# *Behind the Mask:*

# Recognising Genuine and Masked Expressions of Emotion:

# The Effect of Therapists' Training and Experience

**A thesis presented in partial fulfillment of the requirements for the degree of Doctor of Clinical Psychology at Massey University, Palmerston North, New Zealand**

**Alexa Anne Curtis**

**2017**

# ABSTRACT

Accurately recognising facial expressions of emotion can enhance communication and the development of a therapeutic relationship. When emotions are masked or inhibited, duplicity can be betrayed through evidence of leakage of the underlying emotion occurring in the face. Being able to discern when emotions are masked or concealed may also contribute to therapy outcomes by alerting the therapist to areas requiring further exploration. Despite a large body of research on facial expression of emotion, there is a dearth of research into therapists' emotion recognition competencies or ability to detect deception. This study sought to answer the following questions: First, is recognition of facial emotional expressions, including masked expressions, enhanced by training? Second, does clinical experience impact on emotion recognition ability? Finally, does training increase the ability to discern authentic from falsified expressions? In Study 1, 43 clinical psychologists were recruited for the study and undertook an emotion recognition and deception detection task. Twenty-two of the participants completed training in emotion recognition using the Subtle Expression Training Tool (SETT; http://www.paulekman.com). In Study 2, a partial replication of Study 1 was conducted with 25 participants, and an updated version of the SETT. Results for recognition of emotion were mixed, with partial support found for the hypothesis that training would enhance recognition ability. Training was effective for improving recognition of single emotion expressions, but had less impact on recognition for masked or leaked expressions, with the exception of sadness. Efficacy of training was also dependent on experience level, with more experienced participants benefiting the most. All participants were able to detect emotional deception at levels above chance prior to training. Following training, those with a high level of experience demonstrated the greatest improvement in deception detection, with half of this group accurately detecting deception at levels significantly above chance ($M$ = .74).

# ACKNOWLEDGEMENTS

I have been privileged to have had the support of many wise friends, colleagues, family, and affirming mentors throughout my thesis journey. First and foremost, my heartfelt thanks go to my supervisors, Associate Professor John Podd, Dr. Stephen Hill, and Dr. Shane Harvey.

John, I cannot thank you enough. You have been a constant source of encouragement, wisdom, and humour, with an awe-inspiring wealth of knowledge and expertise. Thank you for your positive mentoring style, generous lashings of patience - particularly during the analysis stage, and the ease with which you restored my confidence when it was flagging. Thank you for keeping me on track by continuing to shine the light at the end of the tunnel. You are indeed a gifted supervisor.

To Stephen, thank you for your critical eye and your expert assistance with the analysis and results sections. And to Shane, I appreciate the many informal chats we shared that kept me passionate about the importance of researching emotion, which helped to keep me focused. I also thank Mr Malcolm Loudon and his technical skills which assisted me to analyse my signal detection results and his development of the emotion expression computer programme.

To my family and friends, you have all believed in my ability to complete this thesis, even when I doubted it myself. To my husband Wayne, thank you for your enduring patience and understanding. You had no idea what you were letting yourself in for when you agreed that my decision to complete a doctorate was a good idea. We were both naïve to the challenge that lay ahead of us. Words are not enough to convey the depth of my gratitude to you for your support – emotionally, financially, and your ability to always lift my spirits with your humour. To Michaela and Aimee, I am deeply thankful for having two caring and thoughtful daughters. I am immensely proud of how you have both been inspired by my journey, and achieved your own academic successes. To my parents, Sandy and Jenny, thank you for instilling in me the work ethic necessary to undertake doctoral study.

I would also like to acknowledge the participants who willingly gave their time and engaged so sincerely in the emotion induction process used to create the stimuli for the main study. To all the psychologists who gave up two hours out of their busy schedules to participate in my study, I am most grateful.

Warm thanks to my work colleagues at the Massey Psychology Clinic. Each one of you, in your own way, through encouraging words in the tearoom or some sage advice, helped me to stay on track. I especially owe thanks to Senior Clinician, Jan Dickson, who supervised me throughout my internship which is an integral element of this doctoral degree. Jan, you are an absolute gem, and admired by all clinical students, past and present. I was most fortunate to have your guidance and unwavering support during the most difficult year of the doctorate.

Thank you to all my friends for your support, and to my clinical psychology student colleagues, who through climbing your own Everests, always offered appropriate encouragement. A very special thanks to Rachel Hooks, my fellow intern, for helping me retain my sanity during our intern year and for so willingly agreeing to the arduous task of reading my draft thesis.

I am also grateful for the financial assistance provided to me by Massey University (Doctoral Completion Scholarship) and the Massey Alumni (Doctoral Completion Scholarship).

To the Ekman Group who allowed me free access to their online training programme, the Subtle Emotion Training Tool (SETT), your generosity is most appreciated.

# TABLE OF CONTENTS


**ABSTRACT** ..... ii

**ACKNOWLEDGEMENTS** ..... iii

**TABLE OF CONTENTS** ..... v

**LIST OF TABLES** ..... ix

**LIST OF FIGURES** ..... x

**CHAPTER ONE: EMOTION AND FACIAL EXPRESSIONS** ..... 1
- **What are Emotions?** ..... 1
- **Facial Expressions of Emotion** ..... 4
- **Universality of Expressions** ..... 5
- **Emotion Recognition Abilities (ERA)** ..... 10
  - The Utility of ERA ..... 12
  - Therapists ..... 15
  - Clinical and Non-Clinical Populations ..... 16
  - Acquired Brain Injury and ERA ..... 17
  - ERA Differences Across the Adult Lifespan ..... 17
  - Sex Differences ..... 21
  - ERA and Negotiation Skills ..... 22
- **Measurement of Facial Expressions** ..... 23
  - Electromyography (EMG) ..... 23
  - The Facial Action Coding System (FACS) ..... 25
  - Automated Facial Recognition Systems ..... 28
- **Training in Emotion Recognition** ..... 29
  - Non-Clinical Populations ..... 30
  - Clinical Populations ..... 31
- **Summary of Emotion Research** ..... 32

**CHAPTER TWO: DECEPTION** ..... 34
  - Neurobiological Basis of Emotional Expressions ..... 35
  - The Inhibition Hypothesis ..... 36
  - Inconsistent and Leaked Expressions ..... 36
  - Felt and False Smiles ..... 38
  - Spontaneous and Masked Emotional Expressions ..... 40
  - Reliable Facial Muscles ..... 41
  - Micro-Expressions ..... 43
  - Subtle Expressions ..... 44
  - Verbal Deception ..... 46
  - Emotional Deception ..... 51
  - Facial Behaviour in High-Stakes Deception ..... 54
  - Individual Factors Contributing to Deception Accuracy ..... 55
- **Summary of Deception Research** ..... 57

**CHAPTER THREE: THE CURRENT INVESTIGATION** ..... 60
- **Aims and Objectives** ..... 60

**Rationale** ... **61**
**Hypotheses** ... **63**
Hypothesis 1 ... 63
Hypothesis 2 ... 63
Hypothesis 3 ... 63

**CHAPTER FOUR: METHOD** ... **64**
**Aim** ... **64**
**Stage 1: Creation of Emotion and Deception Stimuli** ... **64**
Participants ... 64
Ethics ... 65
Stimuli ... 65
Apparatus ... 68
Procedure ... 69
**Stage 2: Deception Detection and Emotion Recognition Task and Facial Emotion Recognition Training – Study 1** ... **76**
Participants ... 76
Stimuli ... 77
Procedure ... 78

**CHAPTER FIVE: DATA ANALYSIS** ... **82**
**Overview of the Data Analysis** ... **82**
**Age and Experience Level** ... **83**
**Analysis of Emotion Recognition** ... **86**
**Analysis of Deception Detection** ... **86**

