FaceMaze: An Embodied Cognition Approach To Facial Expression Production in Autism Spectrum Disorder

by

Iris Gordon
M.Sc., University of Victoria, 2010
B.Sc., University of Toronto, 2007

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Psychology

Supervisory Committee

FaceMaze: An Embodied Cognition Approach To Facial Expression Production in Autism Spectrum Disorder

by

Iris Gordon
M.Sc., University of Victoria, 2010
B.Sc., University of Toronto, 2007

Supervisory Committee

Dr. James W. Tanaka (Department of Psychology)
Supervisor

Dr. Daniel Bub (Department of Psychology)
Departmental Member

Dr. Grace Iarocci (Department of Psychology)
Outside Member

Abstract

Dr. James W. Tanaka (Department of Psychology)
Supervisor

Dr. Daniel Bub (Department of Psychology)
Departmental Member

Dr. Grace Iarocci (Department of Psychology)
Outside Member

Individuals with Autism Spectrum Disorder (ASD) are typified by deficits in social communication, including flat and disorganized affect. Previous research investigating affect production in ASD has demonstrated that individuals on the spectrum show impairments in posing, but not mimicking facial expressions. These findings thus point to a deficit in ASD individuals' integration of sensory/motor facets in the cognitive representation of a facial expression, and not a deficit in motor or sensory ability. The goal of the current project was to validate a computer-based intervention that targets facial expression production using methods ground in embodied cognition to connect between the sensory and motor facets of facial displays. The "FaceMaze" is a pac-man like game in which players navigate through a maze of obstacles, and are required to produce high-quality facial expressions in order to overcome obstacles. FaceMaze relies on the Computer Expression Recognition Toolbox (CERT) program, which analyzes user's real-time facial expressions and provides feedback based on the Facial Action Coding System (FACS).

In the first part of this project, the FaceMaze was validated using a typically developing (TD) adult population. In Experiment 1, participants were prompted to produce expressions of "Happy", "Angry" and "Surprise" before and after playing

FaceMaze. Electromyography (EMG) analysis targeted three expression-specific facial muscles: Zygomaticus Major (ZM, Happy), Corrugator Supercilii (CS, Angry) and Obicularis Oculi (OO, Surprise). Results showed that relative to pre-game productions, an increase in activation in the ZM for happy expressions, and an increase in CS response for angry expressions was observed after playing the corresponding version of FaceMaze. Critically, no change in muscle activity for the control expression "Surprise" was observed. In Experiment 2, the perceived quality of facial expressions after FaceMaze/ CERT training was compared to those produced after traditional FACS training. "Happy," "Angry" and "Surprise" expressions were videotaped before and after the FaceMaze game and FACS training, and productions were assessed by a group of naïve raters. Whereas observers rated post-Happy expressions as happier for both FaceMaze and FACS, only the post-Angry expressions in the FaceMaze condition were rated as angrier and less happy after training.

In the second half of this project, the efficacy of the FaceMaze was validated by children with ASD, and age- and IQ-matched, typically developing (TD) controls. In Experiment 3 (in press), children were asked to pose "Happy ", "Angry", and "Surprise" expressions before and after game-play. Expressions were video-recorded and presented to naïve raters who were required to assess video-clips on expression quality. Findings show that the ASD groups' post-FaceMaze "Happy" and "Angry" expressions were higher in quality than their pre-FaceMaze productions. TD children also showed higher expression quality ratings for the "Angry" expression post-gameplay, but no enhancement of the "Happy" expression was found after FaceMaze. Moreover, the ASD groups' post-FaceMaze expressions were rated as equal in quality to those of the TD

group. These findings not only underscore the fidelity of the FaceMaze game in enhancing facial expression production, but also provide support for a theory of disordered embodied cognition in ASD.

Table of Contents

List of Tables

List of Figures

Acknowledgments

I would like to give a special thanks to the members of the Center for Autism Research Technology and Education (CARTE), and the children who participated in CARTE's FaceLabs without whom this project could not have be completed. I would also like to thank my supervisor, James Tanaka, and my committee member, Daniel Bub, for their constant guidance throughout my graduate career. Special thanks goes to my mother and sister for being a continued source of love, support, and insanity that makes me laugh. Finally, I would like to extend a warm thanks to my friends at The Bubble Tea Place for their support in keeping me caffeinated and inspired.

Dedication

I would like to dedicate the following work to my mother and sister; see? I do have a brain.

Chapter 1

General Introduction

Traditional theories of facial displays in animals have argued that facial expressions were instrumental, in that manipulation of facial features was strictly functional for executing a behaviour. For example, facial displays of anger were instrumental in baring teeth in preparation of attack, thus angry facial displays evolved to include grimaces (Darwin, 1872/1965). Darwin, however, was the first to recognize the connection between facial displays and emotional experience. Furthermore, in his publication *The Expressions of Emotion in Man and Animals* (1872), Darwin also argued that facial expressions were not an epiphenomenon of emotional experience, but served a communicative function by conveying the animal's internal state to others. As a result, facial expressions were facilitative in regulating social interaction through signals of approach or avoidance (Darwin, 1872). In accordance with Darwin's theories, more recent research has found that facial expressions help initiate, modify and regulate patterns of social interaction (Barbu, Jouanjean, Allès-Jardel, 2001; Boyatzis, Chazan, & Ting, 1993), by revealing information about a person's momentary affective state. As a result, facial expressions are subject to scrutiny in social situations (Ekman, 1993; Izard and Malatesta, 1987; Fridlund, 1994), thus producing facial expressions that are ambiguous, inconsistent with social expectations, or are 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. The French neurologist Guillaume-Benjamin-Amand-Duchenne (de Boulogne) was the first to demonstrate the dissociation between facial expressions and emotions, using electric stimulation to manipulate facial muscles into recognizable configurations in the absence of emotion (1862/1990). More recent research has also demonstrated that an externalized emotional display, such as a facial expression, can also be expressed in the presence of an incongruent emotion, such as in cases of deception in which participants produce happy facial expressions to mask feelings of disgust (Ekman and Friesen, 1975; Ekman, O'Sulliven, Friesen and Scherer, 1991) or sympathy (Miller and Eisenberg, 1988). Conversely, an emotion can be experienced internally without its externalization as a facial expression or body gesture (Campos, 1985; Camras, Oster, Campos, Campos, Ujiie, Miyake, Wang, Meng, 1998; Hiatt, Campos and Emde, 1979). This dissociation therefore 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 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. How this is possible is best addressed by theories of embodied cognition.

Embodied Cognition and Facial Expressions

According to the embodied cognition approach, the contents of the mind, such as mental representations, are largely influenced by the body, such as perception or moods. In contrast to traditional cognitive theories (Darwin, 1872/1965; Ekman, 1973; Izard, 1977; Tomkins, 1962) that treat the body as an extension of the mind, theories of embodied cognition describe a bi-directional relationship in which the form and function of the body constrain and influence the mind, reciprocally. These theories also assume that the cognitive representations of a process or knowledge also include the sensory and modal information associated with them. Thus, activation of the cognitive process in part re-activates the sensory and motor modalities, and vice-versa (Winkielman, Niedenthal & Oberman, 2009).

Evidence for the bi-directionality between emotions and facial expression production has been demonstrated in studies in which physiological changes in participant's heart rate, skin conductance, body temperature, and muscle tension were recorded in response to evoking emotional states (i.e. "reliving" emotions) or producing constellations of facial movements (i.e. voluntary facial expressions). Findings not only revealed a distinct pattern of autonomic arousal identifying anger, fear, sadness, disgust, and happy and surprise emotions, but also that emulation of facial expressions elicited more potent physiological responses than evoking emotional states without any concurrent facial gestures, providing support for an embodied view of facial expressions (Ekman, Levenson, Friesen, 1983; Levenson, Ekman, Friesen, 1990). Furthermore, the extent to which the motor modality also affects our cognitions in expression production has been demonstrated in studies investigating the effects of voluntary facial displays on perception, or “facial feedback hypothesis” (Strack, Martin, & Stepper, 1988), in neuroimaging research investigating the effects of facial expression inhibition on amygdala activation, (Hennenlotter, Dresel, Castrop, Ceballos-Baumann, Wohlschlager, & Haslinger, 2009), and in studies assessing activation of emotion-related brain regions, such as the somatosensory cortex, in response to volitional facial expressions (Damasio, Grabowski, Bechara, Damasio, Ponto, Parvizi, Hichwa, 2000; Wild, Erb, Eyb, Bartels, Grodd, 2003), providing support for an embodied approach to facial expressions of emotion.

One mechanism that may explain the integration of sensory and motor modalities into the cognitive representation of emotional expressions is mimicry, that is, facial expressions produced in the presence of a model. From a cognitive perspective, mimicry has been interpreted as a “meeting of the minds”, in which re-enactment of other’s behaviors, such as a facial expression, elicits the corresponding physiological state, such as emotions, and thus gives the mimic insight into another’s cognitions (Atkinson & Adolphs, 2005; Dimberg, 1982, Niedenthal, Brauer, Halberstadt, & Innes-Ker, 2001). Evidence for this theory not only comes from research showing unconscious mimicry of other’s facial expressions during an EMG task (Dimberg, Thunberg, & Elmehed, 2000), but also from studies in which inhibiting facial mimicry affects perception. For example, in a study investigating the effects of facial movement on the perception of ambiguous expressions (Niedenthal, Brauer, Halberstadt, & Innes-Ker, 2001), participants were required to report when a morphed face changed from happy to sad, and vice-versa. In one condition, however, subjects were required to hold a pen in their mouths, blocking extraneous facial movement. Findings revealed that participants in the pen condition detected changes in facial expression later than those in a no-pen condition, providing evidence for the importance of mimicry in the interpretation of facial expressions (Niedenthal, Brauer, Halberstadt, & Innes-Ker, 2001). Replicating these findings, Oberman, Winkielman, Ramachandran, (2007) extended the previous study by measuring the specificity of expression blocking in the pen condition by measuring the effects on several facial expressions (happy, sad, fear and disgust). Furthermore, researchers also controlled for the effects of muscle activation by including both a gum-chewing condition that would activate facial muscles intermittently, and a pen biting condition that would continuously activate facial muscles. Findings show that holding a pen in one's teeth disproportionately affected the perception of happy facial displays when compared to disgust, fear and sad facial expressions as a result of engaging the zygomaticus (cheek) muscles involved in the happy facial expressions, and interfering with mimicry (Oberman et al., 2007). Thus, mimicry is shown to be an important facet in the social communicative aspect of facial expressions, by providing a mechanism by which to internalize and interpret other's expressions.

From a developmental perspective, mimicry also provides a learning mechanism that allows for the internalization and fine-tuning of motor behaviours (Piaget, 1951/2013; Vygotsky, 1967). By first presenting an ideal action, and then allowing the child (mimic) to re-enact and refine their performance, mimicry allows for the integration of both motor action and perception within the cognitive representation. Deficits in mimicry can thus result in disorders of cognition by dissociating the sensory processes involved in the perception of facial expressions from the motor action involved in their production.

Autism, Mimicry, and Facial Expressions

Researchers have investigated mimicry of facial expressions in Autism Spectrum Disorder (ASD) by contrasting facial productions made spontaneously and voluntarily (McIntosh, Reichmann-Decker, Winkielman, Wilbarger, 2006). Autism Spectrum Disorder is a pervasive developmental disorder that is typified by deficits in social communication (American Psychological Association, 2000), including facial expressions production (see chapter 3) (Lord, Risi, Lambrecht, Cook, Leventhal, DiLavore, Pickles, & Rutter, 2000). In the McIntosh et al. study, individuals with ASD, and age- and verbal-IQ matched TD controls were first asked to simply watch a screen as pictures of happy and angry faces were presented. Following this, participants viewed the same images, except that they were explicitly prompted to produce a facial expression "just like this one". In light of the fast and subtle nature of micro-expressions that occur in mimicry, expression-specific EMG recordings were obtained from both conditions, and compared across groups. Findings show that when compared to their TD peers, individuals with ASD were less likely to produce any spontaneous muscle response to either happy or angry facial expressions, however individuals with ASD volitionally activated expression-related muscles at similar rates to TD controls when explicitly prompted to mimic an expression. Simply put, individuals with ASD showed impairments in spontaneous mimicry, but not in voluntary mimicry when compared to TD controls (McIntosh et al., 2006). Extending this study, Oberman, Winkielman, Ramachandran (2009) were interested in determining whether the deficits in mimicry of facial affect in ASD resulted from a perceptual inability to recognize expressions quickly, or a general sensory-motor deficit in mimicking facial expressions. Replicating the previous study, researchers expanded the stimulus set to include expressions of happy, sad, angry, fear, disgust, and neutral, and varied stimulus presentations times from extremely short to long. Results revealed a general temporal deficit in ASD participant's spontaneous facial displays, with mimicry occurring significantly later than that of TD controls. In contrast, no differences were found between ASD and TD participants in the timing of mimicked voluntary expressions, providing further evidence that the deficit observed in ASD was sensory-motor based, and not perceptual (Oberman, Winkielman, Ramachandran, 2009).

Taken together, research using an embodied cognition approach has been able to elucidate the mechanism implicated in the facial expression production deficit in ASD. Specifically, individuals with ASD show disordered spontaneous mimicry, and this deficit may subsequently affect the development of the expression concept by dissociating the sensory and motor components involved in facial expression production. Despite this deficit, however, it may be possible to entrain voluntary facial expression production by scaffolding learning on the spared voluntary mimicry abilities in ASD. By explicitly encouraging the mimicry of a readable facial expression, the motoric (muscle activation) and sensory information (proprioceptive information, perception) can then become integrated in the expression concept, allowing for higher quality voluntary expression production. Such facial expression training paradigms exist, however not without their limitations. Most, if not all expression training programs designed to teach facial expression production rely on some variant of the Facial Action Coding System (FACS) training procedure (see chapter 2), in which trainees are shown videos explicating muscle movements, produce facial muscle movements directed by coaches, and receive feedback from instruction and/or mirrors (Charlop, Dennis, Carpenter, and Greenberg, 2010; DeQuiznio, Townsend, Sturmey & Poulson, 2007; Gena, Krantz, McClannahan, & Poulson, 1996; Stewart and Singh, 1995). Whereas these programs have shown positive results, the need for one-on-one tutoring with human therapists over the course of several days was "a tiring procedure for therapists to use, and difficult to use with consistency" (Gena et al., 1996, p. 547), and would be difficult to implement with populations suffering from deficits in language or co-morbid social anxiety. Thus, the goal of the current project is to validate a training paradigm – the "FaceMaze" – that targets facial expression production while circumventing the aforementioned pitfalls. First, Facemaze is computer-based, ensuring reliable training procedures that can be executed consistently over long periods of time. Moreover, Facemaze requires little verbal explanation and does not require any linguistic ability to play, thus can also be used by individuals with deficits in language comprehension or production. Furthermore, computer-based paradigms are less threatening to individuals suffering from social anxiety, thus may present a more effective training paradigm then one-on-one tutelage. Critically, FaceMaze emulates the naturally occurring developmental trajectory by relying on embodied actions, while also increasing their (cognitive) saliency to allow for conscious control, and thus also scaffolds on the natural learning mechanisms.

Training Facial Expressions Through Embodied Cognition

The goal of this project is to validate an interactive, computer-based intervention – the "FaceMaze" – that targets facial expression production using the spared mimicry abilities in ASD. In FaceMaze, players navigate through a maze in order to obtain tokens, while overcoming obstacles by producing matching facial expressions. Player's facial expressions are captured in real-time, using the webcam and the Computer Expression Recognition Toolbox (CERT), which analyses the expression's quality and provides real-time feedback to the player. Critically, FaceMaze allows for the sensory-motor integration of facial expressions by associating the facial configuration (motor) with the feeling (proprioceptive sensation) of producing facial expressions. In Chapter 2, the efficacy of the FaceMaze training paradigm in enhancing facial expression production was validated using physiological measures (electromyography, or EMG), and observer ratings in an adult population. First, participants were prompted to produce expressions of "Happy", "Angry" and "Surprise" before and after playing FaceMaze, while EMG analysis targeted three expression-specific facial muscles: Zygomaticus Major (ZM, Happy), Corrugator Supercilii (CS, Angry) and Obicularis Oculi (OO, Surprise). Results showed that relative to pre-game productions, an increase in activation in the ZM for happy expressions, and an increase in CS response for angry expressions was observed after playing the corresponding version of FaceMaze. Critically, no change in muscle activity for the control expression "Surprise" was observed.

In light of facial expressions' communicative function, a subsequent study was carried out in order to determine if the perceived quality of facial expressions was enhanced after FaceMaze training, as compared to expressions entrained by another validated expression-training paradigm, namely the FACS. Participant's "Happy," "Angry" and "Surprise" expressions were videotaped before and after the FaceMaze game and FACS training, and video-clips were presented to a group of naïve raters, who rated the video-clips for expression quality on six basic emotion scales of happy, angry, sad, surprise, fear and disgust. Whereas observers rated post-Happy expressions as happier for both FaceMaze and FACS, only the post-Angry expressions in the FaceMaze condition were rated as angrier, and less happy after training.

