2*, 217–250.

Kanner, L. (1968). Autistic disturbances of affective contact. *Acta Paedopsychiatr, 35*(4), 100–136.

Langdell, T. (1981). *Face perception: An approach to the study of autism*. Doctoral thesis, University of London.

Lewis, M., Sullivan, M. W., & Vasen, A. (1987). Making faces: Age and emotion differences in the posing of emotional expressions. *Developmental Psychology, 23*(5), 690–697.

Littlewort, G., Whitehill, J., Wu, T., Fasel, I., Frank, M., Movellan, J., et al. (2011). The computer expression recognition toolbox (CERT). In: *Proceeding IEEE international conference on automatic face and gesture recognition*, pp. 298–305.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., et al. (2000). The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*(3), 205–223.

Loveland, K. A., Tunali-Kotoski, B., Pearson, R., Brelsford, K., Ortegon, J., & Chen, R. (1994). Imitation and expression of facial affect in autism. *Development and Psychopathology, 6*, 433–444.

MacDonald, H., Rutter, M., Howlin, P., Rios, P., Le Conteur, A., Evered, C., et al. (1989). Recognition and expression of emotional cues by autistic and normal adults. *Journal of Child Psychology and Psychiatry, 30*, 865–877.

McIntosh, D. N., Reichmann-Decker, A., Winkielman, P., & Wilbarger, J. L. (2006). When the social mirror breaks: Deficits in automatic, but not voluntary, mimicry of emotional facial expressions in autism. *Developmental Science, 9*(3), 295–302.

Miller, P., & Eisenberg, N. (1988). The relation of empathy to aggressive and externalizing/antisocial behavior. *Psychological Bulletin, 103*, 324–344.

Niedenthal, P. M., Brauer, M., Halberstadt, J., & Innes-Ker, A. (2001). When did her smile drop? Facial mimicry and the influence of emotional state on the detection of change in emotional expression. *Cognition and Emotion, 15*, 853–864.

Oberman, L. M., Winkielman, P., & Ramachandran, V. S. (2007). Face to face: Blocking facial mimicry can selectively impair recognition of emotional expressions. *Social Neuroscience, 2*, 167–178.

Odom, R. D., & Lemond, C. M. (1972). Developmental differences in the perception and production of facial expressions. *Child Development, 43*, 359–369.

Rogers, S. J., Hepburn, S. L., Stackhouse, T., & Wehner, E. (2003). Imitation performance in toddlers with autism and those with other developmental disorders. *Journal of Child Psychology and Psychiatry, 44*, 763–781.

Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology, 76*(5), 805–819.

Smith, C. A., & Scott, H. S. (1997). A componential approach to the meaning of facial expressions. In J. A. Russell & J. M. Fernández-Dols (Eds.), *The psychology of facial expression* (pp. 229–254). Cambridge: Cambridge University Press.

Stewart, C. A., & Singh, N. N. (1995). Enhancing the recognition and production of facial expressions of emotion by children with mental retardation. *Research in Developmental Disabilities, 16*(5), 365–382.

Susskind, J. M., Littlewort, G. C., Bartlett, M. S., Movellan, J. R., & Anderson, A. K. (2007). Human and computer recognition of facial expressions of emotion. *Neuropsychologia, 45*(1), 152–162.

Widen, S. C., & Russel, A. R. (2010). Differentiation of preschooler's categories of emotion. *Emotion, 10*(5), 651–661.

Williams, J. H., Whiten, A., Suddendorf, T., & Perrett, D. I. (2001). Imitation, mirror neurons and autism. *Neuroscience and Biobehavioral Reviews, 25*, 287–295.

Winkielman, P., McIntosh, D. N., & Oberman, L. (2009a). Embodied and disembodied emotion processing: Learning from and about typical and autistic individuals. *Emotion Review, 2*, 178–190.

Winkielman, P., Niedenthal, P., & Oberman, L. M. (2009b). Embodied perspective on emotion–cognition interactions. In J. Pineda (Ed.), *Mirror neuron systems* (pp. 235–257). New York: Humana Press.