**CHAPTER SIX: RESULTS FOR STUDY 1** ... **89**
**Emotion Recognition** ... **89**
Single Emotion Expressions ... 90
Masked emotion expressions ... 95
Leaked Emotion Expressions ... 99
**Summary of Results for Emotion Recognition** ... **102**
**Deception Detection** ... **104**
Signal Detection ... 104
Sensitivity - P(A) ... 105
**Summary of Results for Deception Detection** ... **110**

**CHAPTER SEVEN: DISCUSSION - STUDY ONE** ... **111**
**Synopsis of Study One's Aims** ... **111**
**Findings From the Emotion Recognition Task** ... **112**
Single Emotion Expressions ... 112
Masked Emotion Expressions ... 116
Leaked Emotion Expressions ... 118
**Summary of Emotion Findings** ... **118**
**Findings from the Deception Detection Task** ... **120**
Sensitivity *P(A)* ... 120
Bias (*c*) ... 123
**Summary of Deception Detection Findings** ... **123**

**CHAPTER EIGHT: STUDY TWO** ... **125**
**Deception Detection and Emotion Recognition Task and Emotion Recognition Training – Study 2** ... **125**
Participants ... 126
Stimuli ... 126

Procedure ........ 128
**Data Analysis** ........ **128**
**Age and Experience Level** ........ **128**

**CHAPTER NINE: RESULTS FOR STUDY 2** ........ **131**
**Emotion Recognition** ........ **131**
Single Emotion Expressions ........ 132
Masked Emotion Expressions ........ 136
Leaked Emotion Expressions ........ 139
**Analysis of Deception Detection** ........ **142**
**Results for Deception Detection and Bias** ........ **143**
Sensitivity - *P(A)* ........ 143
Response Bias (*c*) ........ 144

**CHAPTER TEN: DISCUSSION – STUDY TWO** ........ **146**
**Synopsis of Study 2'S Aims** ........ **146**
**Findings from the emotion recognition task** ........ **148**
Single Emotion Expressions ........ 148
Masked Emotion Expressions ........ 151
Leaked Emotion Expressions ........ 152
**Summary of Emotion Findings** ........ **153**
**Findings from the Deception Detection Task** ........ **155**
Sensitivity *P(A)* ........ 155
Bias (*c*) ........ 155
**Summary of Deception Detection Findings** ........ **157**
**Evaluation of New SETT Version** ........ **158**

**CHAPTER ELEVEN: LIMITATIONS, STRENGTHS, and SUGGESTIONS for FUTURE RESEARCH** ........ **159**
**Limitations** ........ **159**
**Strengths** ........ **166**
**Suggestions for Future Research** ........ **168**
**Conclusion** ........ **169**

**REFERENCES** ........ **172**

**APPENDICES** ........ **189**
**Appendix A** ........ **190**
Poster Advertisement ........ 190
**Appendix B** ........ **192**
Information Sheet ........ 192
**Appendix C** ........ **195**
Participant Consent Form ........ 195
**Appendix D** ........ **197**
Participant Questionnaire ........ 197
**Appendix E** ........ **199**
Instructions for Powerpoint Slideshow ........ 199
**APPENDIX F** ........ **201**
Images used from the International Affective Picture System ........ 201
**Appendix G** ........ **204**
Statements by New Zealand Rugby Players ........ 204
**Appendix H** ........ **208**
Invitation to Psychologists to Participate in the Study ........ 208
**Appendix I** ........ **210**

Information Sheet ........ 210
**Appendix J ........ 213**
Participant Consent Form ........ 213
**Appendix K(a) ........ 215**
Instructions for Training Group Participants ........ 215
**Appendix K(b) ........ 222**
Instructions for No Training Group Participants ........ 222
**Appendix L ........ 229**
Ethics Approval ........ 229
**Appendix M ........ 231**
ANOVA tables – Study 1 ........ 231
**Appendix N ........ 239**
ANOVA tables – Study 2 ........ 239

# LIST OF TABLES


*Table 1 Emotion Categories and Expressions Requested ...... 69*

*Table 2 Descriptive Statistics Across Experimental Groups ...... 84*

*Table 3 Descriptive Statistics for Split by Level of Experience ...... 84*

*Table 4 Age and Years of Experience Across Experimental Conditions ...... 85*

*Table 5 Mean Percentage Scores or All Emotion Expressions for Each Group ...... 89*

*Table 6 Mean group SDT Values Across Experiemental Groups ...... 104*

*Table 7 Percentage of Participants' Deception Detection Accuracy ...... 106*

*Table 8 Percentage of Participants' Deception Detection Accuracy by Group ...... 107*

*Table 9 Descriptive Statistics Across Experimental Groups ...... 129*

*Table 10 Descriptive Statistics for Split by Level of Experience ...... 129*

*Table 11 Age and Years of Experience for Low and High Experience Groups Across Experimental Conditions ...... 130*

*Table 12 Mean Percentage Scores for All Emotion Expressions for Time 1 and Time 2 for Each Group and Experience Level ...... 132*

*Table 13 Mean Group Values for Time 1 and Time 2 for P(A), and c, for Low and High Experience across Experimental Groups ...... 143*

# LIST OF FIGURES

*Figure 1. Sample of images from eSett 3.0 (http://www.paulekman.com)* .......................................... 10

*Figure 2. Duchenne and non-Duchenne smiles.* ................................................................................ 38

*Figure 3. Sample of images from the Powerpoint slideshow* ............................................................ 67

*Figure 4. Prototypical Expressions* ................................................................................................ 70

*Figure 5. Examples of the stimuli used in this study* ........................................................................ 76

*Figure 6. Screenshot of selection screen.* ........................................................................................ 80

*Figure 7. Distribution of scores for correct identification of a single Happy expression* ....................... 91

*Figure 8. Change in scores for identifying a single Happy expression* ................................................ 91

*Figure 9. Change in scores for identifying a single Sad expression* .................................................... 93

*Figure 10. Change in scores for identifying a single Disgust expression* ............................................ 94

*Figure 11. Distribution of scores for correct identification of a Masked Happy expression* .................. 95

*Figure 12. Change in scores for identifying a masked Happy expression* ........................................... 96

*Figure 13. Change in scores for identifying a masked Sad expression* ............................................... 97

*Figure 14. Change in scores for identifying a masked Disgust expression* .......................................... 98

*Figure 15. Change in scores for identifying a leaked Happy expression* ............................................. 99

*Figure 16. Change in scores for identifying a leaked Sad expression* ............................................... 101

*Figure 17. Change in scores for identifying a leaked Disgust expression* .......................................... 102

*Figure 18. Indices of sensitivity as measured by P(A) for all groups* ................................................ 105

*Figure 19. Change in response bias* ................................................................................................ 109

*Figure 20. Distribution of scores for correct identification of a single Happy expression* ................... 132

*Figure 21. Change in scores for identifying a single Happy expression* ............................................. 133

*Figure 22. Change in scores for identifying a single Sad expression .* ............................................... 134

*Figure 23. Change in scores for identifying a single Disgust expression* ........................................... 135

*Figure 24. Distribution of scores for correct identification of a Masked Happy expression* ................ 136

*Figure 25. Change in scores for identifying a Masked Happy expression* ........................................ *137*

*Figure 26. Change in scores for identifying a masked Sad expression* .............................................. *138*

*Figure 27. Change in scores for identifying a masked Disgust expression* ........................................ *139*

*Figure 28. Change in scores for identifying a leaked Happy expression* ............................................ *140*

*Figure 29. Change in scores for identifying a leaked Sad expression* ................................................ *141*

*Figure 30. Change in scores for identifying a leaked Disgust expression* .......................................... *142*

*Figure 31. Indices of sensitivity as measured by P(A* ......................................................................... *144*

*Figure 32. Change in response bias*...................................................................................................... *145*

# CHAPTER ONE: EMOTION AND FACIAL EXPRESSIONS

*"The face is the dynamic canvas on which humans express emotional states and from which they infer those of others" (ten Brinke & Porter, 2012, p. 469).*