In order to determine the efficacy of the FaceMaze game in changing facial expression quality in children with Autism, facial expression production in ASD children and IQ-matched, typically developing (TD) controls, was compared using observer ratings. In Chapter 3 (Gordon, Pierce, Bartlett, & Tanaka, 2014), ASD and TD children played one five-minute block of FaceMaze containing “Happy” obstacles, and another five-minute block containing “Angry” obstacles. Videotapes of the children posing “happy,” “angry” and “surprise” expressions were recorded before and after each block. Naïve non-ASD adult observers rated the quality of the children’s productions across the six basic emotions of happy, angry, sad, surprise, fear and disgust. The results showed that ASD children’s productions of the “happy” and “angry” expressions were rated as higher in quality after playing the Happy and Angry versions of FaceMaze, respectively, than their pre-FaceMaze versions. For the TD group, only the “angry” expressions were rated higher in quality after playing the Angry version of FaceMaze. Whereas the ASD group’s expression quality ratings were lower than their TD counterparts before the FaceMaze intervention, no differences in expression quality ratings between the ASD and TD children were found after playing the FaceMaze game.

Finally, Chapter 4 will review the previous experiments’ findings with respect to embodied cognition, demonstrating that the facial expression production deficit in ASD does not result from a disorder in motor ability. Rather, deficits in facial expression production are attributed to a disorder in the expression concept that has not fully integrated the sensory and motor components involved in facial displays. By allowing for conscious awareness of facial movement during expression production, FaceMaze allows for the explicit connection between the proprioceptive sensation (sensory) of muscle movements (motor) involved in producing a specific facial expression to be integrated in the player's expression concept as a facet of that expression.

Chapter 2

Experiment 1

Introduction

Facial expressions are not only determined by an individual's expression concept, but are also reliant on our ability to manipulate facial muscles. Whereas Duchenne was the first to demonstrate the relationship between muscle activation and the generation of facial displays (1862/1990), a more systematic investigation of facial expression muscle configuration was carried out by Rusalova, Izard, and Simonov (1975). In this study, trained actors and control participants were asked to produce facial expressions of happy, sad, fear, and angry, while attempting to either re-experience the emotion associated with the expression, mask another emotion with the given expression, or merely produce the expression without emotion. Electromyographic (EMG) activity was recorded from the four separate muscle groups of venter frontalis (forehead), corrugator supercilii (inner brow), masseter (jaw), and depressor anguli oris (cheek), and measures of heart-rate were taken as indicators of emotional experience. Findings show that a similar pattern of muscle activation was observed when the actors were asked to produce facial expressions with and without corresponding emotions. Specifically, when comparing patterns of muscle activation, activation in the venter frontalis was largest for expressions of fear, activation of the corrugator supercilii was largest for expressions of sad, activation of the masseter was largest for expressions of anger, and activation of the depressor anguli oris was largest for expressions of happy. Interestingly, control participants also showed similar patterns of muscle activation, but only for the happy emotion. Expressions of sadness were similar to that of the actors, however this pattern of muscle activation was not different from the other negative emotions of anger and fear in the control participants. Thus, whereas facial expressions lend themselves to particular patterns of muscle activation, their voluntary enactment required explicit training. Furthermore, changes in heart-rate were similar for both the actors and control participants, with fluctuations observed only for the condition in which participants were required to re-live the emotion, underscoring the similarity of facial expression production in situations where expressions are produced with and without emotions.

Studies investigating the production of facial expressions using EMG have highlighted the importance of the zygomaticus major in the production of facial expressions associated with positive emotions, and the currogator supercilii in the generation of facial displays associated with negative emotions (Cacioppo & Petty, 1981; Schwartz, 1975; Hjortsjö, 1970; Schwartz, Fair, Salt, Mandel, & Klerman, 1976; Schwartz, Brown, & Ahern, 1980; for a review, see Fridlund and Izard, 1983; Dimberg 1990). The zygomaticus major muscle is found bilaterally on the face, attached to the cheekbone and the upper corner of the lip. Contraction of zygomaticus is responsible for flexing the lips superiorly and posteriorly, resulting in a "smile". The corrugator supercilii is located in the middle portion of the eyebrow, spanning diagonally until the top of the nose arch. Constriction of the corrugator supercilii results in the furrowing of the brow, a critical part in the production of a "scowl". The sensitivity with which facial EMG can detect activation in these facial muscles has been demonstrated in a study by Cacioppo, Petty, Losch & Kim (1986), in which researchers used facial EMG to detect changes in facial muscle movement across lower, non-visible expression intensities. Participants were shown either mildly or moderately positive images accompanied by a pleasant tone, and mildly or moderately negative images paired with a negative tone and were required to rate how much they liked the image on a 9-point Likert scale, ranging from 1 (disliked) to 9 (like) to further corroborate stimulus pleasantness. Meanwhile, EMG measures were obtained from the zygomaticus major (cheek), corrugator supercilii (eyebrow), obicularis oculi (lower eyelid), medial frontalis (forehead), and the obicularis oris (lip), and participants' facial expressions were also video-recorded. In order to determine the extent to which expressions were visually discernable, video-recordings of participants' faces were subsequently presented to naïve raters who were asked to determine if the participants were viewing affectively positive or negative scenes. Consistent with previous research, EMG results revealed that activation of the zygomaticus major was enhanced when participants viewed a mildly or moderately positive scene, and activation of the corrugator supercilii occurred during presentations of mildly or moderately negative stimuli. In addition, activation was correlated with stimulus intensity, such that moderately affective scenes generated more EMG impulse than those of mildly affective scenes. Furthermore, participant's stimulus ratings showed higher Likert ratings for affectively positive scenes, and lower ratings for negative scenes, corroborating EMG results. More importantly, accuracy of naïve raters' categorizations was at chance, underscoring the fidelity of EMG recording in detecting facial muscle activation despite participants not producing overt facial expressions (Cacioppo, Petty, Losch and Kim, 1986). Thus, EMG is a reliable and highly sensitive measure of facial muscle activation that has shown great consistency with respect to facial affect categorization and intensity.

In the current experiment, participants were asked to pose the facial expressions of “happy,” “angry” and “surprise” while muscle activity was recorded with EMG. Participants then played the Happy or Angry version of the FaceMaze game, followed by the EMG post-training assessment that was identical to the EMG pre-training assessment. If FaceMaze selectively enhances expression production, we predicted increased post-game activity of the zygomaticus major after playing “happy” maze and increased activity of the currogator supercili after playing “angry” maze, and no change in EMG activity when posing surprise. Alternatively, if training has no effect on the voluntary execution of happy and angry expressions, we would expect little or no difference between pre- and post-training productions. Critically, if changes are observed in the happy and angry conditions but surprise expressions are not altered, then the changes detected reflect an implicit learning process and not an artifact of repeated muscle movement. Moreover, this experiment would serve as a physical confirmation that the activation of specific muscle groups the CERT was targeting was being altered.

Methods

Participants

Thirty-six undergraduate students from the University of Victoria participated in this study. Six participants were discarded as a result of technical issues, one was excluded as a result of attrition, and another four were removed from analysis because of inability to perform a facial expression (see procedure for further discussion). Of the remaining 25 participants (six male), ages 18 to 24 years (M = 18.9 years), six participants were left-handed. All had normal or corrected-to-normal visual acuity. None had any history of brain injury or trauma. Informed consent was obtained from all participants prior to the experiment, and students were given two credits toward class requirements.

Materials

Stimuli comprised of the emotion words “Happy” “Angry” and “Surprised” presented on a 14” computer monitor. Words were in white type-font on a black background. Words were 126 x 46 pixels in size, allowing for a visual angle of 6.39 degrees in the horizontal field and 2.34 degrees in the vertical field.

The Computer Expression Recognition Toolbox (CERT). To implement our training program, we employed the Computer Emotion Recognition Toolbox (CERT) developed by Bartlett and colleagues (Littlewort et al., 2011; Bartlett, Littlewort, Frank, Lainscsek, & Movellan, 2006; Bartlett et al., 2005). To maximize the capabilities of CERT, we designed the “FaceMaze” game in which a player navigates a pac-man-like figure through a series of corridors, and removes face tokens by producing the appropriate happy or angry expressions (Cockburn, Bartlett, Tanaka, Movellan, & Schultz, 2008). CERT detects the target expression via webcam input, rates the quality of the expression and provides real time feedback to the player.

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) (see Figure 1). 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 AUs according to 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. 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).

CERT implements a set of 6 basic emotion detectors, and an additional neutral expression detector, by feeding the final AU estimates into a multivariate logistic regression (MLR) classifier. The classifier was trained on the AU intensities, as estimated by CERT, on the Cohn-Kanade dataset (Tian, Kanade, & Cohn, 2001), and its corresponding emotion labels. MLR outputs the posterior probability of each emotion given the AU intensities as inputs. Performance of the basic emotion detectors was measured on the 26 subjects in the updated CK+ database that were not in the CK training set. Accuracy was measured in two different ways: (a) mean percent correct using a 2-alternative forced choice (corresponding to area under the ROC curve) was 98.8%. Mean percent correct on a 7-alternative forced choice was 87.2% correct. More information on CERT design and performance is available in Littlewort et al., (2011).

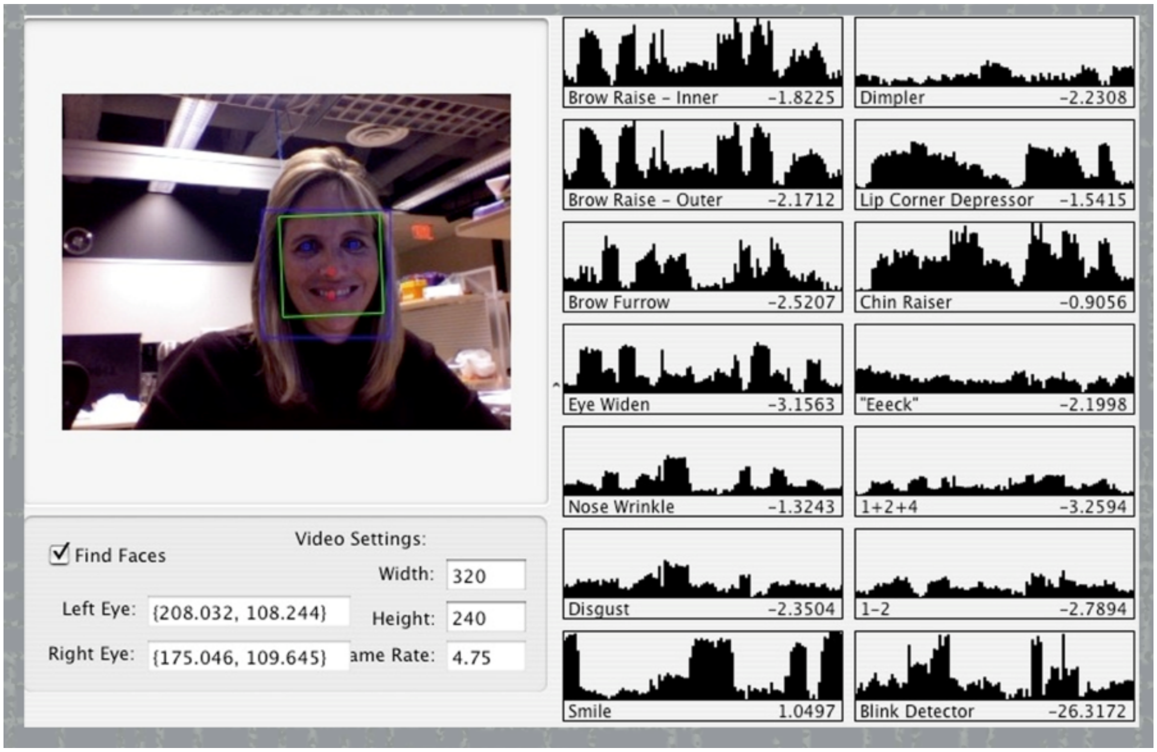

Figure 1. Computer Expression Recognition Toolbox interface. The image of an individual's face is captured in real-time, via live video-stream (left). The face is detected (green and blue squares), and analysis of FAUs is performed while output is presented (right).

FaceMaze. "FaceMaze" is a computer game in which users navigate through a maze with a PacMan-like character (blue colored neutral face) using the arrow keys, with the goal of collecting as many tokens littered about the maze as possible. The challenge of the game is to overcome the barriers blocking one's path (see Figure 2). The obstacles are differently colored faces depicting expressions (such as a yellow happy face or a red angry face), which are removed when the user correctly produces the obstacle's facial expression (see Figure 2). 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 seen and not one that may be easier to produce for the player.

The FaceMaze game was divided into two levels: HappyMaze and AngryMaze. In HappyMaze, the facial expression to be performed in order to remove game obstacles was a smile, interpreted as activation of the "smile detector". Activation of the smile detector was operationalized as the tensing of the zygomaticus major, resulting in a visible upturned inflection of the lip detected by CERT. A scowl, operationalized as the tensing of the 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.

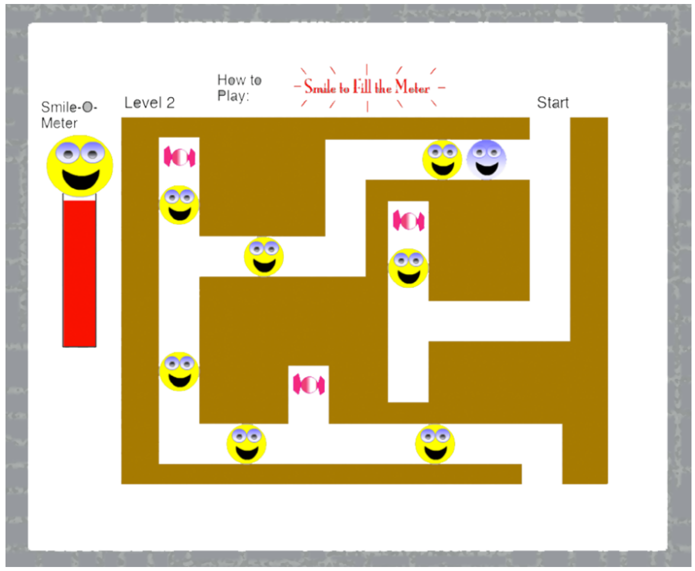

Figure 2. The "Happy" level of FaceMaze. The player moves a blue, neutral PacMan-like face throughout the maze, with the goal of collecting tokens (pink candy wrappers). In order to remove obstacles in their path, players must mimic the facial expression displayed by the obstacle. In HappyMaze, obstacles 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.

Procedure

EMG methods

The participants' muscle activation was recorded from 3 pairs of 4mm electrodes placed over the zygomaticus major (cheek), orbicularis oculi (eye) and corrugator supercilii (eyebrow) as a measure of happy, angry and surprise expressions respectively (see Figure 3) using the Brain Vision Recorder software (Version 1.3, Brainproducts, GmbH, Munich, Germany). Channels were referenced to a common ground placed on the forehead, away from the measured muscle groups. All impedances were sampled digitally at 1000 Hz with a bandpass filter of 0.017 Hz to 250 Hz online (Quick Amp, BrainProducts, GmbH, Munich, Germany).

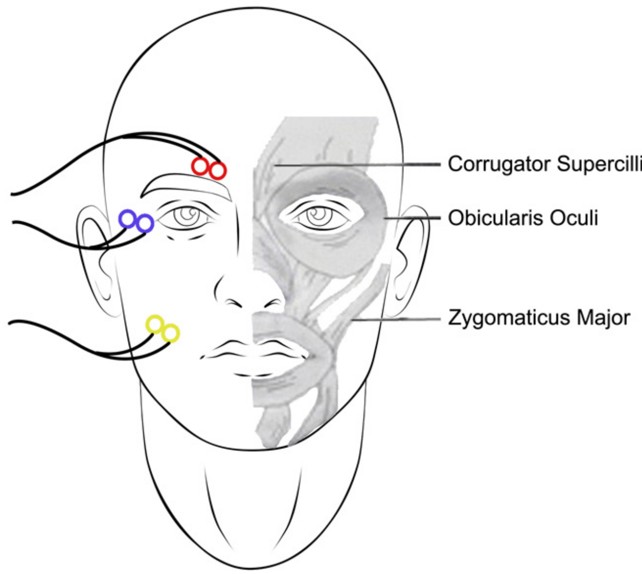

Figure 3. Diagram presenting musculature of face on the left half, with corresponding EMG electrode placement on the right half.

Data obtained was then subjected to several filtering process offline as follows; first, EMG data was segmented into 2000 ms epochs, beginning 500 ms before stimulus onset and subsiding 1500 ms after stimulus onset, thus the start of each epoch coincided with the blank screen stimulus that preceded the presentation of the emotion word stimulus. Epochs began at 500 ms before stimulus onset in order to generate a baseline comparison. EMG epochs were then filtered using a pass-band filter of 10 Hz – 200 Hz, rectified and integrated.