*Emotion is a fundamental component of human experience and interest in emotion is evident in history in the writings of many great philosophers such as Plato and Aristotle, and the evolutionist, Charles Darwin. Over the years, a number of theories of what constitutes emotion have abounded from various disciplines. Although a comprehensive review of the emotion literature is beyond the scope of this thesis, a brief summary of emotion theories is outlined in the first section. This is followed by a review of studies of facial expressions of emotion and the overwhelming support for the universality of emotion expressions. Then background information on emotion recognition abilities (ERA) is provided, along with a review of the research that has examined ERA in various populations. Concluding the chapter, systems for objectively measuring facial behavior are described, with a detailed description of the system used in the current study – the Facial Action Coding System (FACS: Ekman, Friesen, & Hager, 1978/2002), as well as a summary of studies investigating the impact of training on emotion recognition abilities.*

## WHAT ARE EMOTIONS?

The word 'emotion' is a relatively new term, only coming into use during the nineteenth century. Previously, words such as *passions* and *sentiments* were used to describe feelings we now think of as emotions (Oatley, 2004). In 1884, Carl Lange and William James independently published the same theory of emotion, stating that emotions are "the effect of the organic changes, muscular and visceral, of which the so-called 'expression' of the emotions consists. It is thus not a primary feeling, directly aroused by the exciting object or thought, but a secondary feeling indirectly aroused; the primary effect being the organic changes in question, which are immediate reflexes following upon the presence of the object" (James, 1894, p. 516). According to this theory, it is the physical arousal in response to a stimulus that creates the feeling of emotion.

In contrast to the Lange-James theory of emotion, Walter Cannon argued that emotions emerge from lower parts of the brain and are mediated by the cortex. In the 1920s, Walter Hess added further weight to Cannon's theory by eliciting emotional responses by electrically stimulating the hypothalamus and limbic system (Oatley, Keltner, & Jenkins, 2006).

The cognitive psychology movement led to the development of theories focusing on an individual's subjective evaluation or appraisal of environmental events. These can trigger various emotions such as anger, joy, or sorrow, which are influenced by experience. The Schachter-Singer two-factor theory of emotion proposes that emotions arise as a function of an interaction between physiological arousal and an individual's cognitive attributions for that arousal. Schachter and Singer propose that the same state of physiological arousal can be labeled with a different emotion, depending on the cognitive interpretation of the situation. It is the cognitions that determine whether "joy", "anger", or "fear", or some other emotion, is attributed to the state of physiological arousal (Schachter & Singer, 1962).

Contemporary approaches to understanding emotions range from theories of emotion rooted in biology (e.g., Buck, 1999), those claiming an interaction between culture and biology (Scherer, 2009), and others that argue that emotions are socially constructed, developing in social contexts which shape and define them (Boiger & Mesquita, 2012).

Although it has been claimed that emotion is too broad a concept to define (Oatley et al., 2006), a definition offered by Ekman and Cordaro (2011, p. 364) described emotions as “discrete, automatic responses to universally shared, culture-specific and individual-specific events. … These affective responses are preprogrammed and involuntary, but are also shaped by life experiences.” Izard (1992, p. 561) stated that emotion theorists, particularly those influenced by Darwin, “assume that emotions are specific neuropsychological phenomena, shaped by natural selection, that organise and motivate physiological, cognitive, and action patterns that facilitate adaptive responses to the vast array of demands and opportunities in the environment.”

Finally, the facial feedback hypothesis was introduced by Silvan Tomkins, which derives in part from Darwin (1872/1955) who claimed that the intensity of emotions could be influenced by enhancing or inhibiting facial expressions. William James (1894) also introduced the idea that visceral reactions and facial expressions provide sensory input, influencing emotional experience. Tomkins (1962) stated that one’s own face not only expresses affect to others but also to the self, via feedback from activity of facial muscles.

Few people would require instruction to understand the feelings of sadness, joy, desire, or anger. Emotions have evolved as a way in which humans can rapidly respond to events affecting their welfare. They occur in every relationship in our lives and can be triggered so rapidly as to precede conscious thought. The outward signs of emotions occur uniquely, most identifiably in the face and the voice (Ekman, 2007).

## FACIAL EXPRESSIONS OF EMOTION

*"I do believe thee; I saw his heart in his face"*

*Shakespeare (Winter's Tale).*

Charles Darwin, in his book *The Expression of the Emotions in Man and Animals* (1872/1955), argued that emotion expressions are adaptive and have evolved to serve a communicative function. Darwin's book has been highly influential in research on emotion, with citations recorded by the Institute for Scientific Information numbering almost 3,000 (Hess & Thibault, 2009). Darwin considered that facial expressions of emotions are external representations of an internal emotional state, and expression of these states is adaptive and contributes to an organism's chance of survival by organising behaviour in response to environmental demands. Tomkins (1962) also believed that the face has evolved to be increasingly more expressive due to increased visibility of the facial musculature in order to transmit affective information.

Facial expressions are one of the primary ways that emotions are expressed socially, signalling feelings to others which may influence their responses (Ekman, Freisen, & Ancoli, 1980). Negatively valenced expressions such as anger and fear, may signal impending danger; sadness may elicit empathy (Kunz, Peter, Huster, & Lautenbacher, 2013); and disgust can convey a warning not to approach a particular food as it may be contaminating (Blair, 2005). The expression of disgust may also be used to show objection to another's actions (Ekman, 2007). Using electromyography (EMG) to detect subtle facial movements, Chapman, Kim, Susskind, and Anderson (2009) tested the origins of moral disgust compared with disgust elicited by disease vectors and bad tastes. They found that all three disgust-evoking states activated the same facial muscles. Therefore, the expression of moral disgust triggered by social transgressions is indistinct from that evoked by bad tastes and potential toxins.

Although expression of emotions has communicative value, facial expressions also occur in solitary, private situations, being expressed involuntarily without conscious effort. Dimberg, Thunberg, and Elmehed (2000) demonstrated, using EMG, that it is possible to evoke distinct facial reactions without conscious awareness in response to positive and negative emotional displays.

## UNIVERSALITY OF EXPRESSIONS

Although the notion that basic emotions exist is ubiquitous in the emotion literature, not all theorists agree on what the basic emotions are, with some proposing as few as two emotions, whilst others suggest there are at least 18 (Ortony & Turner, 1990). Despite the divergence of opinion on what constitutes a basic emotion, evidence has been found to support the idea that there are at least seven basic emotions that are recognised universally: anger, fear, surprise, sadness, disgust, contempt, and happiness (Ekman & Cordaro, 2011). These have been established by pioneering research beginning in the late 1960s conducted with literate and pre-literate cultures (Ekman & Friesen, 1971; Ekman, Sorenson, & Friesen, 1969). For each of these emotions, a specific pattern of facial expressions has been observed. Earlier studies provided evidence of six universal expressions (anger, disgust, fear, happiness, sadness, and surprise), but more recent research has found universal expressions for the original six, as well as contempt, shame, and pride. Shame and pride differ from the other seven expressions in that in addition to facial expressions, they also involve changes in body posture and head positioning (Tracy, Randles, & Steckler, 2015).

The face is a commanding source of information, partly because it is always visible. The face commands attention because it symbolises the self, conveying messages about an individual (Cohn & Ekman, 2005), and is the primary source of outward expression of emotion. Emotion can also be expressed via the voice, gestures, and posture, but facial expressions are an aspect of human communication that provides a powerful way to express emotion. Although facial expressions can also be used to add emphasis to speech, such as raising of the eyebrows to emphasise a particular word (Ekman, 2007), the term “facial emotion expression” is used in this present study to describe the movements of facial muscles that spontaneously occur when an emotion is experienced, or if an expression is generated intentionally to communicate an emotion.

Early studies of facial expression began in the 19$^{th}$ century with Sir Charles Bell's interest in the physiology of facial expressions and the way in which these linked to emotions**.** The French neurologist, Duchenne de Boulogne (1862/1990), advanced the study of facial musculature using electrical stimulation. Following Bell and Duchenne came Charles Darwin's treatise on emotions, *The Expression of the Emotions in Man and Animals* (1872/1955). Darwin explored a number of questions related to facial expressions of emotion and was the first to suggest they occur universally, and also to question why particular expressions are formed in connection with certain emotions. For example, why do just the inner corners of our brows raise in sorrow instead of the whole brow? Are these expressions learned or are they innate? Are they the same regardless of culture, location, or language? Darwin proposed that there are distinct facial expressions that occur across cultures in response to the experience of specific emotions. He supported this proposition by gathering cross-cultural observations of expressions from 36 observers, mainly missionaries and traders, whose responses noted uniformity of expressions across various cultures, many of which had little or no communication with European cultures. However, his methodology used a questionnaire outlining the facial components he expected to occur thereby limiting the objectivity of the untrained observers, who also likely had a cultural bias. Also, without photographs of expressions, it is difficult to know what expression the observer actually perceived, and whether the supposed emotion matched the subjective experience of the expresser. Although Darwin's early studies lacked methodological rigor, his hypothesis inspired the genesis of research on facial expressions, as well as neuroanatomy and the biological and social functions of emotions (Izard, 1994). Following Darwin's work, the view that facial expressions of emotion occurred universally was challenged by anthropologists who proposed a culture-relativistic view that emotion-facial expression was culture dependent and expressions were learned through observation. The relativist approach viewed emotions as being embedded in socially constructed categories, mediated through language and culture, and defined by socially influenced judgements (Lutz & White, 1986). This view prevailed for some time.

In the early 20th century, studies of facial emotion expression were also hindered by research suggesting that the face did not accurately portray emotion. Landis (1924) exposed 25 participants (12 women, 12 men, and one boy, aged 13 years who suffered from high blood pressure and emotional instability) to emotionally evoking stimuli whilst being photographed.[1] The photographs were analysed and no correlation was found between facial expressions and subjective reports of emotion experienced by the participants. Landis concluded, therefore, that there was no specific facial expression characteristic of a particular emotional state. However, no objective measurement system was used and the emotion induction techniques were highly questionable.

After Landis's (1924) ethically dubious study, interest in researching facial expression declined just as Behaviourism was emerging, a school of thought which considered observable behaviour as the only legitimate 'data'. Interest in the study of emotion declined, and for many years the subject of facial expression and studies of emotion went unexplored. However, a resurgence of interest in emotion emerged in the 1960s with Silvan Tomkins reviving Darwin's theory about the evolutionary origins of facial expression. In one study, Tomkins and McCarter (1964) presented a range of facial affect pictures to participants who accurately identified emotions with above-chance accuracy. Ekman and Friesen (1971) then ventured into the study of cross-cultural facial expression. Initially they believed that facial expressions were culture specific, and so investigated the differences in facial expression across various cultures, expecting to disconfirm Darwin's theory. However, their findings were contrary to what they had expected. Ekman and Friesen studied remote tribes in New Guinea, the South Fore and the Dani. Both tribes were unexposed to Western

[1] Ethical issues were pretty much ignored in these times. As well as including a vulnerable child in the study, to evoke a strong emotional response, participants were shown pornographic photographs, a live rat having its head severed, were asked to put a hand into a bucket of live frogs, or were electrically shocked whilst trying to complete a mathematical task.

culture but were able to discriminate between happiness, sadness, disgust, and surprise, and to discriminate these from anger and fear. Although the South Fore did not distinguish between fear and surprise, the Dani tribe did. A similar study by Ekman et al. (1969) also found strong evidence to support Darwin's claim of universality in expressions. They found that facial expressions of emotion in images showing both Caucasian and Melanesian faces were recognised by people living in remote areas in Papua New Guinea, Borneo, Japan, and the United States.

A review by Russell (1994) notes research carried out in a variety of Western and non-Western cultures supports the claim that the facial expressions of basic emotions are fundamentally the same across all cultures. However, Russell criticised the validity of some of the methods used, such as biased response forms, forced-choice response formats, and posed rather than spontaneous expressions. These criticisms have been addressed in more recent research (Frank & Stennett, 2001; Matsumoto, Olide, Schug, Willingham, & Callan, 2009; Matsumoto & Willingham, 2009) with the results reconfirming support for the universality of facial emotional expression.

Frank and Stennett (2001) addressed the criticism regarding forced-choice formats by offering the escape option '*none of these terms are correct',* and four additional emotion labels. They found that, despite these modifications, agreement on emotion labels for universal expressions was still greater than chance.

A recent study by Matsumoto et al. (2009) found members of various cultures observing spontaneous facial expressions of athletes reliably judged the emotions expressed. Matsumoto and Willingham (2009) also compared the emotional expressions of congenitally blind and non-congenitally blind athletes at the 2004 Athens Paralympic Games with spontaneous expressions of sighted athletes at the 2004 Athens Olympic Games. Facial expressions of emotions for all athletes were the same and no cultural differences were noted. The fact that blind athletes spontaneously produced the same facial expressions as sighted athletes in response to the same emotionally evocative situations provides compelling support for expressions occurring as innate responses to emotions rather than through observational learning. Another study that lends support to the hypothesis that facial emotional expressions are innate rather than learned was conducted by Camras, Oster, Campos, Miyake, and Bradshaw (1992). Camras and her colleagues studied Japanese and American infants' responses to aversive stimuli and found no cultural differences in their expressions of negative facial affect.

Although there is strong evidence for the universal expression of emotions, culture can also influence the management of emotion expressions, which Ekman termed 'display rules'. For example, in one study, described by Ekman et al. (1987), participants from Japanese and American cultures watched an emotionally evocative film whilst alone and unaware of being filmed. The facial expressions emitted were the same, irrespective of culture. However, when a researcher was present, the Japanese were more likely than the Americans to mask negative expressions with smiles.

In contrast to Ekman et al.'s (1987) concept of display rules which involves deliberate manipulation of emotion expressions, cross-cultural differences in emotion recognition have been attributed to in-group familiarity. A meta-analysis by Elfenbein and Ambady (2002) of 168 data sets examining judgements of emotion within and across cultures found that emotions were universally recognised at levels well above chance. The degree of cross-cultural recognition varied across emotions, with fear and disgust being the most poorly recognised while recognition was highest for happiness. However, they also reported that there is an "in-group" advantage when emotions were both expressed and received by members of the same cultural group. This also varied according to the emotion expressed with an in-group advantage being highest for fear and disgust, and lowest for happiness and anger. Findings such as these suggest that there are key components of emotion expression which are universal and likely biologically based, but which are also modulated by culture and experience.

Notwithstanding the influence of culture, there is moderately strong evidence for the universality of at least seven facial expressions of emotions - anger, disgust, fear, happiness, sadness, contempt, and surprise. Each of these emotions is presented by distinct facial patterns, and are the ones used in the current study. A sample of each expression is shown in Figure 1.

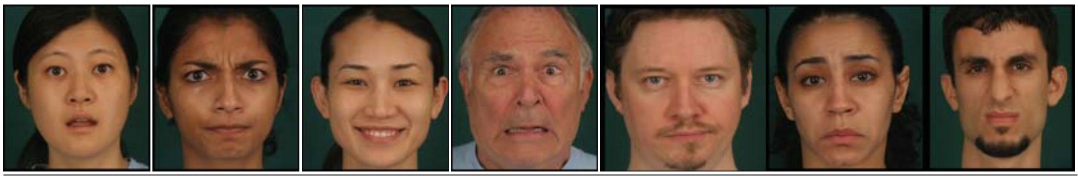


*Figure 1.* Sample of images from eSett 3.0 (http://www.paulekman.com). From left to right: Surprise, anger, happiness, fear, contempt, sadness, disgust.