For both “Happy” and “Angry” blocks, trials were divided into pre- and post-training productions of Surprised, Happy, and Angry, exclusively, and EMG activation was then averaged across trials. As a result, averaged segments representing the mean muscle activation of each emotion word in all presented conditions were created; for the "Happy" block, pre-training Happy, post-training Happy, pre-training Surprised, and post-test Surprised, were averaged. For the "Angry" Block, pre-training Angry, post-training Angry, pre-training Surprise, and post-training Surprise were averaged. These averages were utilized in subsequent analysis techniques.

Pre- and post-training expression production.

After the electromyography (EMG) electrodes were applied, subjects were read instructions presented on-screen, directing them to "make the facial expression they would naturally make if they were feeling the presented emotion word". Participants then received a practice trial wherein an emotion word was shown, and subjects performed the associated expression. Following completion of the practice phase, participants were given an opportunity to ask the experimenter any questions they may have had and proceeded to the experimental phase.

The pre-test/post-training productions used in the experiment consisted of two blocks, one "Happy" and one "Angry" that was counterbalanced across participants. In each block, participants completed a pre-training assessment, the FaceMaze activity, and a post-training assessment. Pre/post training assessments were similar, consisting of 30 trials of which half the emotion word (i.e. "Happy" or "Angry") was presented, and the other half the control word, "Surprised". The emotion words were presented on a computer monitor. Trials consisted of a blank screen with a fixation cross at the center for 1000 ms, followed by a blank screen lasting 500 ms, and followed by an emotion word for 1500 ms. No feedback was given with regards to expression produced.

Results

EMG measures were subjected to a 2 (time: pre, post) x 3 (expression: happy, angry, surprise) x 3 (muscle: zygomaticus major, obicularis oculi, corrugator supercilii) within-subjects, repeated-measures ANOVA. All within-subjects factors were Greenhouse-Geisser corrected, and Bonferonni adjustments were performed.

A significant main effect of time F (1, 24) = 9.01, $p < 0.01$, $\eta_p^2 = 0.273$, was found, with pre-FaceMaze muscle activation (M = 19.47, Se = 1.63) reliably smaller than post-FaceMaze muscle activation (M = 21.67, Se = 1.77). A significant main effect of Emotion was also found, F (1.754, 42.092) = 41.00, $p < 0.00$, $\eta_p^2 = 0.631$, with muscle activation significantly larger for the Happy expression (M = 37.36, Se = 3.77) than that of the Angry expression (M = 22.22, Se = 2.19), the Happy-control Surprise expression (M = 11.46, Se = 1.19), and the Angry-control Surprise expression (M = 11.25, Se = 1.33). Furthermore muscle activation for the Angry expression was reliably larger than the Happy-control Surprise expression and the Angry-control Surprise expression, and no significant difference was found in muscle activation between the Happy-control and Angry-control Surprise expressions. In order of magnitude, muscle activation was largest for the Happy expression, followed by the Angry expression, then the control Surprise expressions. A reliable main effect of Muscle, F (1.62, 38.85) = 8.85, $p = 0.01$, $\eta_p^2 =$ 0.269, was also found, with activation in the zygomaticus major (M = 26.23, Se = 2.87) reliably larger than that of the obicularis oculi (M = 15.38, Se = 1.75), and similar to that of the corrugator supercilii (M = 20.11, Se = 1.92). Furthermore, currugator supercilii activation was significantly larger than that of the obicularis oculi. In order of magnitude, activation of the zygomaticus major was the largest, followed by the currogator supercilii and then the obicularis oculi. No significant interaction of Time x Expression, F (1.79, 42.92) = 2.42, $p = 0.11$, $\eta_p^2 = 0.09$, and no significant interaction of Time x Muscle was observed, F (1.87, 44.79) = 1.48, $p = 0.24$, $\eta_p^2 = 0.06$, were found. A reliable interaction of Time x Expression, F (2.29, 55.00) = 62.85, $p < 0.00$, $\eta_p^2 = 0.72$, as well as a significant interaction of Time x Expression x Muscle, F (3.69, 88.65) = 4.67, $p < 0.005$, $\eta_p^2 = 0.16$, was also observed.

Consistent with our prediction, in the HappyMaze condition, greater activation was found in zygomaticus major channels, t (24) = -2.21, $p < 0.05$, during the post-HappyMaze block (M = 81.54, Se = 8.43) when compared to the pre-HappyMaze block (M = 73.29, Se = 8.87). However, no differences were found for orbicularis oculi activation, t (24) = -0.64, $p = 0.53$, nor for the corrugator supercilii, t (24) = -0.56, p = 0.58, between pre-HappyMaze, and post-HappyMaze activation (see Figure 4).

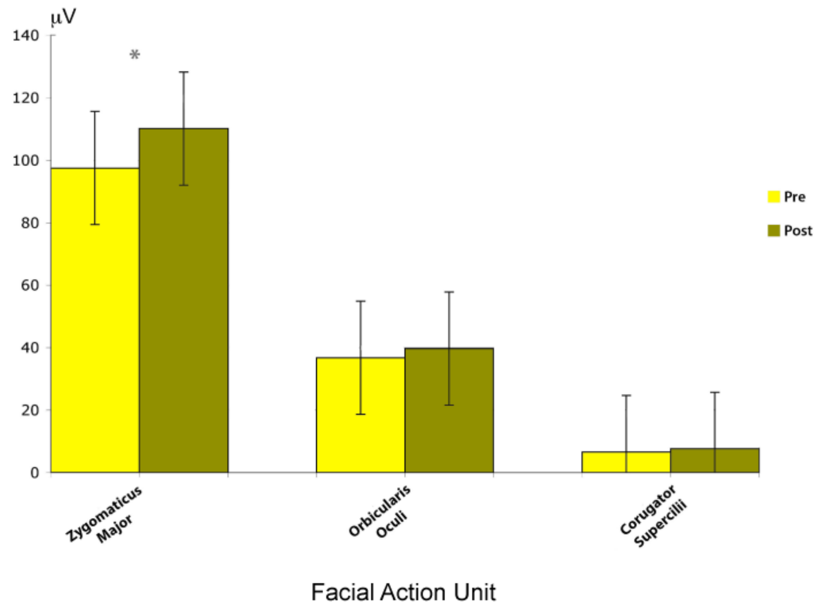

Figure 4. Bar-graph showing levels of activation for the zygomaticus major, orbicularis oculi, and corrugator supercilii activation during the Happy expression, before and after training. Asterisk represents significant difference at p < 0.01.

In the AngryMaze condition, significant differences were found in corrugator supercilii activation, t (24) = -2.70, $p < 0.05$, with greater activation recorded during the post-AngryMaze block (M = 42.75, Se = 4.65) when compared to the pre-AngryMaze block (M = 33.92, Se = 3.95). No significant differences were found for the zygomaticus major, t (24) = -1.60, p = 0.12, or the orbicularis oculi, t (24) = -1.32, p = 0.20, between the pre-AngryMaze and post-AngryMaze measures (see Figure 5).

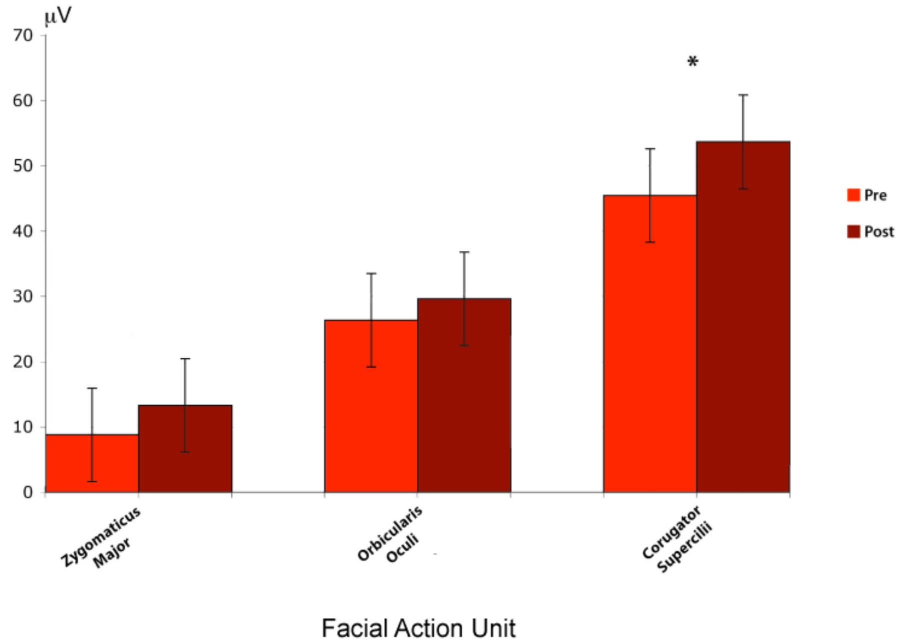

Figure 5. Bar-graph showing levels of activation for the zygomaticus major, orbicularis oculi, and corrugator supercilii activation during the Angry expression, before and after training. Asterisk represents reliable difference at $p < 0.01$.

In contrast, no differences were observed between pre- and post-FaceMaze productions for the Surprise expression as measured by zygomaticus major, orbicularis oculi, or corrugator supercilii activity, $p > 0.10$. Similarly, after playing AngryMaze, zygomaticus major, orbicularis oculi and corrugator supercilii activity was not reliably different from pre-game play levels, $p > 0.10$.

Finally, post-hoc comparisons revealed a significant difference in activation of the zygomaticus major pre-FaceMaze, between the Happy ($M = 73.28$, $Se = 8.87$) and Angry ($M = 7.35$, $Se = 2.14$) expressions, $t(24) = 7.56$, $p < 0.001$, as well as between the Happy and the control Surprise expression ($M = 11.04$, $Se = 2.17$), $t(24) = 7.79$, $p < 0.001$. Activation in the zygomaticus major post-FaceMaze also showed a reliable difference between the Happy expression (*M* = 81.54, *Se* = 8.43) and the Angry (*M* = 10.08, *Se* = 2.42), *t* (24) = 8.90, $p < 0.001$, and control Surprise (*M* = 10.93, *Se* = 1.76) expression, *t* (24) = 9.18, $p < 0.001$, with activation highest in the Happy expression (See Figure 6).

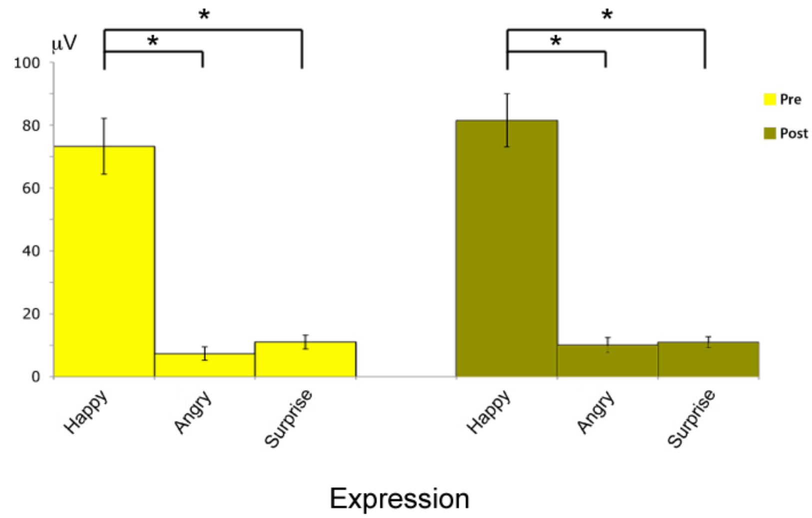

Figure 6. Bar-graph showing levels of activation for the zygomaticus major, during the Happy, Angry and control Surprise expression, before and after training. Asterisk represents reliable difference at p < 0.01

Pre-FaceMaze activation of the corrugator supercilii was significantly larger for the Angry expression (*M* = 33.92, *Se* = 3.95) when compared to the Happy expression (*M* = 5.57, *Se* = 0.64) , *t* (24) = -7.39, $p < 0.00$, and the control Surprise expression (*M* = 18.77, *Se* = 2.44), *t* (24) = 4.22, $p < 0.00$. Furthermore, post-FaceMaze activation of the currugator supercilii was significantly larger for the Angry expression (*M* = 42.75, *Se* =4.65) when compared to the Happy expression (*M* = 5.99, *Se* =0.75), *t* (24) = -8.23, $p < 0.00$, and control Surprise expression (*M* = 20.51, *Se* = 2.87), *t* (24) = 5.46, $p < 0.00$ (see Figure 7).

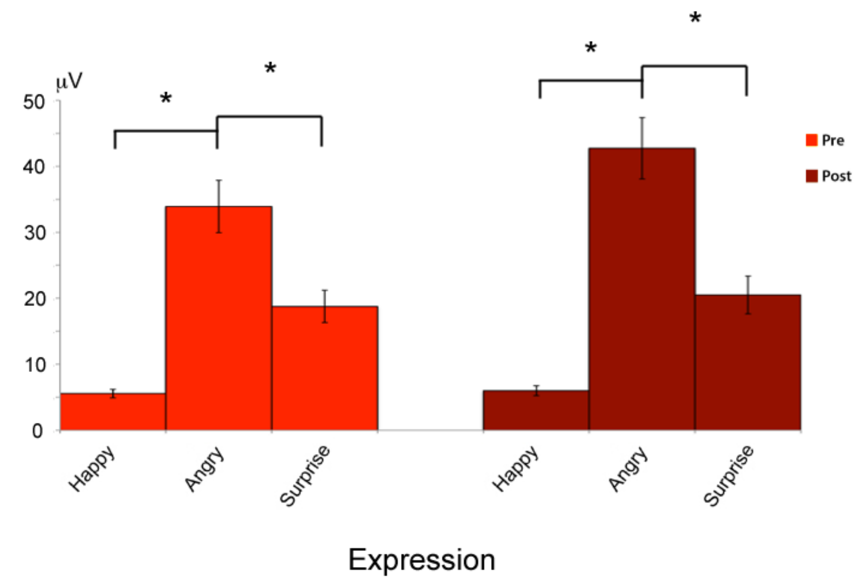

Figure 7. Bar-graph showing levels of activation for the currogator supercilii, during the Happy, Angry and control Surprise expression, before and after training. Asterisk represents reliable difference at $p < 0.01$

In sum, activation of the zygomaticus major (cheek) muscle was larger for the Happy expression, and activation of the currogator supercilii (eyebrow) was larger for the Angry expression pre-training. Furthermore, activation of the zygomaticus major in the HappyMaze condition, and the corrugator supercilii in the AngryMaze condition was significantly larger in the post-training phase when compared to activation in the baseline pre-training phase. These findings indicate that whereas participants were able to activate expression-specific muscles differentially before training, activation of expression-specific muscles was further enhanced by playing the corresponding FaceMaze game.

Discussion

The goal of the current experiment was to provide a physiological check of the CERT module by using traditional EMG methods to detect if muscle activation was enhanced as a result of playing FaceMaze. EMG results revealed that when compared to pre-FaceMaze expressions, facial emotions displayed post-FaceMaze showed enhanced activity in expressions-specific muscles, with greater activity in the zygomaticus major associated with the Happy expression, and enhanced corrugator supercilli activity associated with the Angry expression. Critically, no changes in obicularis occuli activity associated with the Surprise expression were found, underscoring that changes observed in facial expressions post-training did not result from merely activating facial muscles indiscriminately.

It is important to note that pre-Facemaze muscle activation was congruent with that of previous literature showing that spontaneous positive expressions activate the zygomaticus major, whereas spontaneous negative expressions elicit corrugator supercilli activation (Fridlund and Izard, 1983; Dimberg, 1990). Thus, changes in zygomaticus and corrugator activation post-FaceMaze did not result from participants voluntarily producing incorrect facial expressions pre-training, but rather voluntary muscle activation was *enhanced* as a consequence of targeted training. More-so, enhanced activation was limited to the specific muscles associated with a target expression, thus the CERT module was not encouraging more flamboyant expressions (i.e. quantitative change) but encouraging more pointed displays (i.e. qualitative change).

Whereas EMG provides a sensitive, direct measure of muscle activation, one problem in attempting to measure facial muscle activation is that electrode placement may bias participants' activation of certain facial muscles. Furthermore, there is also a lack of ecological validity with respect to expression quality, as facial expressions are not naturally measured via electric impulse, but rather are assessed visually, during social interaction with observers. Whereas the current experiment served a preliminary check to efficacy of CERT as indexed by muscle activation, the goal of the next experiment focuses on verifying the efficacy of the CERT module from the perspective of the observer's judgment of facial expression quality.

Experiment 2

Introduction

Facial expressions are communicative, providing to those around us a signal of our internal state (Buck, 1984; Ekman, 2006; Dimberg, 1983). Despite the relation between facial muscle activity and facial movement, facial expressions are not naturally decoded using measures of electrical impulse but are deciphered visually within the context of human interaction. It is therefore more appropriate to determine the efficacy of FaceMaze in enhancing facial expression production as judged by naïve raters. Previous research has employed the subjective judgments of observers in order to determine expression quality. In one such study (Macdonald, Rutter, Howlin, Rios, Le Conteur, Evered, & Folstein, 1989), participants were asked to produce the corresponding facial expressions in response to vignettes describing an emotional situation. Photographs of participants' productions were taken, and these images were subsequently shown to naïve raters who were required to label the expression. Results show that accuracy ratings differed by expression, with expressions of Happy being correctly categorized 86% of the time, while negative expressions such as Angry were accurately categorized in only 35% of cases (Macdonald et al., 1989). These findings substantiate those of previous EMG studies demonstrating participants' superior abilities in portraying happy expressions.

In another attempt to quantify facial expression production, researchers were interested in determining the effects of sightedness on voluntary productions. Galati, Scherer, and Ricci-Bitti, (1997) compared both blind and sighted participant's abilities in producing voluntary facial expressions by subjecting photographs of their participants' productions to observer judgment, as well as FACS rating (see below). Naïve raters were required to either select or produce a label describing the facial expression seen in each photograph, and findings with respect to sighted participants revealed that only half of all facial expressions voluntarily produced were properly categorized. Specifically, Happy expressions were categorized correctly 83% of the time, while Angry expressions were identified as such in only 33% of the cases. According to the authors, voluntary expression quality was influenced by cultural display rules, thus positive expressions were more easily recognized than negative ones (Galati, Scherer, and Ricci-Bitti, 1997). Furthermore, these findings replicated those of Macdonald et al., (1989), providing support for a disproportionate ability in producing voluntary Happy displays than voluntary negative displays such as Angry.

Other research in facial expression production has relied on more objective, muscle activation coding systems in order to describe facial expression quality. For example, the Facial Action Coding System (FACS) (Ekman and Frisien, 1978) is an anatomically based coding system that allows for the description of facial muscles, or Facial Action Units (AUs), at discrete levels of activation. Research using FACS has been able to determine specific patterns of activation involved in facial expression production (Ekman and Frisien, 1978; Ekman, Frisien & Hager, 2002). Furthermore, the Maximally Discriminative Affect Coding System (MAX) (Izard, 1979; Izard 1983) is also an anatomically based coding system that describes facial movements in three separate regions of the face (brows and forehead, eyes and cheeks, mouth), and determines the production of a facial expression based on a constellation of specific movements in each of these three regions (Izard, 1979; Izard 1983). Whereas FACS describes the facial expression quality with respect to the appropriateness of muscle movement, MAX categorizes facial expressions with respect to muscle movement. As a result, FACS coding has been used to determine a single expression's quality, whereas MAX coding has been used to describe the kinds of pure or blended expressions seen.

Research using FACS has been used to quantify voluntary facial expression production in adults. As previously mentioned, Galati, Scherer, and Ricci-Bitti, (1997), not only compared blind and sighted participants' abilities in producing voluntary facial expressions using observer judgment, but also subjected photographs of their participants' productions to FACS assessment. In the study, blind and sighted participants were given short vignettes that described a situation in which an emotion was elicited, and participants were required to produce the corresponding facial expression. Photos of the participants taken during production were then subjected to FACS coding, which involved cataloguing the observable muscle activation in each photograph and comparing that to the FACS verified expression activation codes. Results of the FACS analysis revealed that both normal and blind participants failed to activate all the appropriate AUs associated with any specific emotion, with the exception of Happy. Happy expressions elicited activation in AU 12 (zygomaticus major) in 100% of sighted cases, and in 86% of blind participants. Elicitation of other expression-specific AUs occurred in less than half of all participants, such as in the case of AU 4 (brow lowerer; depressor glabellae, depressor supercilii, corrugator supercilii), which was activated in expressions of Anger in only 21% of sighted individuals, and 6% of blind participants (Galati, Scherer, and Ricci-Bitti, 1997). These findings thus provide objectively measurable support for the superiority of voluntary Happy facial displays, and the poor expression quality characteristic of voluntary Angry facial expressions.

In another study examining the coordination of facial AUs in typical adults (Gosselin, Beaupré and Perron, 2010), participants were required to activate individual AUs after receiving written descriptions and video demonstrations of the AU movement, and practicing the AU movements with feedback from both a mirror and the researchers. The main finding showed that whereas adults were adept at activating AUs involved in Happy expressions both in isolation and in combination with other AUs, they were less adept at activating AUs involved in expressions of Anger, Disgust or Sad, in isolation or in combination with other AUs (Gosselin, Beaupré and Perron, 2010).

Similar results have also been obtained using the MAX coding system. Lewis, Sullivan and Vasen (1987) compared adult and children's performance in producing voluntarily expressions of happy, angry, surprise, fear, disgust and sad. In this study, the participants' expressions were video-recorded and scored by two independent raters using the MAX (Izard, 1979). Expressions were scored as "complete" when all three correct facial muscle components were activated, "partial" when two or fewer correct facial muscle components were present, or "incorrect" if all facial muscle components activated were inappropriate for the expression requested. Results of the coding revealed that whereas adults were able to produce more complete facial expressions, children under the age of 4 could only produce partial expressions, and children between the ages of 4 years and 10 years produced a mixture of both complete and partial facial expressions. With respect to type of expression, positive expressions (happy, surprise) were rated as complete more often than negative expressions (angry, sad, fear, disgust), and this trend was observable even into adulthood wherein only the positive expressions were consistently produced this way (Lewis, Sullivan and Vasen, 1987).

It is important to emphasize that the above research alludes to only one type of voluntary expression produced without an external model, namely posed expressions. Studies in another type of voluntary expression, mimicry, have examined facial productions in which participants are provided with a human model or photograph of the target expression to imitate. Not surprisingly, the quality of the expression was enhanced when an external example is provided (Dimberg 1982; McIntosh 2006), and subsequent expression training paradigms, such as the FACS, have made use of mimicry in training facial expressions (Ekman & Friesen, 1978). This line of research thus implies that the discrepancy in expression production is not related to an inability to activate facial muscles, but is associated with a deficit in an expressions' internal representation. Thus, expressions can also be improved when participants are provided with an external representation to model their productions. From a theoretic standpoint, however, providing an external representation is a more indirect method to increasing expression fidelity as mimicry requires first a visual representation, and then matching between the actor and mimic before proprioceptive mechanisms may come into play. Posed expressions, however, are more direct in that they rely strictly on proprioceptive mechanisms that are generated from an internal representation of the emotion.

In sum, previous research has shown that typically developing (TD) adults can, within limits, produce facial expressions that are consistent with both subjective and objective interpretations of an external observer. Specifically, positive displays such as Happy are successfully decoded, while negative displays such as Angry are not as efficiently interpreted. The goal of the current chapter is to assess the efficacy of FaceMaze in enhancing the perceptibility of facial expressions by altering the expression concept using mimicry. The methods and procedures used in the current experiment were similar to those of the previous experiment, except that no EMG measures were taken. Instead, facial expression production was assessed using observer ratings. Furthermore, facial expressions entrained by FaceMaze were compared to those entrained by a previously validated facial expression-training paradigm, namely the FACS. In the first part of this experiment, participants were assigned to either a Facemaze or FACS training group, and videos of their happy and angry expressions were recorded before and after training. In the second part of the experiment, naïve participants were asked to rate the videos of the Facemaze and FAC groups for expression quality. If FaceMaze selectively enhances facial expression production, then target expression ratings should increase after game-play, with ratings of “happy” increasing after playing “HappyMaze” and ratings of “angry” increasing after playing “AngryMaze”. Critically, no changes in ratings of “surprise” for the control expression of Surprise should be observed following game-play. Alternatively, if training has no effect on expression production, we would expect to see no changes in expression quality ratings after playing FaceMaze.

Furthermore, in order to quantify the efficiency of targeting the expression concept directly through proprioceptive mechanisms, the results of FaceMaze were compared to the instructional and mimicry approach of the FACS (Ekman & Friesen, 1978), with the supposition that directly altering the expression concept will result in more identifiable expressions, as determined by larger increases in expression quality ratings for the FaceMaze condition when compared to FACS.

Method

Part 1 – Stimulus Generation

Participants

Four participants (2 male) comprised the Facemaze group, mean ages 19 to 21 (M = 20.2). Four participants (2 male) comprising the FACS group, mean ages 19 to 21 (M = 19.5) All students were from the University of Victoria and were compensated with course credit for their time.

Materials

Video recordings of frontal facial expressions produced by the participants (see procedure) were recorded using a Canon Powershot i-780, mounted above the computer monitor that displayed the expression cues.

Procedure

Consent to the use of video recordings was obtained both before and after the experiment from all participants, with video recordings of the participants' expressions obtained before and after training. The FaceMaze group played the FaceMaze game during the training period as described in the previous experiment. The FACS group underwent a modified FACS training procedure in which participants were first shown the separate muscle groups involved in making either the Happy or Angry facial expressions, with emphasis on the orbis obiccularis and zygomaticus major for the Happy expression, and the corrugator supercilli for the Angry expression. The experimenter explained the movement of the corresponding muscle for each expression, and then demonstrated the facial expression. Participants were encouraged to mimic the experimenter in moving the corresponding muscle groups for the happy and angry expressions but were not provided any feedback. Following this, participants were oriented to the computer screen and were given a practice trial in which a fixation cross appeared for 2 seconds, followed by an image of a FACS-verified exemplar producing the target expression that they were to mimic. The training session consisted of showing 24 FACS-verified exemplars on a computer screen, and participants were told to mimic the facial expression they saw. The exemplar images featured an individual producing the corresponding Happy or Angry facial expression corresponding to block condition, with arrows pointing to the corresponding muscle groups implicated in production. Blocks were counterbalanced across participants. For the Happy exemplars, the arrows pointed to the orbis obiccularis and zygomaticus major. For the Angry exemplars, the arrows pointed to the corrugator supercillii and the buccinator (see Figure 8).

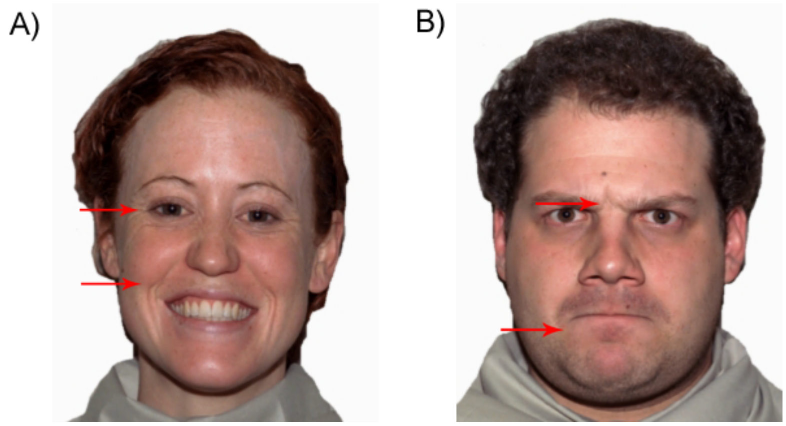

Figure 8. Examples of stimuli used in the FACS condition, depicting A) happy facial expression and b) angry facial expression. Arrows point to the corresponding FACS verified AUs.

Before and after FACS training, participants were instructed to produce the happy, angry and surprise facial expression a total of 15 times during each assessment block. The facial expression selected for the stimulus set was the last production of "Happy" "Angry" and "Surprise" during the pre- and post-training assessment. The video recordings were edited into 2.7-second clips that captured one facial expression. Thus, each clip showed the participant's expression moving through first neutral, then emotive (i.e. Happy, Angry or Surprised), and then neutral expressions. Four critical (i.e., pre-training Happy and Angry, post-training Happy and Angry) and two control (i.e., pre-training and post-training Surprise) video clips were obtained from each participant. In a total of 48 video clips were used for the expression rating phase of the experiment, 24 clips from the FaceMaze group, and 24 clips from the FACS group.

Part 2 – Expression Rating

Participants

23 naïve undergraduate participants (5 male), ages 18 to 32 (M = 21.22) from the University of Victoria took part in this portion of the experiment. All participants had normal or corrected-to-normal vision. Participants received course credit as compensation for their time.

Materials

Rating scales consisted of emotion labels with a Likert-scale ascending from 0 (not at all) to 4 (very much). All six basic emotions of happy, angry, surprise, fear, disgust, and sad were presented for each video, with each emotion label corresponding to one rating scale. A total of 48 rating scales were given to participants, to be filled out manually.

Stimuli

The 48 video clips of the happy, angry and surprise expressions were presented on a computer screen with viewers sitting 1 meter away, resulting in an image of 16.51 x 10.16 centimetre on a white screen, creating a visual angle of 44.7 degrees in the horizontal plane and 27.64 in the vertical plane.

Procedure

After obtaining their consent, participants were seated in front of the computer. Participants were told that they would be viewing a series of video clips and were required to rate each video clip on a series of scales. Participants were then given the rating scale form and instructed to rate the video for the quality of happy, angry, surprise, fear, disgust, and sad on a scale of "0" (not at all) to "4" (very much) Following this, a practice trial was given in which participants were to rate one video clip. If there were no further questions, participants proceeded to the experiment phase.

The experiment phase consisted of 48 video clips, with presentation of one video-clip constituting a trial. The video clips were divided into two blocks of 24 videos. Each block contained 12 videos from the FaceMaze-trained participants (6 pre-training and 6 post-training videos) and 12 videos from the FACS-trained participants (6 pre-training and 6 post-training videos). Presentation of the blocks was counterbalanced across participants and the order of the clips within each block was randomized.

At the beginning of each trial, participants saw a screen reading "get ready…" for 2 seconds before the video clip was shown. Immediately after presentation of the video clip, a screen reading "Please rate the video now. When you are finished, please press "spacebar" to continue…" appeared. Participants then could fill out the ratings, and proceed to the next video at their own pace. After the presentation of the video clips was complete, participants were debriefed and thanked for their time.

Results

Separate repeated measures ANOVA's were performed for the Happy, Angry, and Surprise expressions, with 2 (group; FaceMaze, FACS) x 2 (time; pre, post) x 6 (emotion; happy, angry, surprise, fear, disgust, sad) as within-subjects factors. All within-subjects factors were Greenhouse-Geisser corrected, and Bonferonni adjustments were performed.

Happy Training Condition. A main effect of Group, *F* (1, 91) = 12.45, p < 0.005, η_p^2 = 0.12, was observed where the FaceMaze group (*M* = 1.69, *Se* = 0.03) produced reliably larger ratings than the FACS group (*M* = 1.61, *Se* = 0.02). A main effect of Time, *F* (1, 91) = 11.44, *p* < 0.005, η_p^2 = 0.11, was also found, such that expression quality ratings of post-training happy expressions (*M* = 0.68, *Se* = 0.02) were reliably greater than those of pre-training productions (*M* = 0.62, *Se* = 0.02) A main effect of Emotion was also found, *F* (2.47, 225.00) = 913.93, *p* < 0.001, η_p^2 = 0.91, with expression quality ratings of happy (*M* = 3.00, *Se* = 0.06) larger than angry (*M* = 0.14, *Se* = 0.03), surprise, (*M* = 0.39, *Se* = 0.05), fear (*M* = 0.10, *Se* = 1.02), disgust (*M* = 0.09, *Se* =0.03), and sad (*M* = 0.17, *Se* = 0.03). Surprise expression quality ratings were also reliably larger than those of Angry, Fear, Sad and Disgust. In order of magnitude, expression quality ratings of happy were the largest, followed by surprise, and all other ratings lesser than surprise showing no difference from each other. The Group x Emotion interaction was reliable, *F* (2.39, 216.88) = 5.06, *p* = 0.005, η_p^2 = 0.05 as well as a Time x Emotion interaction was observed, F (3.28, 298.34) = 12.24, p < 0.001, η_p^2 = 0.12. No reliable interaction of Group x Time x Emotion, F (2.55, 232.22) = 2.06, p = 0.12, η_p^2 = 0.02, was observed.

Collapsing across groups, post-hoc analysis revealed a significant increase in expression quality ratings of happy post-training (*M* = 3.21, *Se* = 0.07) when compared to pre-training expression quality ratings (*M* = 2.79, *Se* = 0.08), *t* (183) = -4.87, *p* < 0.001. Furthermore, a reliable decrease in post-training expression quality ratings of sad (*M* =

0.11, *Se* = 0.03) when compared to pre-training expression quality ratings, (*M* = 0.22, *Se* = 0.04) was also found, *t* (183) = 2.56, *p* = 0.01 (see Figure 9).