In summary, although there have been challenges to the evidence of universal expressions (e.g., Lutz & White, 1986; Russell, 1994), there is a large body of research spanning the last four decades supporting the universal view (e.g., Camras et al., 1992; Ekman & Cordaro, 2011; Ekman & Friesen, 1971; Elfenbein & Ambady, 2002; Matsumoto et al., 2009). These studies have used a variety of approaches including judgements of posed expressions by both literate and preliterate cultures; judgements of spontaneous expressions within and across cultures; spontaneous expressions of blind individuals; and facial responses of infants. Many of the studies have been replicated several times and a number have also addressed the earlier criticisms levelled at cross-cultural research. For example, Ekman et al. (1987) replicated a number of earlier studies whilst addressing some of the shortcomings such as limited numbers of emotion terms; posed rather than spontaneous expressions; judgements of intensity level; and use of blended facial expressions. After tackling these limitations, it was found that agreement across cultures in interpretation of facial expressions of emotion remained consistent.

## EMOTION RECOGNITION ABILITIES (ERA)

This section provides an overview of research investigating the ERA of various populations. ERA has been theorised as being both a specific ability, i.e., a particular skill, as well as a component of an integrated approach to Emotional Intelligence (EI; Mayer, Roberts, & Barsade, 2008). EI directs attention to the role of emotions in many facets of life – at home, at school, and in the workplace. EI is described by Mayer, DiPaolo, and Salovey (1990, p. 772) as “the accurate appraisal and expression of emotions in oneself and others and the regulation of emotion in a way that enhances living”.

Outcome studies have found that EI is positively correlated with better social, occupational, and intimate relationships; being perceived more positively by others; greater academic achievement; and enhanced psychological well-being (Elfenbein, Marsh, & Ambady, 2002). EI involves recognising emotion expression in others and is strongly associated with empathy, described by Blair (2005) as being aware of and responding to the emotional displays of others. The ability to recognise facial expressions of emotion is considered a vital component of EI, and promotes interpersonal behaviour that improves social interactions (Elfenbein, Foo, White, Tan, & Aik, 2007). Mayer et al. (2008) describe several approaches to conceptualizing EI with one approach being a four-branch model of EI, where abilities from four areas integrate, with one of these branches being the accurate perception of emotion in others. This model views all EI abilities as interconnected with development of a skill in one area contributing to development of skills in other areas. For example, skill in recognising emotions from the facial expressions of others is likely to precede the ability to detect false emotion expressions. The other branches of the four-branch model include: using emotions to facilitate thought; understanding emotion; and emotion management. Another related conceptual approach to defining EI that Mayer and colleagues describe is the Specific-Ability approach, which focuses on particular skills that are considered fundamental to EI, with one of these abilities being the accurate perception of emotion from non-verbal information. This includes recognising emotion from facial expressions, gestures, vocal cues, and posture. Other specific abilities include: use of emotions to facilitate thinking and decision making; emotional reasoning and understanding; and emotional self-management and regulation.

## The Utility of ERA

It has been established that there are unique patterns of facial behaviour that signal particular emotions. However, this is only useful if the observer is able to infer the internal state of another by accurately recognising the distinct facial expressions that accompany particular emotions. However, like many skills, ability ranges from highly adept to less proficient, and this skill is more important in some occupational settings than in others. For example, in a therapeutic setting, the ability to recognise the emotional state of others from facial expression is an important factor in communication, and may be necessary to develop empathy and obtain valid and reliable information from clients (Machado, Beutler, & Greenberg, 1999). However, it may not be so important for people in occupations with less interpersonal involvement, a computer analyst, for instance. For therapists, accurate empathy was identified by Rogers (1957) as a prerequisite to establishing a therapeutic relationship with a client, and research on therapist empathy has identified it as possibly being related to positive outcomes in therapy (Cain, 2007). Seligman (2011) believes empathic accuracy is one of the basic therapeutic essentials along with warmth, genuineness, and rapport.

In therapy, attending to emotions is a vital component of emotion-based therapies involving emotional processing. These are used for treatment of fear reduction in anxiety disorders such as Post-Traumatic Stress Disorder. Emotional processing is defined by Foa and Kozak (1986) as what occurs when the memory structures underlying the fear are modified. It is therefore important for the therapist to be able to accurately observe and reflect a client’s feelings in order to assist them through this process to achieve positive outcomes. For example, Pos, Greenberg, Goldman, and Korman (2003) found that emotional processing skills as measured by the Experiencing (EXP) scale (Klein, Mathieu, Gendlin, & Kiesler, 1969) contributed to decreases in symptoms and increases in self-esteem in depressed clients.

An exploration of sex offenders’ emotion recognition skills and empathy was carried out by Hudson et al. (1988). The authors proposed that sexual offenders are deficient in empathy, which may enable them to violate others. As emotion recognition is considered to be a key component of empathy, it was hypothesized that offenders would have greater difficulty in recognising facial expressions of emotions relative to other individuals. The results confirmed that sex offenders were significantly less accurate than community controls at recognising emotional facial expressions. They also scored significantly lower than controls on an empathy measure – the Interpersonal Reactivity Index (IRI: Davis, 1983), supporting the link between this skill and the construct of empathy.

The ability to accurately identify peoples' underlying emotional state from their facial expression is a valuable skill for therapists for a number of reasons. Being able to accurately recognise another person's emotional state can evoke that emotion in oneself (Haxby, Hoffman, & Gobbini, 2002; Niedenthal, Brauer, Halberstadt, & Innes-Ker, 2001) which enhances communication and the development of a therapeutic relationship with a client (Machado et al., 1999). Additionally, clients conceal emotions in therapy for a wide range of reasons. These include fear of being judged, avoidance of painful or embarrassing information, maintaining a favourable self-image, and fear of mandatory reporting (Grohol, 2008). Being alert to emotional concealment would benefit therapy as it signals emotions the client is experiencing and consciously or unconsciously avoiding. Being able to discern when emotions are masked or concealed can contribute to therapy outcomes by alerting the therapist to areas that may need further exploration. For example, a depressed inpatient may conceal anguish with a mask of positive expression in order to expedite discharge in order to commit suicide (Ekman, 2009). In the context of risk assessments, clients may be deceptive in responding to questions concerning abuse in order to protect a perpetrator or to conceal their own intentions. In forensic settings, detecting deception in emotional expression may contribute to more accurate risk assessment and assist in detecting feigned motivation to attend rehabilitative programmes. As well, perceived credibility of an offender's remorse is a mitigating factor in sentencing decisions in New Zealand (Sentencing Act, 2002) and in Canada (Porter, ten Brinke, & Wallace, 2011). In child custody evaluations, obtaining accurate information from parents and children will be assisted if the clinician is alert to signs of deception and avoidance of painful or distressing emotions.

Emotional expressions can also assist in alerting therapists to information that may be painful to disclose verbally. For example, a study by Bonanno et al. (2002) found child sexual abuse (CSA) survivors who did not voluntarily disclose CSA showed greater facial expressions of shame, whereas CSA survivors who did disclose abuse expressed more disgust. Expressions of disgust also signalled violent abuse. Facial emotional expressions can also be helpful in predicting the course of grief in bereaved spouses (Bonanno & Keltner, 1997). Bereaved spouses, who showed more negative expressions, particularly anger (but also contempt and fear) six months after their loss, experienced more severe grief and poorer perceived health after 25 months. Interestingly, expressions of sadness were unrelated to outcome. In contrast, the spouses who displayed more positive expressions, particularly amusement, experienced less grief after 25 months. Gottman, Levenson, and Woodin (2001) studied the facial expressions of 79 married couples whilst engaging in marital conflict discussions. They found that facial expressions such as fear, feigned happiness, disgust, and contempt, significantly predicted marital outcomes such as separation and illness.

In summary, it has been established that the ability to recognise emotions in others is important in therapy to build empathy, which is considered fundamental for establishing a therapeutic relationship with a client, and has been linked to positive therapy outcomes. Emotional processing in therapy has also been associated with benefits for clients, and accurately observing and reflecting a client’s feelings can assist them in this process. There is also evidence that emotion expressions can provide information that is not disclosed verbally, as well as predict the course of therapy.

### Therapists

Despite the importance of facial emotion recognition abilities (ERA) for therapists, little research has been conducted to date on this topic. The only study located that investigated ERA of therapists is one by Machado et al. (1999) which explored whether level of training and experience enhanced judgements of both verbal and nonverbal expressions of emotion. Participants were therapists (*n*=36) compared with a control group of undergraduate students (*n*=36). The therapist group was comprised of 26 advanced doctoral students, five clinical or counselling psychologists, one psychiatrist, and four practicing Masters-level therapists.

In this study, representative segments of psychotherapy sessions were presented in one of three ways to separate the effects of verbal and non-verbal cues. The first was a video recording featuring both verbal and non-verbal cues. The second was the same video recording with audio content filtered out, and the third was a verbatim transcript of the therapy session, representing only lexical cues. Participants were randomly assigned to one of these conditions and completed ratings of emotional quality (the prevalent emotion observed in the client) and emotional intensity (modal and peak levels of emotional intensity exhibited by the client). It was found that experience increased accuracy in recognising the specific emotion expressed by the client but there were no differences in ratings of emotional intensity between the low and high experience groups. Interestingly, accuracy in the therapist group increased when verbal content was filtered out, whereas the lowest level of accuracy occurred in the undergraduate group for this condition. This suggests that for those with therapy experience, non-verbal clues increase the accuracy of emotion perception, and for those with less experience, there is a reliance on verbal information to determine the emotion expressed. Participants also completed an emotional awareness scale. Results showed that awareness of their own emotions was correlated positively with emotion identification but not emotion intensity. Although the two groups were comparable on demographics such as sex and ethnicity, there were differences in level of education (undergraduate students vs. professionals). Also, the mean age of the therapist group (34.80 years, *SD* = 6.49) was significantly higher than the undergraduate group (25.17 years, *SD* = 8.11). This presents a considerable confound to the variable of interest, years of experience, as in addition to the effects of experience, older adults may outperform their younger counterparts because of reasons unrelated to working as a therapist, such as practice effects. Presumably, being older, they would have accumulated more experience of social interaction with a wide range of individuals. Indeed, age-related differences in emotion recognition have been reported in a number of studies (e.g., Calder et al., 2003; Murphy, Lehrfeld, & Isaacowitz, 2010). Furthermore, the 'experienced' group only had an average of 5.60 years (*SD* = 4.09) of experience. A more experienced group may have provided more illuminating results.

A study by Hutchison and Gerstein (2012) looked at counselling trainees' ability to recognise facially expressed emotions. This study found no difference between trainees (*n*=54) and undergraduate students (*n*=54) in emotion recognition ability. Given the trainees had had limited experience in working with clients, this finding is unsurprising. Another explanation provided by the authors was that the study yielded high accuracy rates for both groups (approximately 82% for undergraduates and trainees); therefore, a ceiling effect may have occurred.

## Clinical and Non-Clinical Populations

Studies of therapists' ERA are limited, although there are large numbers of studies that have investigated ERA in a range of clinical populations including, but not limited to, autism spectrum disorders (Ashwin, Chapman, Colle, & Baron-Cohen, 2006), Alzheimer's disease (Henry et al., 2008), schizophrenia (Russell, Chu, & Phillips, 2006), sex offenders (Hudson et al., 1988), methamphetamine abusers (Kim, Kwon, & Chang, 2011), alcohol dependence (Maurage et al., 2011), Huntington's disease (Henley et al., 2012), Parkinson's disease (Herrera, Cuetos, & Rodríguez-Ferreiro, 2011), major depressive disorder (Dalili, Penton-Voak, Harmer, & Munafo, 2015), and acquired brain injury (Adolphs, Damasio, Tranel, & Damasio, 1996; Barod, Koff, Perlman Lorch, & Nicholas, 1986). All of these studies have reported impairments in ERA compared to healthy controls.

Studies have also examined ERA in non-clinical populations such as medical students (Endres & Laidlaw, 2009), salespeople and trial attorneys (Matsumoto & Hwang, 2011), sales negotiators (Elfenbein et al., 2007). Calder et al. (2003) and Murphy et al. (2010) investigated differences in ERA between young and older adults, and others have examined sex-related differences in emotion recognition (Hall, 1978; Hampson, van Anders, & Mullin, 2006; Hoffmann, Kessler, Eppel, Rukavina, & Traue, 2010; Sasson et al., 2010). These studies will be discussed in more detail in the following sections.

## Acquired Brain Injury and ERA

There is evidence to suggest that the right hemisphere (RH) shows some predominance for processing emotions. In a study examining ERA in patients with brain-damage, Barod et al. (1986) found those with damage to the RH were significantly impaired at ERA when compared with left hemisphere (LH) damaged patients and healthy controls. Adolphs et al. (1996) also investigated ERA in patients with focal brain damage and found impairments in ERA where lesions occurred in the RH, but no ERA impairment when lesions were restricted to the LH. A review by Najt, Bayer, and Hausmann (2013) reported hemispheric asymmetry in processing positive and negative expressions, with a RH advantage for recognition of anger, fear, and sadness, and a LH advantage for recognising happy expressions. More recently, using fMRI, Srinivasan, Golomb, and Martinez (2016) report that the posterior superior temporal sulcus (pSTS) in the right side of the brain may be specialised for recognising and decoding Action Units (AUs) of emotion expressions.

## ERA Differences Across the Adult Lifespan

An investigation of ERA across the adult lifespan was conducted by Calder et al. (2003). Participants aged between 20 and 75 years were divided into five age groups (20-30, 31-40, 41-50, 51-60, and 61-75 years). It was found that from around the age of 40 years, people began to demonstrate a decrease in the ability to recognise facial expressions of fear. Also, but to a lesser extent, recognition of sadness and anger decreased with increasing age. However, older age was associated with improved ability in recognising disgust. Using a dynamic emotion recognition task which displayed morphed emotional facial expressions at various levels of intensity, West et al. (2012) examined ERA in 482 people ranging in age from 20 to 89 years of age. They reported similar findings with age-related declines observed for recognition of fear, anger, and sadness. However, in their study, these declines first emerged in the 60s, declining further in the 70s and 80s. West et al. found no difference across ages for recognition of happiness or surprise, and, similar to Calder et al., recognition of disgust was observed to improve with increasing age. Suzuki, Hoshino, Shigemasu, and Kawamura (2007), compared younger adults (*M* = 20.6 years, *SD* = 1.8) with older adults (*M* = 69.7 years, *SD* = 4.8). They found a significant age-related decline in sadness recognition, marginal declines for surprise and anger recognition, no differences for happiness and fear, but significantly improved disgust recognition with increasing age. However, the authors stated this may be due not to older adults being more sensitive to disgust, but, instead, a result of younger people being more likely to mistake disgust for anger. They speculated this may be due to younger people being more sensitive to the emotion of anger. Henry et al. (2008) examined ERA in a sample of people with relatively mild Alzheimer's disease and reported significant declines in recognition of all emotions relative to controls, with the exception of disgust, which remained intact.