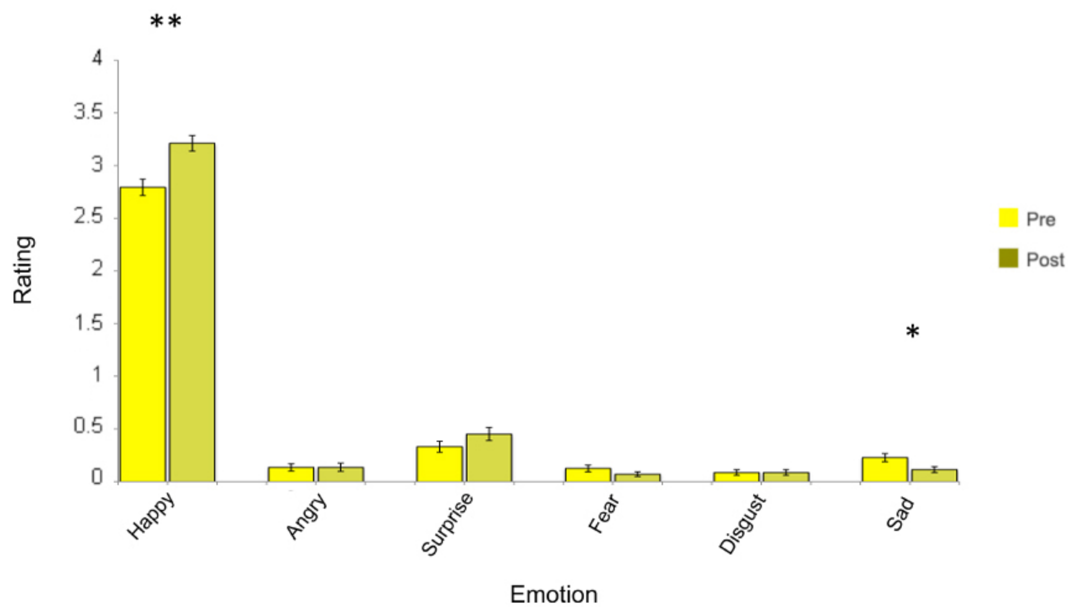

Figure 9. Bar-graph of expression quality ratings for the Happy expression before and after training, collapsed across the Facemaze and FACS groups. Error bars represent Standard Error of the mean. Asterisk represents significant difference at p < 0.05, double asterisk represents significant difference at p < 0.005.

Angry Training Condition. For the Angry condition, a reliable main effect of Emotion, *F* (3.31, 301.00) = 117.87, p < 0.001, η_p^2 = 0.56, was found, with expression quality ratings of angry (*M* = 2.10, *Se* = 0.10) significantly larger than all other quality ratings of happy (*M* = 0.25, *Se* = 0.04), surprise (*M* = 0.56, *Se* = 0.06), fear (*M* = 0.27, *Se* = 0.04), disgust (*M* = 1.58, *Se* = 0.09) and sad (*M* = 0.91, *Se* = 0.07). Furthermore, quality ratings of disgust were also reliably larger than ratings of happy, surprise, fear and sad, and ratings of sad significantly larger than ratings of happy, surprise, and fear. In order of magnitude, expression quality ratings of angry were the largest, followed by disgust, then sad, then all other expression ratings. There was no reliable main effect of Group, F (1, 91) = 2.03, $p = 0.16$, $\eta_p^2 = 0.02$, or Time, F (1, 91) = 0.88, $p = 0.35$, $\eta_p^2 = 0.01$, however there was a reliable interaction between Group x Emotion, F (3.70, 337.07) = 22.37, $p < 0.001$, $\eta_p^2 = 0.20$, and a significant interaction between Time x Emotion, F (3.90, 355.28) = 7.69, $p < 0.001$, ηp2 = 0.08. Finally, a reliable Group x Time x Emotion interaction was found, F (3.57, 324.38) = 4.98, $p < 0.005$, $\eta_p^2 = 0.05$.

Post-hoc analysis focusing on the FaceMaze group revealed a significant increase in quality ratings of angry post-AngryMaze (M = 2.63, Se = 0.12) when compared to pre-AngryMaze ratings (M = 2.07, Se = 0.15), t (91) = -3.56, $p < 0.005$. Furthermore, ratings of happy significantly decreased post-AngryMaze (M = 0.01, Se = 0.01) as compared with pre-AngryMaze ratings (M = 0.73, Se = 0.13), t (91) = 5.46, $p < 0.001$ (see Figure 10).

As can be seen in Figure 8, for the FACS group, quality ratings of fear significantly decreased post-training (M = 0.18, Se = 0.05) as compared with pre-training ratings (M = 0.37, Se = 0.08), t (91) = 2.19, $p < 0.05$, and a reliable increase in quality ratings of sad post-training (M = 1.51, Se =0.15) when compared to pre-training ratings (M = 1.14, Se =0.13) was observed, t (91) = -2.09, $p < 0.05$.

Finally, no reliable difference in pre-training quality ratings of angry between the FaceMaze and the FACS groups was found, t (91) = 1.34, $p = 0.18$, however a reliable difference was observed for post-training quality ratings between the two groups, t (91) = 5.20, $p < 0.001$. Specifically, post-training quality ratings of angry were higher for the

FaceMaze group ($M = 2.63$, $Se = 0.13$) than those of the FACS group ($M = 1.83$, Se = 0.14) (see Figure 11).

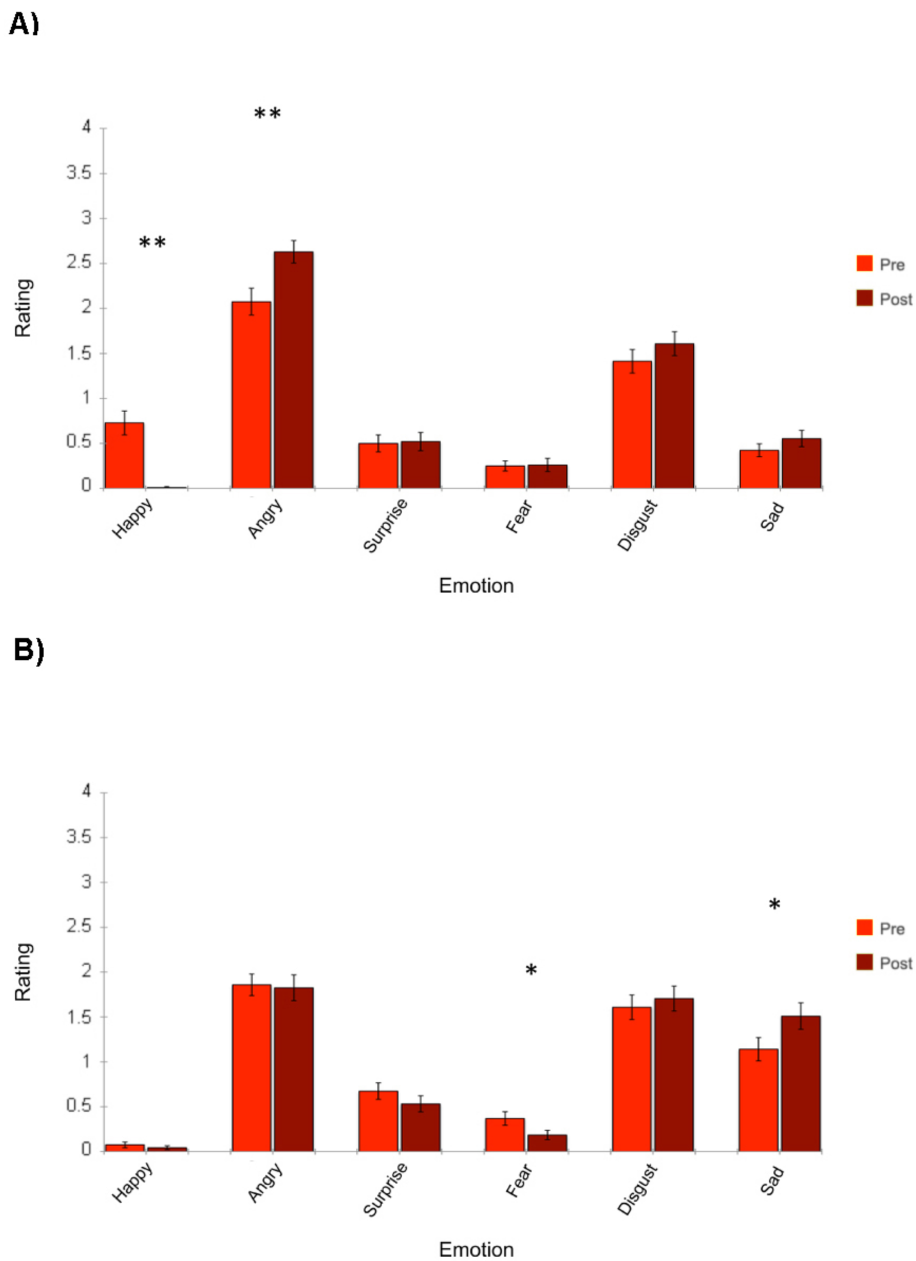

Figure 10. Bar-graph of expression quality ratings for the Angry expression before and after training for a) the Facemaze condition, and b) the FACS condition. Error bars represent Standard Error of the mean. Asterisk represents significant difference at $p < 0.05$.

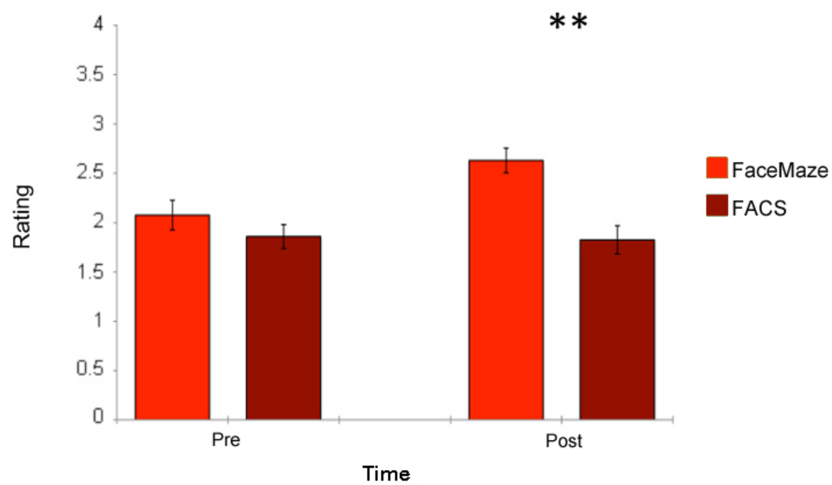

Figure 11. Bar-graph of the angry expression quality rating for both FaceMaze and FACS groups in the Angry condition, before and after training. Error bars represent Standard Error of the mean. Asterisk represents a significant difference at $p < 0.05$.

Surprise. With regards to the Surprise condition, a main effect of Emotion, F (2.99, 269.93) = 411.44, $p < 0.00$, $\eta_p^2 = 0.82$, was found, such that expression quality ratings of surprise ($M = 3.09$, $Se = .08$) were reliably larger than all other ratings of happy ($M = 0.85$, $Se = .07$) angry ($M = 0.15$, $Se = .03$), fear ($M = 0.63$, $Se = .07$), disgust ($M = 0.35$, $Se = .05$) and sad ($M = 0.11$, $Se = .02$). Furthermore, expression quality ratings of happy were also reliably larger than those of angry, disgust and sad, and ratings of sad were reliably lower than all other ratings, except for angry. In order of magnitude, expression quality ratings of surprise were the largest, followed by happy, fear, disgust, and ending with ratings of angry and sad. No main effect of Group, F (1, 90) = 0.27, $p = 0.63$, $\eta_p^2 < 0.00$, or Time, F (1, 90) = .05, $p = 0.83$, $\eta_p^2 < 0.00$, was found. Finally, an interaction effect of a Group x Time x Rating was also observed, F (3.84, 345.20) = 5.54, $p < 0.00$, $\eta_p^2 = 0.06$.

Subsequent post-hoc analysis revealed no difference in pre-training expression quality ratings of surprise between the FaceMaze (M = 3.16, Se = 0.10) and FACS groups (M = 3.12, Se = 0.11), t (91) = 1.08, p = 0.28, nor any difference between the FaceMaze (M = 3.03, Se = 0.12) and FACS groups (M = 3.12, Se = 0.11) post-training, t (91) = -0.66, p = 0.51.

To summarize, no differences between the FaceMaze and FACS training groups was found for the Happy expression. Naïve observers rated the post-training productions of the Happy expression as higher in quality than pre-training productions. Furthermore, a decrease in expression quality ratings of sad was also observed for post-training productions compared to pre-training productions. Differences between the FaceMaze and FACS conditions were observed for the Angry expressions. For the FaceMaze condition, naïve observers rated post-FaceMaze productions of the Angry expression as higher in quality than pre-FaceMaze productions. Furthermore, post-FaceMaze productions showed a decrease in expression quality ratings of happy as well, indicating that the expressions were becoming more differentiated. Post-training productions of Angry for the FACs group, however, were rated higher on sad, and lower on fear than pre-FACS productions, and no changes in the target angry expression quality ratings were observed. No differences between the FaceMaze and FACs group was also observed for pre-training ratings of the Angry expression, however post-training productions were rated higher in expression quality for the FaceMaze group when compared to the FACS group. Critically, no changes in the quality of expressions were found between pre- and post-training productions of the FaceMaze or FACs trained participants for the control expression of Surprise.

Collectively, these findings show that the FaceMaze game increases the perceptibility and fidelity of the targeted Happy and Angry expressions as well as, and more efficiently than FACS training.

Discussion

The goal of the current experiment was to first test the efficacy of FaceMaze in changing facial expression production using the more ecologically valid approach of observer judgment, and second to compare post-training changes to those of the more traditional FACS training paradigm. Findings show that naïve raters not only judged post-FaceMaze expressions of “happy” and “angry” as higher in quality than pre-game expressions, but also demonstrated more gains in expression quality when compared to ratings of the more traditional FACS training. These results provide evidence for the utility of FaceMaze in enhancing facial expression production.

Naïve raters judged post-FaceMaze expressions of “happy” and “angry” as higher in quality than pre-game expressions, however no changes in observers’ ratings for the control Surprise expression were found post-game-play. This finding underscores that the differences found in productions of Happy and Angry expressions after playing FaceMaze were due to directed training in the production of specific expression and was not a by-product of simple practice in making generalized facial movements. Observer judgments showing an increase in ratings of disgust for Angry displays are also not unusual given previous research showing that angry expressions have been categorized as expressions of disgust, and vice-versa, by naïve raters (Aviezer, Hassin, Ryan, Grady, Susskind, Anderson, Moscovitch, & Bentin, 2008; Bullock & Russel, 1984; Widen &

Russell, 2010). This result thus suggests that posed angry expressions also looked more genuine after FaceMaze training. Lastly, ratings of happy were also eliminated for the "Angry" expression after game-play, suggesting that the target expressions were also being differentiated conceptually after training. Together, these results validate the value and utility of both the CERT system and the FaceMaze game in training facial expression production.

Naïve raters' judgment of FACS trained happy expressions also showed an increase in expression quality, however this was not the case for FACS trained angry expressions that showed no enhancement after training. This finding is not necessarily surprising given the results of previous research showing a large disparity between the voluntary production of positive (happy) and negative (i.e. angry) facial displays across the developmental spectrum, with a bias toward more identifiable positive expressions (Lewis, Sullivan and Vasen, 1987). As a result of this discrepancy in baseline performance, some facial displays may be easier to enhance through training than others. It is important to note, however, that training using the FaceMaze module was able to enhance both Happy and Angry facial expressions. Hence, the two training techniques impacted the production quality of the happy expression. FACS training targets facial movement by providing explicit coaching in the form of verbal instruction, and visual aids such as showing participants pictures of facial expressions, modeling AU movement and combinations, and verbal feedback to encourage correct facial movement and matching. Consequently, participants in the FACS condition that relied on exogenous cues and mimicry to produce facial expressions may not have internalized changes in facial movement.

In contrast, FaceMaze does not provide the extensive and explicit direction, relying instead on endogenous cues such as proprioceptive information generated from the entire group of facial muscles and not individual AU activation. As a result, participants in the FaceMaze condition were forced to rely on their expression concept when generating a facial display, and any alterations made to elicit feedback could have then been immediately incorporated into this internal representation. To be clear, changes in the expression concept can also occur during FACS training as evidenced by other research showing that FACS-based training has been shown to enhance facial expression production (Charlop, Dennis, Carpenter, and Greenberg, 2010; DeQuiznio, Townsend, Sturmey & Poulson, 2007; Gena, Krantz, McClannahan, & Poulson, 1996; Stewart and Singh, 1995). From a training perspective, however, these findings point to the efficacy of the Facemaze as a training paradigm, in that the FaceMaze was able to produce enhancements of the angry expression in the allotted, albeit short game-play time whereas no improved performance was observed for the same amount of traditional FACS training. It is also important to underscore that changes in facial expression production were measured across only four participants for each training condition. Thus, future research may look to increasing the stimulus set to include more models to verify the extent to which FACS and FaceMaze training differ.

Findings from this experiment provide further evidence for the efficacy of FaceMaze as a training paradigm by extending findings to include observer ratings. Chapter 3 (in press) further investigates the training potential of both the CERT module and FaceMaze game by testing the training paradigm's potential with younger, and clinical populations.