A summary of studies focusing on healthy, community-dwelling people reported by Isaacowitz et al. (2007) also suggests age-related decline in recognition of negative emotions, with the exception of disgust. In 11 out of the 13 studies reviewed, recognition of happiness remained stable, although it was noted that many of the studies showed ceiling effects with regard to recognition of this emotion due to the ease of the task. However, in the study conducted by Isaacowitz et al. on three groups of adults (Young, *M* = 27.05 years; Middle-aged, *M* = 48.01 years; Older, *M* = 71.90 years) they found significant age differences in recognition of anger and fear from facial expressions, but no differences in recognition of facial expressions of sadness. Contrary to previous research, they also found age-related differences in recognition of happiness, suggesting that previous research reporting no differences may have been affected by ceiling effects. Murphy et al. (2010) investigated age-related differences in recognition of posed and spontaneous smiles. Although there was no difference between groups in the recognition of expressions in young people, older adults were better able to discriminate between smile types when targets were a mix of young and older adults. Demenescu, Mathiak, and Mathiak (2014) studied ERA in young, middle-aged, and older adults. They found that older adults were less accurate, irrespective of emotion type, than younger adults in recognising emotion. They also found that middle-aged females performed better relative to males. A meta-analytic review of ERA in younger (< 45 years) and older (> 55 years) adults was conducted by Ruffman, Henry, Livingstone, and Phillips (2008). This review noted a predominant pattern of age-related decline in the recognition of facial expressions, including happiness. However, the effect sizes for age differences in happiness and surprise recognition were much smaller than those found for sadness, anger, and fear. It was reported that younger adults found sadness easiest to identify, whereas for older people, this appeared most difficult. Consistent with Calder et al. and Suzuki et al., the review reported that older adults appear to be better than their younger counterparts in recognising the expression of disgust, although the effect size was not statistically significant. All of the studies investigating the impact of age on facial emotion recognition, except for two, used static images of facial affect.

In summary, it appears that recognition of facial emotion expressions shows some decline with increasing age, although most studies have found an improvement in the recognition of disgust. There are some inconsistencies reported, but this may be due to the different age groups being investigated, with some studies comparing young with middle-aged adults, whilst others had a larger mean age difference.

One possible reason for age-related decline in emotion recognition is the habituation to emotion as we age. Therefore, older people may be less attuned and pay less attention to emotions as they are feeling emotions less intensely themselves (S. T.

Harvey, personal communication, 26 May 2016). For instance, Calder et al. (2003) reported that older adults experience and express fewer negative emotions, but not positive emotions. It is suggested that the same factors which to decreased experience/expression, may account for reduced processing of certain emotions. However, a study by Phillips, Henry, Hosie, and Milne (2008), which investigated the regulation of expression and experience of negative affect in old age, found that older adults subjectively reported higher levels of emotional reactivity and expressed higher levels of negative affect when compared with their younger counterparts. Despite this, the older adults were more effective at using emotion-regulation strategies to reduce the experience of negative emotions resulting in decreased expression. It is possible that development of an automatic practice of utilising emotion-regulation strategies to reduce negative affect may influence the recognition of emotions in others.

Another possible reason for age-related changes is presented by Wong, Cronin-Golomb, and Neargarder (2005). They propose that age-related changes in visual scanning of faces may account for older adults' deficits in emotion recognition ability. Their study found that older adults' deficits in recognizing anger, fear, and sadness, were significantly correlated with fewer gaze fixations to the top region of the face. It has previously been suggested that correct recognition of these emotions requires visual scanning patterns which include fixations on the upper halves of expressive faces (Calder et al., 2003). Therefore, Wong, et al. propose that age-related changes in visual scanning patterns may contribute to emotion recognition deficits in older adults.

Finally, Ruffman et al. (2008) propose what accounts for changes in emotion recognition abilities with age are changes in brain structure, with declines in the orbitofrontal cortex, cingulate cortex, and the amygdala, contributing to reduced recognition of anger, sadness, and fear. The authors also propose that older adults' relative strength in recognising disgust may be related to preservation of the basal ganglia.

## Sex Differences

The term sex differences, rather than gender differences, is used here. This is because the literature, with a few exceptions, refers to the differential ability between men and women at recognising emotions as sex differences. Those that have used the term *gender* (e.g., Hall, 1978) have used this interchangeably with sex, when referring to males and females. This is possibly because Hall's 1978 meta-analysis surveyed research conducted between the 1920s to the 1970s, a time before gender became a term to describe a social construct rather than biological sex. ERA between sexes have been investigated with mixed results. An early study by Guilford (1929) reported no sex differences in facial expression recognition, albeit with a small sample of eight women and seven men. Guilford also referred to an earlier study by Allport who found no difference between sexes in judgement accuracy but did find that women responded twice as quickly as the men, leading to the inference that women's judgements are made more intuitively. Grimshaw et al. (2004) also found no sex differences for perception of emotional faces using signal-detection analysis.

Meta-analytic studies have reported significant but small effect sizes for a female advantage in emotion recognition (Hall, 1978; Thompson & Voyer, 2014). Hampson et al. (2006) also reported a female advantage in reaction times responding to visual facial expressions, but due to ceiling effects, no differences between men and women could be assessed for accuracy of emotion recognition. Likewise, Rahman, Wilson, and Abrahams (2004) found women had faster reaction times than men but there were no overall differences in accuracy of emotion recognition. An early study by Zuckerman, Lipets, Koivumaki, and Rosenthal (1975) reported that females were significantly better than males at recognising facial expressions. It was suggested by the authors that this may be due to women being more socialised to perceive emotions than men. The study also found that for males, but not females, familiarity with the expresser increased accuracy; however, it was possible that this finding was an artifact caused by a female accuracy ceiling effect. Regarding intensity of facial expressions, Hoffmann et al. (2010) found women were better than men at accurately recognising subtle facial expressions but no differences were found when recognising full-intensity facial expressions.

In a meta-analysis by Ickes, Gesn, and Graham (2000), it was found that women outperformed men in tasks involving empathic accuracy only in situations when they were motivated to present themselves as being highly empathic, suggesting differential motivation accounts for sex differences, rather than differential ability.

Sex differences in people with schizophrenia have shown biases in males towards misinterpreting neutral faces as angry, while females misinterpret neutral faces as being sad (Marsh, Luckett, Russell, Coltheart, & Green, 2012).

In summary, meta-analyses have found significant, but small effect sizes in favour of women's ERA, particularly with regard to negative emotions. A possible explanation for this is offered by Hall (1978) who suggests differences in emotion recognition could be due to biological factors resulting from evolution. Nonverbal sensitivity in mothers may make them more sensitive and responsive to distress in their babies, which would enable them to interpret threat from other adults more readily. This would contribute to enhanced survival of their offspring. Hall also discusses how sex differences may be due to gender stereotypes which promote women as caregivers, a role which requires one to be nurturing and is likely to be assisted by sensitivity to non-verbal clues. Although babies are pre-verbal, from an early age they are able to produce a variety of facial emotion expressions, so if women are culturally and socially encouraged to become caregivers, they are more likely to develop better emotion recognition skills to enable them to communicate with non-verbal infants. Being sensitive to facial emotion expression is therefore likely to enhance the caregiving relationship.

## ERA and Negotiation Skills

Elfenbein et al. (2007) examined the relationship between ERA and an objective measure of a goal-oriented interpersonal interaction. As negotiation can be a highly emotional process involving a range of emotions such as pleasure, surprise, fear and anger, Elfenbein et al. explored whether success in negotiating a sales transaction is enhanced by ERA using a measure where participants identified emotion expressions from still photographs. Participants were randomly assigned to be either buyer or seller in a negotiation exercise. It was found that there was a positive correlation between ERA and successful negotiation skills for participants who assumed the role of sellers, but not for buyers. The authors speculated that this effect of ERA on only one member of the dyad was due to the buyers having more control over the negotiation process and therefore they would have had more opportunity to read their counterparts whilst bargaining.

In summary, it appears that a number of factors, such as age, vocational experience, mental illness, and possibly gender, can influence a person's ability to recognise facial expressions. Although ERA have been studied in various populations as outlined above, in order to accurately study the recognition of emotion, a measurement system that objectively and consistently distinguishes between various facial expressions is necessary. The most frequently used systems will now be explored.

## MEASUREMENT OF FACIAL EXPRESSIONS

The following section provides an overview of several ways in which facial behaviour has been objectively measured. These measurement systems examine the muscular activity underlying a facial expression, as well as the intensity of the expression. They also enable the temporal dynamics of expressions to be assessed, which can be useful for distinguishing between genuine and posed facial expressions.