Chapter 3

Experiment 3

Introduction

Autism Spectrum Disorder (ASD) is a pervasive developmental disorder typified by deficits in social communication and restricted or repetitive behaviours (American Psychological Association, 2013). One deficit in social communication is flat (lack of) or disorganized (ambiguous) facial expression production (Lord, Risi, Lambrecht, Cook, Leventhal, DiLavore, Pickles, & Rutter, 2000). Although Kanner (1945, 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, Rutter, Howlin, Rios, Le Conteur, Evered, and Folstein (1989) assessed high-functioning ASD adults' abilities to recognize and produce 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 mislabelled the ASD expressions more frequently than the TD expressions. A higher proportion of labelling 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, Tunali-Kotoski, Pearson, Brelsford, Ortegon, and Chen (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 displays 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 labelled 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 their 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, Reichmann-Decker, Winkielman & Wilbarger, 2006; Rogers, Hepburn, Stackhouse & Wehner, 2003; Williams Whiten, Suddendorf & Perrett, 2001), however research has shown that these individuals retain their ability to mimic others when explicitly prompted (McIntosh, Reichmann-Decker, Winkielman & Wilbarger, 2006; Rogers, Hepburn, Stackhouse & Wehner, 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 individuals with ASD perform on par with TD children when they are required to mimic a facial expression in the presence 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 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. 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 (2 female), aged 6 to 18 years (M = 10.89, SD = 3.39), were recruited from the Centre for Autism Research Training and Education (CARTE) database. Parents of participating children verified children’s diagnosis by providing documentation of assessment from 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 to 18 years (*M* = 10.76, *SD* = 3.59). 23 typically developing (TD) aged 8 to 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.

Table 1. Average K-Bit-2 scores for ASD and TD participants. Parentheses denote standard errors.

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

Materials

Kaufman Brief Intelligence Test (2nd edition). The Kaufman Brief Intelligence Test (2nd edition) (Kbit-2) yields both a verbal (crystallized) and non-verbal (fluid) intelligence score. The verbal portion of the Kbit-2 collapses across two separate tests of verbal knowledge which test for comprehension (definitions) and production (riddles). The non-verbal portion assesses pattern recognition using pictorial matrices.

The Computer Expression Recognition Toolbox (CERT). To implement our training program, we employed the Computer Emotion Recognition Toolbox (CERT) developed by Bartlett and colleagues (Littlewort, Whitehill, Wu, Frank, Movellan, Bartlett, 2011; Bartlett, Littlewort, Frank, Lainscsek, & Movellan, 2005, 2006). 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, Bartlett, Hager, Ekman, & Sejnowski, 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) according to 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. 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, Bartlett, Tanaka, Movellan, & Schultz, 2008) (see Figure 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.

The CERT program interprets "happy" as the upward inflection of the lip produced by zygomaticus major facial muscle. The CERT operationalizes "angry" as the tensing of the 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, 8 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 Kbit-2.

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 three seconds 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 playing the happy version of FaceMaze (and before playing the angry version of FaceMaze and 3) after playing the angry version of FaceMaze (see Figure 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. Of the 11 excluded participants, 4 were removed because of extraneous hand gestures that interfered with the image of a face, another 6 were removed because participants had moved out of the camera focus, and 1 was removed as a result of producing facial expressions toward the experimenter and no the camera. Participants must have had all 6 videos in order to be included in the stimulus set, thus 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 minutes 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 to 22 (M = 20.04, SD = 1.19) rated the videos obtained from the ASD children, and 24 participants (6 male), aged 18 to 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 meter away from participants, resulting in a retinal image of 16.51 x 10.16 cm on a white screen, creating a visual angle of 44.7 degrees 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 8 participants (48 videos) from the 3-level condition, 4 participants (24 videos) from the 2-level condition, and 5 participants (30 videos) from the 1-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 start of the experiment. 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 seconds, 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.

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 x post) x 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, 1108.84) = 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 x Emotion interaction was observed, *F* (3.55, 1445.67) = 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 .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 Figure 12).

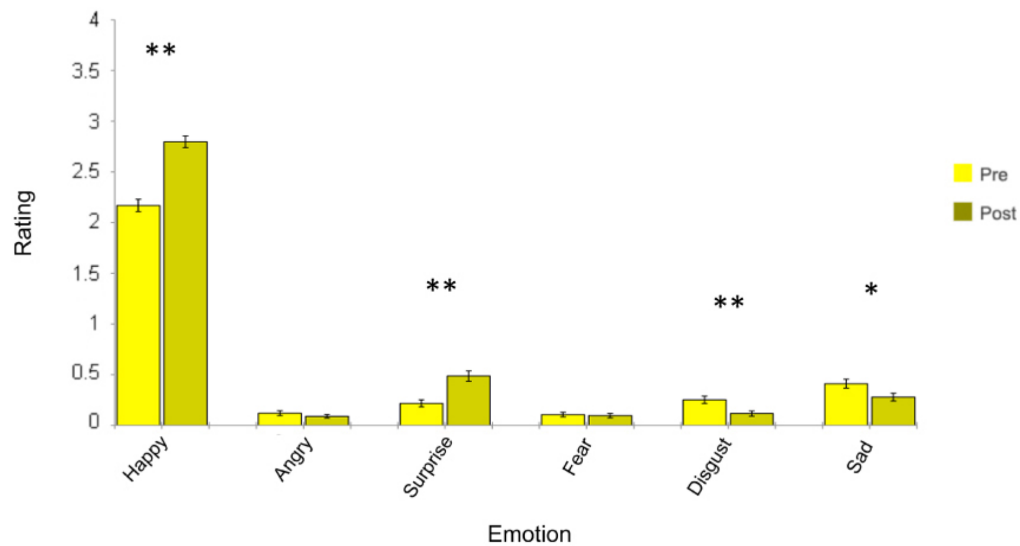

Figure 12. Bar-graph of expression quality ratings for the Happy expression, before and after training. Error bars denote Standard Error of the mean. Asterisk represents a significant difference at p < 0.05. Asterisk represents a significant difference at p < 0.05. Double asterisk represents significance at p < 0.01.

AngryMaze. In order to determine the effects of playing AngryMaze on facial expression production, expression quality ratings were subjected to a 2 (time: pre x post) x 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, η_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, 1442.04) = 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 x Emotion interaction was found, F (3.72, 1514.17) = 14.67, $p < 0.001$, $\eta_p^2 = 0.04$.

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 Figure 13).

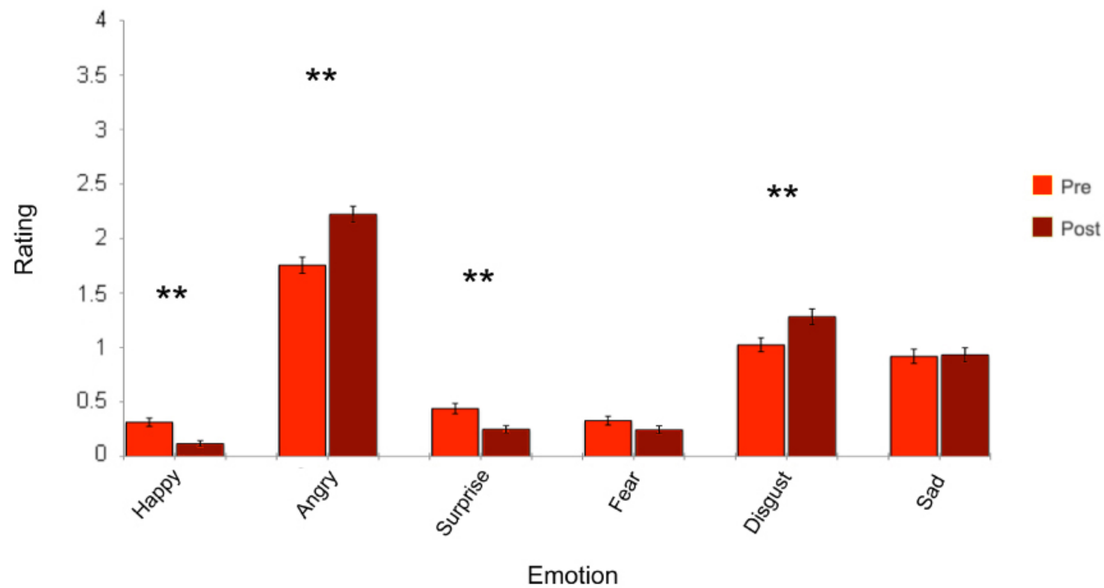

Figure 13. Bar-graph of expression quality ratings for the Angry condition, before and after training. Error bars denote Standard Error of the mean. Asterisk represents a significant difference at $p < 0.05$. Double asterisk represents significance at $p < 0.01$.

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 1 level of the AngryMaze , 2 levels of AngryMaze, or 3 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 Figure 14). Importantly, no differences were found between pre- and post-training productions for the target expression of angry.

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 Figure 14).

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 Figure 14).

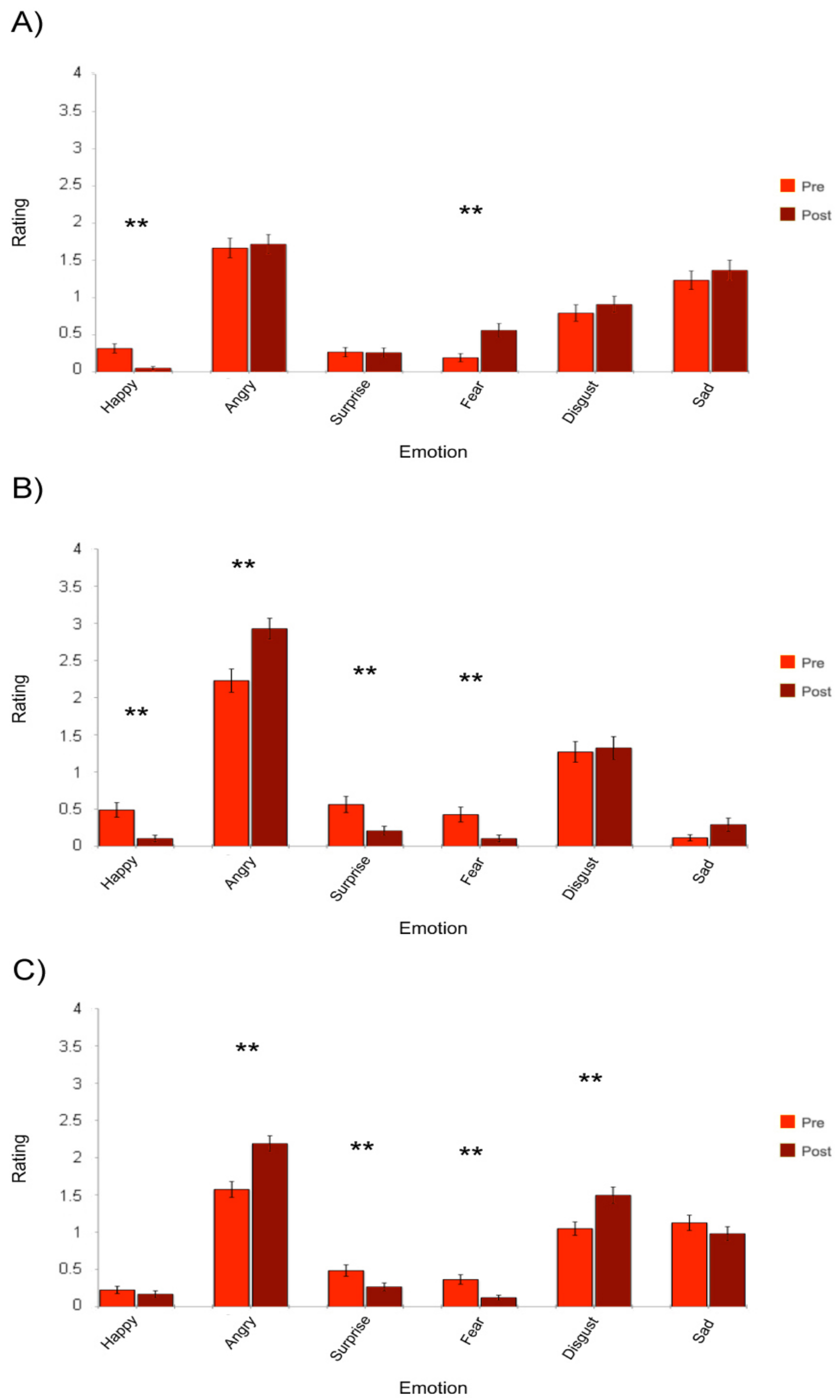

Figure 14. Bar-graph of expression quality ratings for the a) 1-level, b) 2-level, and c) 3-level Angry condition, before and after training. Error bars denote Standard Error of the mean. Asterisk represents a significant difference at $p < 0.05$. Double asterisk represents significance at $p < 0.01$.

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 x post) x 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, 1335.92) = 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 x Emotion, F (4.40, 1790.12) = 3.87, $p < 0.005$, $\eta_p^2 = 0.01$, was found (see Figure 15).

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, $p < 0.005$.

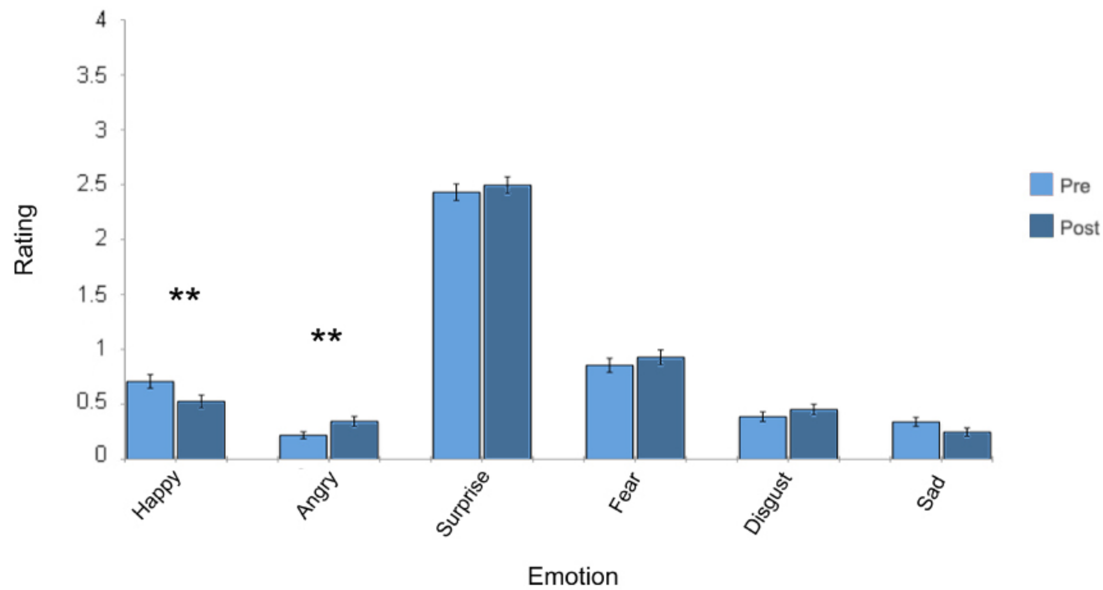

Figure 15. Bar-graph of expression quality ratings for the Surprise condition before and after training. Error bars denote Standard Error of the mean. Asterisk represents a significant difference at $p < 0.05$. Double asterisk represents significance at $p < 0.01$.

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 x post) x 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) = 2077.74, $p < 0.001$, $\eta_p^2 = 0.84$, 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 x Rating interaction, *F* (2.64, 1075.80) = 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 Figure 16).

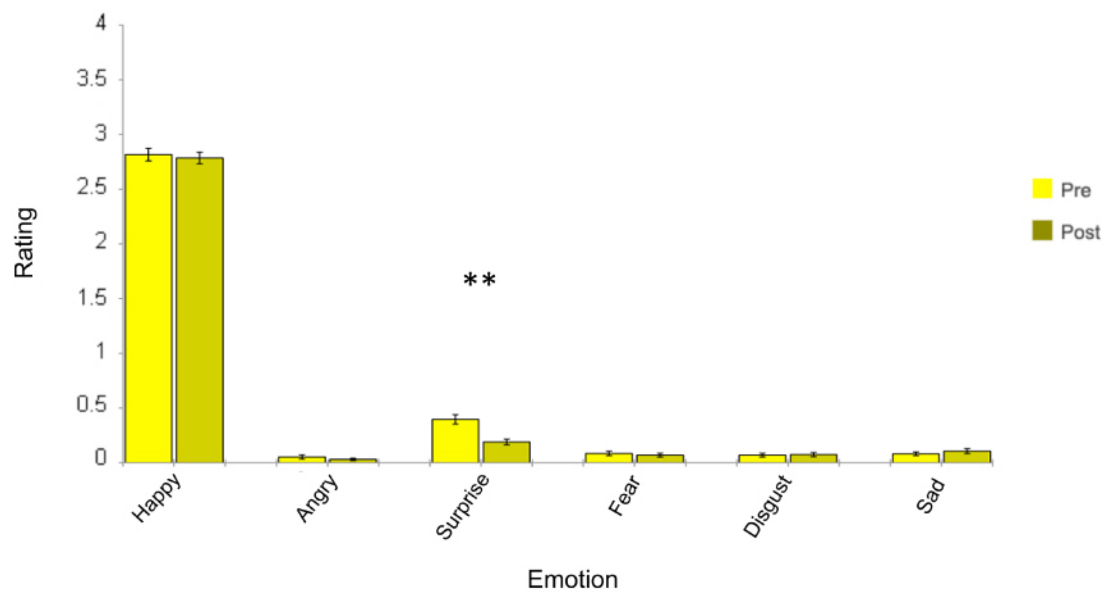

Figure 16. Bar-graph of expression quality ratings for the Happy expression, before and after training. Error bars denote Standard Error of the mean. Asterisk represents a significant difference at p < 0.05. Double asterisk represents significance at p < 0.01.

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 x post) x 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, 1356.69) = 241.41, $p < 0.001$, $\eta_p^2 = 0.37$, was found, 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 was observed, F (1, 407) =, $p = 1.56$, $\eta_p^2 = 0.21$, however a significant Time x Rating interaction was found, F (3.77, 1532.56) = 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 Figure 17).

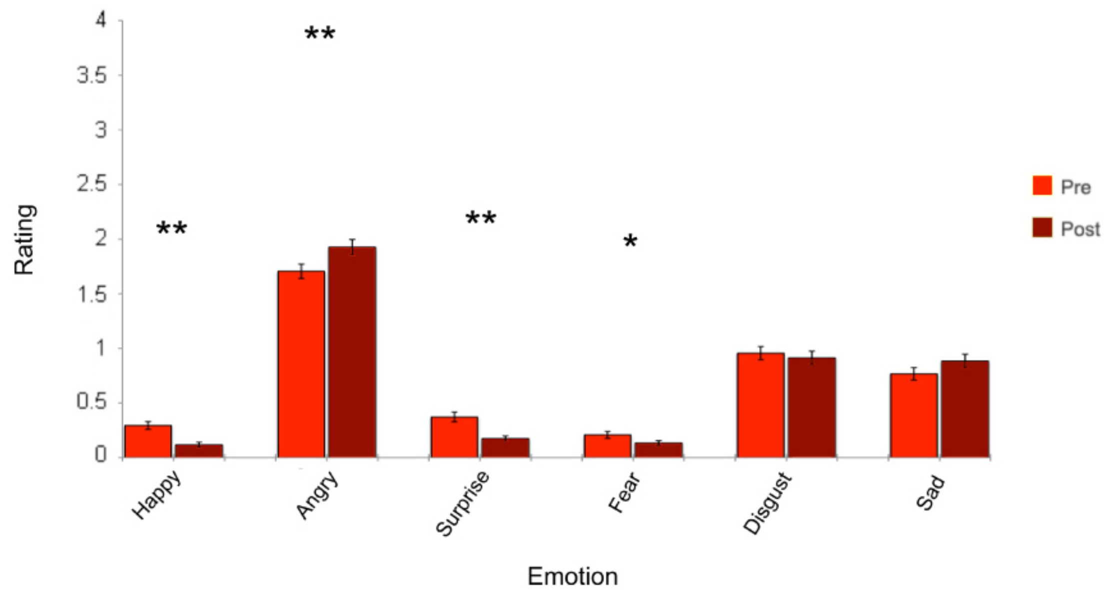

Figure 17. Bar-graph of expression quality ratings for the Angry expression, before and after training. Error bars denote Standard Error of the mean. Asterisk represents a significant difference at $p < 0.05$. Double asterisk represents significance at $p < 0.01$.

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 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 Figure 18).

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 Figure 18).

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 Figure 18).

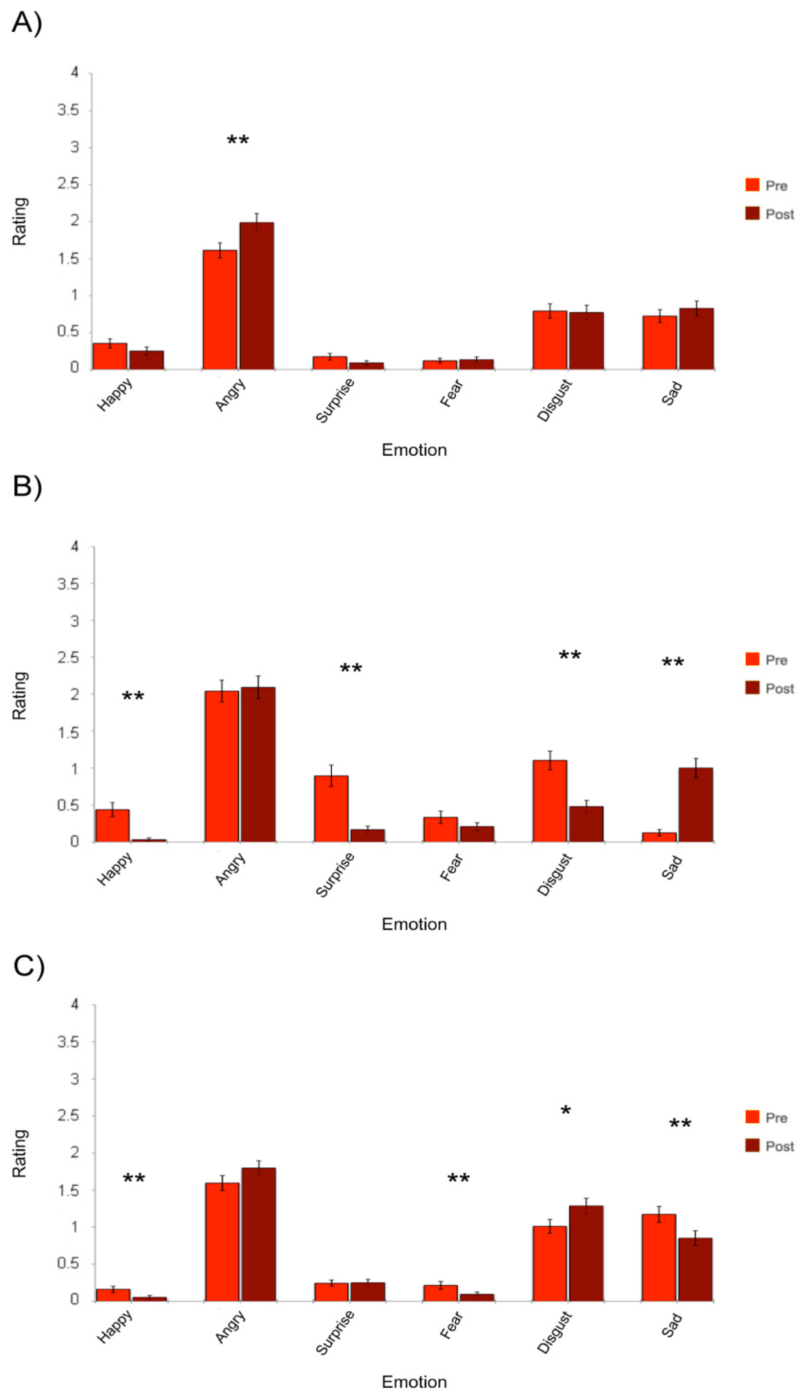

Figure 18. Bar-graph of expression quality ratings for the a) Level 1 group, b) Level 2 group and c) Level 3 group, before and after training. Error bars denote Standard Error of the mean. Asterisk represents a significant difference at $p < 0.05$. Double asterisk represents significance at $p < 0.01$

Surprise. For the control Surprise expression, analysis revealed a significant main effect of Rating, F (3.18, 1293.66) = 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 x Rating was observed, F (3.54, 1441.48) = 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 Figure 19).

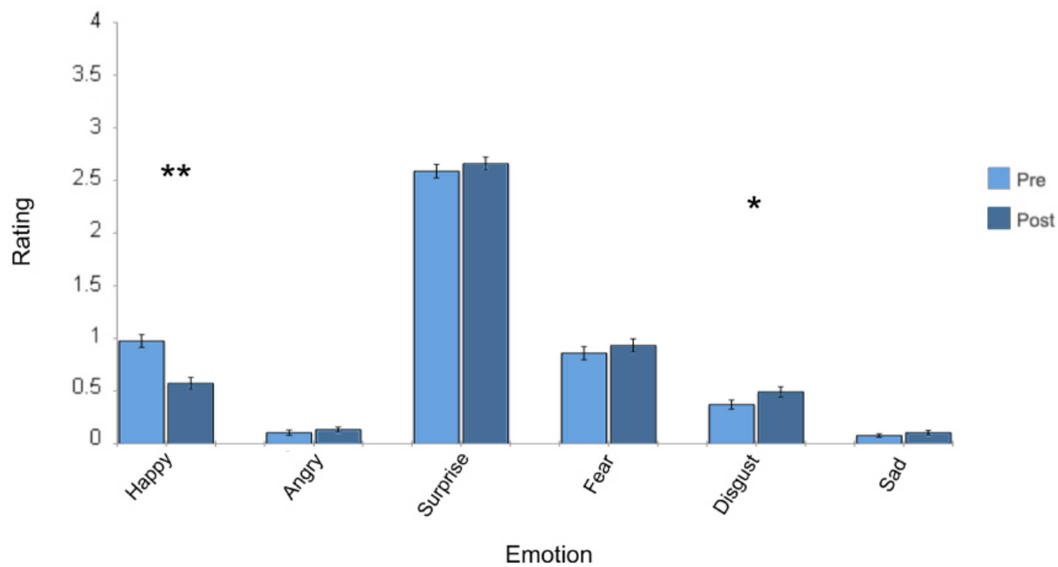

Figure 19. Bar-graph of expression quality ratings for the Surprise expression, before and after training. Error bars denote Standard Error of the mean. Asterisk represents a significant difference at $p < 0.05$. Double asterisk represents significance at $p < 0.01$.

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 FaceMaze Video Ratings of ASD 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 Figure 20). 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 Figure 21). 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.

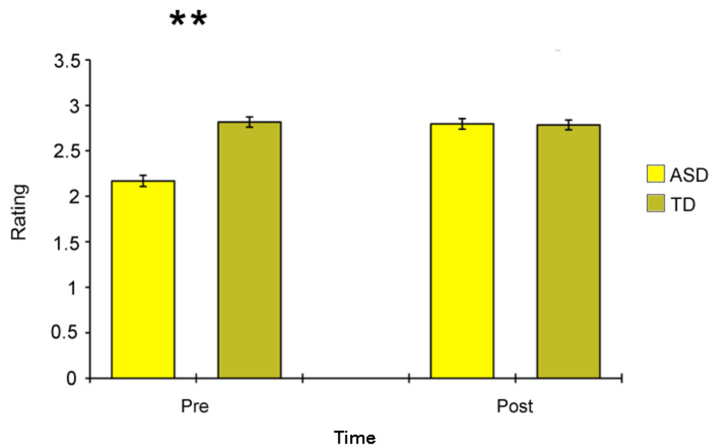

Figure 20. Bar-graph of happy expression quality ratings for the HappyMaze condition, before and after training, for both ASD and TD groups. Error bars denote Standard Error of the mean. Double asterisk represents a significant difference at $p < 0.005$.

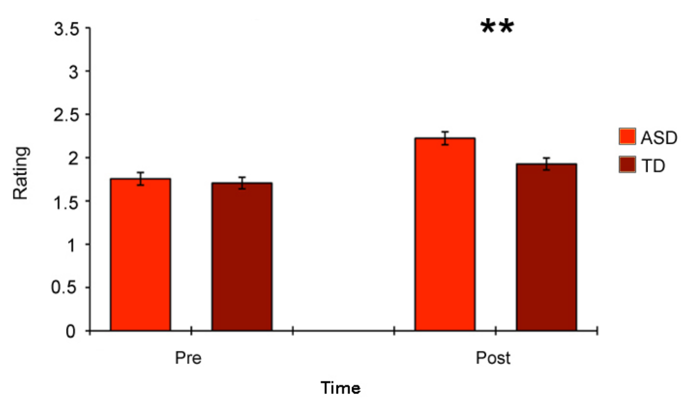

Figure 21. Bar-graph of angry expression quality ratings for the AngryMaze condition, before and after training, for both ASD and TD groups. Error bars denote Standard Error of the mean. Double asterisk represents a significant difference at $p < 0.005$.

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 & 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 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, Sullivan and Vasen, 1987; Odom & 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 adults (Lewis, Sullivan and Vasen, 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 developmental trend in the perception of both TD and ASD children's negative displays (Lewis, Sullivan and Vasen, 1987; Macdonald, Rutter, Howlin, RiosLe Conteur, Evered, and Folstein, 1989; Odom & 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, Cottrell, Padgett, & Adolphs, 2002; Smith & Scott, 1997; Susskind, Littlewort, Bartlett, Movellan, & Anderson, 2007), and perception (Aviezer, Hassin, Ryan, Grady, Susskind, Anderson, Moscovitch, Bentin, 2008; Bullock & Russel, 1984; Ekman & Friesen, 1975; Widen & 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 (Russel and Barrett, 1999). Results from the current study help clarify the relationship between Angry and Disgust facial expressions, by showing that the expression quality of disgust is part of Angry facial displays.

Findings from the current experiment underscore the efficacy of FaceMaze in training facial expression production by providing evidence for the paradigm's effectiveness with younger and clinical populations. Chapter 4 further discusses these findings with respect to embodied cognition and the developmental trajectory.

Chapter 4

General Discussion

The goal of the current project was to validate the efficacy of the FaceMaze game in facilitating higher quality facial expressions. Findings from adult, TD, and ASD children participants show that FaceMaze not only enhances the quality of targeted facial expressions, but also decreases the production of competing facial displays. Critically, no changes were observed with respect to the control Surprise expression, underscoring that the gains observed in entrained facial displays resulted from directed training, and not as an epiphenomenon to manipulating one's facial muscles.

Previous developmental research has shown that positive and negative displays follow different developmental trajectories (Ekman, Roper, Hager, 1980; Field and Waldmen, 1982; Lewis, Sullivan and Vasen, 1987; Odom and Lemond, 1972). Positive expressions, such as happy, are produced more efficiently than negative expressions across the age spectrum (Ekman, Roper, Hager, 1980; Field and Waldman, 1982; Odom and Lemond, 1972), and also show high quality productions as early as 5 years of age (Odom and Lemond, 1972; Lewis, Sullivan and Vasen, 1987). Results from the current project are in line with these findings, with TD children and adults showing less ambiguity and high target expression quality ratings for Happy expressions even before training. Despite adults' proficiency in producing Happy facial expressions pre-FaceMaze, increases in target expression quality ratings after game-play serve to underscore FaceMaze's efficacy as a training module. Interestingly, TD children showed no further enhancement with respect to the happy expression quality rating, however this finding may not be unusual given research demonstrating that children between the ages of 8-12 tend to inhibit spontaneous and voluntary facial displays as a result of modesty or embarrassment, but adults do not (Izard, 1971; Yarczower, Kilbride, & Hill, 1979). Since most of the child participants in the current study were between the ages of 8-12 years, it is likely that children were producing inhibited displays, especially with the experimenter present.

Current findings with respect to the Angry facial displays are in line with those of previous research showing an increase in expression quality with development (Lewis, Sullivan, Vassen, 1987; Odom & Lemond, 1972). Despite showing poorer performance than Happy displays, adults not only produced more complete Angry displays when compared to children, but also produced higher-quality partial displays. Children, in comparison, showed a marked deficit in expression quality, with less than half producing an expression that could be categorized as Angry, and even fewer producing partial or complete displays (Lewis, Sullivan, Vassen, 1987; Odom & Lemond, 1972). Results from the current project corroborate these findings, with children showing more ambiguous Angry facial displays pre-FaceMaze than adults, as indicated by post-FaceMaze gains. That is, whereas both TD children and adults showed increases in the target angry expression quality rating, only children showed a decrease in other competing expression quality ratings after training. It is important to note that the lack of change in competing expression quality ratings in adults' post-Facemaze ratings did not result from a lack of training efficacy, as evidenced by the decrease in happy expression quality ratings. Rather, adult's pre-FaceMaze Angry expressions already showed the characteristic expression quality pattern that typifies these displays, which include aspects of disgust (Aviezer et al., 2008; Galati, Scherer, Ricci-Bitti, 1997; Russel &

Barrett, 1999). Expression quality ratings of disgust were also enhanced for TD children after playing Facemaze, demonstrating that the program facilitated more genuine-looking expressions.

ASD children showed the most gains post-FaceMaze, with higher happy and angry expression quality ratings for the Happy and Angry facial expressions, respectively. Furthermore, both Happy and Angry expressions showed a reduction in competing expression quality ratings, underscoring that not only were the targeted facial displays enhanced, but expressions had also become less disorganized as well. More importantly, increases in expression quality ratings of disgust for the Angry expression further highlight the natural tone with which these expressions were produced, emphasizing that the facial expressions entrained were not mechanical or artificial looking. From a technical perspective, the efficacy of FaceMaze in treating ASD symptoms may emerge from its similarity to interventions using Applied Behavioral Analysis (ABA), that have shown great proficiency in Autism intervention (Meyers, 2007/2010). ABA relies on operant conditioning techniques, such as prompting and reinforcement, to shape a desired behavior without necessarily relying on verbal instruction. In line with these techniques, FaceMaze shapes facial displays by providing visual prompts (emotive icons), and immediate reinforcement in the shape of a visual cue (filling of the expression meter) in response to desired behavior (a high-quality expression). Although other treatment programs targeting facial expression production have shown positive results, they necessitate specialized training and are labor-intensive, requiring one-on-one tutoring with human therapists over the course of several days (Charlop, Dennis, Carpenter, and Greenberg, 2010; DeQuiznio, Townsend, Sturmey &

Poulson, 2007; Gena, Krantz, McClannahan, & Poulson, 1996; Stewart and Singh, 1995), and present a "tiring procedure for therapists to use and difficult to use with consistency" (Gena et al., 1996, p. 547). Furthermore, these treatments may also be more difficult to implement as a result of co-morbid social anxiety in ASD. In contrast, FaceMaze relies on a computerized detection module that reliably encodes behavior, and presents consistent reinforcement with respect to behavioral response (expression quality), eliminating confounds of therapist fatigue, participant's social anxiety, and potential linguistic barriers. Furthermore, Facemaze is a cost-effective training program in facial expression production that is not only engaging for the child, but can be conducted in a safe, familiar setting such as the child's home.

Whereas previous research has been able to demonstrate the effects of facial expression production on the perception of facial emotion (Niedenthal, Brauer, Halberstadt, & Innes-Ker, 2001; Oberman, Winkielman, Ramachandran, 2007; Strack, Martin, & Stepper, 1988), findings from the current study have shown that the opposite also holds true, with perception (in the form of the "expression meter") influencing facial affect production quality. Furthermore, patterns of facial muscle movement (motor) were explicitly connected to proprioceptive (sensory) feedback that resulted from a deliberate, sustained pattern of muscle contraction (i.e. facial expressions), which further enhanced sensory-motor integration in the expression concept, as evidenced by the quality of facial expression posed after playing FaceMaze. This integration not only provides a more naturalistic process to expression training, but would also allow for an enhancement of all cognitive aspects of emotion expression, such as expression recognition (Atkinson & Adolphs, 2005; Deriso, Susskind, Tanaka, Winkielman, Herrington, Schultz, & Bartlett, 2012; Goldman & Sripada, 2005; Niedenthal, Brauer, Halberstadt, & Innes-Ker, 2001; Oberman, Winkielman, Ramachandran, 2007). Previous research investigating the expression production-perception link has demonstrated that individuals who underwent targeted expression production training were more efficient in detecting changes, and at categorizing dynamic displays of facial expressions than individuals who merely engaged in a control task. Furthermore, findings were titrated based on level of engagement, with participants who had high engagement in the training protcol outperforming those who had low engagement in the training protocol (Deriso, Susskind, Tanaka, Winkielman, Herrington, Schultz, & Bartlett, 2012). Additionally, research has shown a positive correlation between the quality of voluntary and spontaneous expressions (Berenbaum & Rotter, 1992) suggesting that the two processes may rely on similar cognitive mechanisms (Winkielman, McIntosh, & Oberman, 2009; Winkielman, Niedenthal & Oberman, 2009). Thus, 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 Facemaze shows great promise, some limitations are evident. First, all of our ASD and TD participants had some verbal proficiency, with most participants scoring in the average range. In order to verify the efficacy of the FaceMaze paradigm, future research may look to measure changes in facial expression production in low-functioning or preverbal children with ASD. Second, whereas post-FaceMaze facial expressions showed enhanced readability, whether or not these changes were maintained in the children's daily lives is still yet to be seen. Furthermore, how much training would be required to maintain these changes is still in question. Follow-up studies may look to measure changes in facial expression quality after implementing different FaceMaze training schedules, as well as sample facial expressions produced in a variety of settings, such as classroom or in the home. Finally, previous research has underscored gender differences in the perception of facial expressions, with females outperforming males on several measures of recognition (Hall, 1978; Hal, Carter, & Horgan, 2000; Hal & Matsumoto, 2000). Since majority of the raters in our studies were female, it is possible that the differences found resulted from the sensitivity of female judges to expression quality, thus future research may also look to more balanced designs with respect to expression rating.

From a theoretic standpoint, it is difficult to determine the extent to which the FaceMaze resulted in changes in the expression concept, as both pre and post-training measures only assessed changes in expression production. According to the embodied cognition approach, changes in the expression concept should result in changes in all related facets of cognition, thus changes in facial expression perception should be evident as well. Whereas the current study did not take these measures into account, future research could provide support for the embodied approach by demonstrating a causal link between the increase in facial expression production quality and the enhancement of facial expression detection.

Finally, one factor that had not been addressed in the current project is the effect of social norm on facial expression production. Previous research has demonstrated that facial expressions are subject to display rules, that is, cultural norms governing when it is appropriate to display certain facial expressions, and what is considered an acceptable display (Camras et al., 1999; Camras, Chen, Bakeman, Norris & Cane, 2006; Ekman, 1973; Matsumoto, 2001). Such display rules have been shown to affect children's abilities in producing readable facial expressions, with cultures that encourage positive displays, such as Happy, fostering more Happy expressions in children as young as 3 years of age (Camras et al., 1999; Camras et al., 2006; Matsumoto, 2001). With respect to the current experiment, it is possible that the FaceMaze was encouraging certain facial expression displays by demonstrating the "appropriate" display through the emotive icons. As a result, changes in expression quality post-FaceMaze could have resulted from children attempting to meet demand characteristics, and not from changes in the expression concept per se. Future research may look to include a separate control condition where children are not required to produce facial expressions during training, but are exposed to pictures of "appropriate" displays in order to determine the effect of demand characteristics on mimicry.

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, NewYork: GuilfordPress

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., Moscovitch, M. & Bentin, S. (2008). Angry, disgusted, or afraid? *Psychological Science, 19*(7), pp. 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, pp. 75–92.

Bartlett, M.S., Littlewort, G., Frank, M.G., Lainscsek, C., Fasel, I., and Movellan, J.R. (2005). Recognizing Facial Expression: Machine Learning and Application to Spontaneous Behavior. *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), pp. 22-35.

Bartlett, M., Whitehill, J. (2011). Automated facial expression measurement: Recent applications to basic research in human behavior, learning, and education. In *Handbook of Face Perception, Andrew Calder, Gillian Rhodes, James V. Haxby, and Mark H. Johnson (Eds)*. 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*, pp. 179 –190.

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

Buck. R, (1984). *The communication of emotion.* NewYork: Guilford Press.

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

Cacioppo, J. T., & Petty, R. E., (1981). Electromyograms as measures of extent and affectivity of information processing. *American Psychologist, 36,* pp. 441-456.

Cacioppo, J.T., Petty, R.E., Losch, M.E., & Kim, H.S., (1986) Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. *Journal of Personality and Social Psychology*, 50, pp. 260–268.

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., Chen, Y., Bakemen, R., Norris, K., & Cain, T.R., (2006). Culture, Ethnicity, and Children's Facial Expressions: A Study of European American, Mainland Chinese, Chinese American, and Adopted Chinese Girls. *Emotion, 6*(1), pp. 103–114.

Camras, L. A., Oster, H., Campos, J., Campos, R., Ujiie, T., Miyake, K., Wang, L., & Meng, Z. (1998). Production of emotional facial expressions in European American, Japanese, and Chinese infants. *Developmental Psychology*, *34*(4), pp. 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), pp. 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. *Proceedings from the IEEE International Conference on Automatic Face & Gesture Recognition* (peer-reviewed conference proceeding), pp. 978–986.

Damasio, A. R., Grabowski, T. J., Bechara, A., Damasio, H., Ponto, L. L., Parvizi, J., & Hichwa, R. D. (2000). Distinctive patterns of subcortical and cortical brain activation associated with self-generated emotions and feelings. *Nature Neuroscience*, *3*(10), pp. 1049-1056.

Darwin, C. (1872/1965). *The expression of the emotions in man and animals* (Vol. 526). University of Chicago press.

Dailey, M.N., Cottrell, G.W., Padgett, C., & Adolphs, R. (2002). EMPATH: A neural network that categorizes facial expressions. *Journal of Cognitive Neuroscience, 14,* pp. 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), pp. 755–759.

Deriso, D. M., Susskind, J., Tanaka, J., Winkielman, P., Herrington, J., Schultz, R., & Bartlett, M. (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). Springer Berlin Heidelberg.

Dimberg. U., (1982). Facial reactions to facial expressions. *Psvchophysiology, 19,* pp. 643-64.

Dimberg, U., (1983). Emotional conditioning to facial stimuli: A psychobiological analysis, *Acta Universitati Upsaliensis. Abstracts of Uppsala dissertations from faculty of Social Sciences, 29.* Uppsala: Almqvist &Wiksell.

Dimberg, U., (1990). Facial electromyography and emotional reactions. Psycholphysiology, 27(5), pp. 481-494.

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological science*, *11*(1), pp. 86-89.

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

Duchenne De Boulogne, G.B. (1990). *The Mechanism of Human Facial Expres-*

sion. A. Cuthbertson, Trans. & Ed. Cambridge Universitiy Press. New York.

Ekman, P. (1973). Cross-cultural studies of facial expression. *Darwin and facial expression: A century of research in review*, pp. 169-222.

Ekman, P. (1993). Facial Expressions of Emotion. *American Psychologist*, *43*, pp. 384–392.

Ekman, P., (2006). *Darwin and facial expression: a century of research in review*. Cambridge, MA: Malor Books.

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

Ekman, P., & Friesen, W, V. (1978). *The facial action coding system (FACS).* California: Consulting Psychologists Press.

Ekman, P., Friesen, W.V., & Hager, J.C., (2002). *The Facial Action Coding System.* Research Nexus eBook. Salt Lake City, UT.

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science*, *221*(4616), pp. 1208-1210.

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

Fridlund A. (1994). Human Facial Expression: An Evolutionary View. New York: Academic Press.

Fridlund, A.J., & Izard, C.E. (1983). Electromyographic studies of facial expression of emotions and pattems of emotion. In J.T. Cacioppo & R.E. Petty (Eds.), *So-cial psychophysiology: A sourcebook* (pp. 243-286) New York: The Guilford Press.

Galati, D., Scherer, K.R., & Ricci-Bitti, P.E., (1997). Voluntary facial expressions of emotion: Comparing congenitally blind with normal sighted encoders. *Journal of Personality and Social Psychology*, *73*, pp. 1363-1379.

Gena, A., Krantz, P. J., McClannahan, L. E., & Poulson, C. L. (1996). Training and generalization of affective behavior displayed by youth with autism. *Journal of Applied Behavior Analysis*, *29*(3), pp. 291-304.

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

Gordon, I., Pierce, M.D., Bartlett, M.S., & Tanaka, J.W. (2014) Training facial expression production in children with autism spectrum disorder. *Journal of Autsim and Developmental Disorders*. doi 10.1007/s10803-014-2118-6

Gosselin, P., Perron, M., & Beaupré, M. (2010). The voluntary control of facial action units in adults. *Emotion*, *10*(2), pp. 266.

Hall, J.A., (1978). Gender effects in decoding non-verbal cues. *Psychological Bulletin, 85*, pp. 845-857.

Hall, J.A., Carter, J., Horgan, T. (2000). Gender differences in the nonverbal communication of emotion. In: *Fischer A, editor. Gender and emotion: Social psychological perspectives.* Paris: Cambridge University Press. pp. 97–117.

Hall, J.A., Matsumoto, D., (2004). Gender differences in judgments of multiple emotions from facial expressions. *Emotion*, *4*, pp. 201–206.

Hennenlotter, A., Dresel, C., Castrop, F., Ceballos-Baumann, A. O., Wohlschläger, A. M., & Haslinger, B. (2009). The link between facial feedback and neural activity within central circuitries of emotion—New insights from Botulinum toxin–induced denervation of frown muscles. *Cerebral Cortex*, *19*(3), pp. 537-542.

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

Hjortsjö, C.H., (1970). *Man's face and mimic language*. Malmo, Sweden: Nordens Boktryckeri.

Izard, C.E., (1971). The face of emotion. New York:Appleton - Century Crofts

Izard, C. E. (Ed.). (1977). *Human emotions*. Boom Koninklijke Uitgevers.

Izard, C. E., (1979). *The maximally discriminative facial movement coding system (Max).* Newark, DE: University of Delaware, Office of Academic Computing and Instructional Technology.

Izard, C.E., (1983). *The Maximally Discriminative Facial Movement Coding System (revised edition).* Instructional Resources Center, Newark, DE.

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. (1968). Autistic disturbances of affective contact. *Acta Paedopsychiatr, 35* (4), pp. 100–36.

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), pp. 690-697.

Levenson, R. W., Ekman, P., & Friesen, W. V. (1990). Voluntary facial action generates emotion‑specific autonomic nervous system activity. *Psychophysiology*, *27*(4), pp. 363-384.

Littlewort G, Whitehill J, Wu T, Fasel I, Frank M, Movellan J, and Bartlett M. (2011). The Computer Expression Recognition Toolbox (CERT). *Proc. IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 298–305.

Lord C, Risi S, Lambrecht L, Cook EH, Jr, Leventhal BL, DiLavore PC, Pickles A, Rutter M. (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), pp. 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*, pp. 433–444.

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

Matsumoto, D. (2001). Culture and emotion. In D. Matsumoto (Ed.), *Handbook of culture and psychology*. New York: Oxford University Press, pp. 171–194.

McIntosh, D. (2006). Spontaneous facial mimicry, liking, and emotional contagion. *Polish Psychological Bulletin, 37,* pp. 31– 42.

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), pp. 295–302.

Meyers, S.M. (2007) Management of Children With Autism Spectrum Disorders. *Pediatrics, 120*(5), pp. 1162–1182. Reaffirmed September 2010.

Miller, P., & Eisenberg, N. (1988). The relation of empathy to aggressive and externalizing/antisocial behavior. *Psychological Bulletin*, *103*, pp. 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*, pp. 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*, pp. 167–178.

Oberman, L. M., Winkielman, P., & Ramachandran, V. S. (2009). Slow echo: facial EMG evidence for the delay of spontaneous, but not voluntary, emotional mimicry in children with autism spectrum disorders. *Developmental science*, *12*(4), pp. 510-520.

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

Piaget, J. (1951/2013). *Play, dreams and imitation in childhood* (Vol. 25). Routledge.

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*, pp. 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), pp. 805–819.

Rusalova, M. N., Izard, C. E., & Simonov, P. V. (1975). Comparative analysis of mimical and autonomic components of man's emotional state. *Aviation. Space, and Environmental Medicine. 46,* pp. 1132-1134.

Schwartz, G. E. (1975). Biofeedback, self-regulation, and the patterning of physiological processes. *American Scientist, 63,* pp. 314-324.

Schwartz, G. E., Brown, S., & Ahern, G., (1980). Facial muscle patterning and subjective experience during affective imagery: Sex differences. *Psychophysiology*, *17*, pp. 75– 82.

Schwartz, G. E., Fair, P. L., Salt, P., Mandel, M. R., & Klerman, G. L. (1976). Facial muscle patterning to affective imagery in depressed and nondepressed subjects. *Science*, *192*, pp. 489– 491.

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), pp. 365–382.

Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: a nonobtrusive test of the facial feedback hypothesis. *Journal of personality and social psychology*, *54*(5), pp. 768.

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), pp. 152–162.

Tian, Y. L., Kanade, T., & Cohn, J. F. (2001). Recognizing action units for facial expression analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *23*(2), pp. 97-115.

Tomkins, S. S. (1962). Affect, imagery, consciousness: Vol. I. The positive affects.

Vygotsky, L. S. (1967). Play and its role in the mental development of the child. *Journal of Russian and East European Psychology*, *5*(3), pp. 6-18.

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

Wild, B., Erb, M., Eyb, M., Bartels, M., & Grodd, W. (2003). Why are smiles contagious? An fMRI study of the interaction between perception of facial affect and facial movements. *Psychiatry Research: Neuroimaging*, *123*(1), pp. 17-36.

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

Winkielman, P., Niedenthal, P., & Oberman, L. M. (2009). Embodied Perspective on Emotion-Cognition Interactions. In *J. Pineda (Ed.) Mirror Neuron Systems.* (pp. 235-257). Humana Press, New York.

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

Yarczower, Matthew Kilbride, Janet E. Hill, Linda A. (1979) Imitation and inhibition of facial expression. *Developmental Psychology*, *15*(4), pp. 453-454.