### Electromyography (EMG)

Since the 1970s facial electromyography (EMG) has been applied to the study of facial behaviour (Ekman & Rosenberg, 2005). EMG involves placing surface electrodes on the skin in order to measure nerve action potentials associated with muscular contractions. This is a sensitive measure able to detect muscular activity that is otherwise unobservable. EMG measures have shown that patterns of facial muscle activity reliably accompany the experience of specific emotional states, even when there is little or no movement of facial skin (Rinn, 1984). Recent studies of facial emotion expressions using EMG have included investigating distinct expressions for schadenfreude (Boecker, Likowski, Pauli, & Weyers, 2015), patterns of facial reactions to stimuli showing expressions in people with borderline personality disorder (Matzke, Herpertz, Berger, Fleischer, & Domes, 2014), and facial mimicry responses in children (Deschamps, Schutte, Kenemans, Matthys, & Schutter, 2012). Dimberg, Andréasson, and Thunberg (2011) used EMG to compare the facial reactions of people high in emotion empathy with those low in empathy when presented with happy and angry facial expressions. EMG demonstrated that people high in emotion empathy react with stronger facial expressions that mimic the expression observed. It is proposed that facial mimicry is a key component of empathy as the activation of facial muscles functions as a feedback system for the person's own experience of emotion. This is also consistent with the facial feedback hypothesis (Tomkins, 1962), which suggests that facial emotional expressions play a causal role in subjective experiences of emotion.

Although EMG is highly sensitive and able to detect facial muscle activity that is unobservable to the naked eye, it also has several limitations. For example, it is obtrusive and may increase self-conscious behaviour; and cross-talk between muscle action potentials can occur, interfering with the signal and possibly leading to misinterpretation of the expression (Ekman & Rosenberg, 2005).

As well as EMG, observational coding systems have been used to study facial behavior such as the Maximally discriminative facial movement coding system (MAX), the Facial Affect Scoring Technique (FAST), and the most widely used, the Facial Action Coding System (FACS: Ekman et al., 1978/2002). Although MAX is an anatomically-based system, it was developed to include only facial displays related to emotion, and is less comprehensive than FACS. FACS scoring and EMG measures of intensity are highly correlated (Pearson's $r$ = .85) (Ekman, Hager, & Friesen, 1981). Although EMG may be more sensitive to low level changes in muscular activity, because surface EMG is imprecise, it cannot always distinguish the wide range of facial expressions that FACS can. Furthermore, as mentioned previously, EMG is more obtrusive which can affect participants' awareness of what is being measured, which could lead to modification of expressions (Hager & Ekman, 2005).

## The Facial Action Coding System (FACS)

FACS (Ekman et al., 1978/2002) is a facial measurement system originally developed by Ekman and Friesen in 1976. The system was developed to measure facial movement based on anatomical analysis of expressions. FACS represents a breakthrough in the study of facial expression as previously, research on facial expressions relied on observers' inferences of expressions, which were subject to the influence of contextual knowledge, cultural interpretation, and other information such as the voice. The original development of FACS relied on motion picture film and still photographs. Ekman et al. (1978/2002) have since updated the FACS system. FACS 2002 includes significant improvements in scoring criteria, and didactic materials, utilising advancements in technology such as embedded video links and hyperlinked cross-referenced text (Cohn & Ekman, 2005). FACS measures individual facial movements from distinct momentary expressions, deconstructing them into specific Action Units (AUs), representing muscular changes in facial expression. FACS distinguishes 44 AUs, each of which is anatomically separate and visually distinguishable (Ekman et al., 1980). Individual AUs do not signal specific emotions; rather, combinations of AUs represent emotions and the FACS manual lists what combinations of AUs are most frequently observed for common emotional expressions. Intensity of AUs is scored on a scale ranging from A to E, with A referring to trace evidence of the expression, and E denoting maximum expression (Ekman et al., 1978/2002). Trained observers can detect subtle changes in facial expressions from slow motion video-recording of facial behaviour and code all facial displays as AUs (Kanade, Cohn, & Tian, 2000). Although the muscular basis of AUs is well established, AUs do not correspond precisely to specific facial muscles, but are strictly defined by the observable movements of the facial skin (Rinn, 1984).

A system to assess inter-rater reliability was developed by Wexler (1972, cited in Ekman, Friesen, & Simons, 1985) and is calculated according to the following formula: Total number of AUs that both coders agree on multiplied by two, divided by the total number of AUs scored by the two coders. For example, if an image was coded as 1 + 4 + 11 + 15 by one coder and as 1 + 4 + 15 by the other, the agreement index would be 3 x 2/7 = 0.86. The intensity of the AUs is not included in the calculation; the focus is only on whether the AU is active or not (Lewinski, den Uyl, & Butler, 2014). To achieve certification as a coder, manual FACS coders must reach an agreement index of 0.70 with the master coding on the final test.

There are several studies that have demonstrated the validity of FACS for measuring spontaneous emotional expressions. FACS coding has provided accurate information about subjective experiences of emotion, and is able to discriminate the intensity of the emotion experienced (Ekman et al., 1980), confirming coherence between expressed and subjective experiences of emotion.

Two reviews of FACS have been published in The Buros Mental Measurements Yearbook (Benes, 1995; Donlon, 1995). Benes (1995) described FACS as a 'highly complex research tool' with remarkable overall reliability scores. Inter-rater reliability was reported to be 0.76 and measures of intensity correlated highly with EMG readings (Pearson's $r$ = .85). Donlon's (1995) review described FACS as a highly comprehensive taxonomic system relevant for coding facial expressions.

A lot of the research of facial emotion expression using FACS has involved posed expressions. Spontaneous expressions are believed to differ from posed expressions in symmetry and velocity (Ekman et al., 1981). So to examine the reliability of FACS when studying dynamic expressions, an evaluation of the reliability of FACS was conducted by Sayette, Cohn, Wertz, Perrott, and Parrott (2001). Their research evaluated inter-observer reliability of FACS using video-taped spontaneous emotion expressions obtained through three different emotion eliciting procedures. Facial behaviour was independently coded by two FACS-certified coders. Inter-observer agreement was quantified with coefficient kappa (*k)*, with coefficients of 0.60 to 0.75

reflecting good reliability, and coefficients of 0.75 and above indicating excellent reliability. Sayette and colleagues reported 'good to excellent reliability' between coders for measuring spontaneous emotional expressions for all but two AUs, which showed only fair reliability ($k$ = 0.53, 0.56). As posed emotional expressions differ from spontaneous expressions in terms of the action units (AUs) involved, the difference between genuine and posed expressions can be distinguished using FACS.

FACS has since been used in various studies and is considered to be the gold standard in facial expression research (Kunz et al., 2013). Kunz et al. used FACS to study facial expressions of pain and disgust. They found that although pain and disgust experiences elicited the same AUs, clear differences were found between pain and disgust with the intensity varying for different AUs. In pain, the contraction of muscles around the eyes is more intense, whereas in disgust, the muscles around the eyebrows are more strongly contracted and the upper lip and nose wrinkling occurs more with disgust than with pain. In a study of whether FACS could discriminate levels of pain in people with Alzheimer's Disease who may have difficulty communicating pain due to cognitive impairment, it was reported that differences in facial expressions were associated with varying levels of pain. FACS was found to be sensitive to detecting expressions of low levels of pain, enabling assessment to occur without relying solely on self-report, which may be difficult for people with Alzheimer's Disease (Lints-Martindale, Hadjistavropoulos, Barber, & Gibson, 2007).

FACS has also been found to accurately identify facial expressions in people with Rett's Syndrome, a neurodevelopmental disorder leading to communication impairment and the inability to use expressive language. FACS was able to distinguish between facial expressions that stem from emotions and those caused by abnormal brain activation due to Rett's Syndrome. This distinction is useful as, due to communication difficulties, carers rely on facial expressions to interpret clients' needs, and without this knowledge may incorrectly interpret these expressions (Bergstrom-Isacsson, Lagerkvist, Holck, & Gold, 2013).

FACS was used by Archinard, Haynal-Reymond, and Heller (2000) to analyse the differences in facial expressions in patients admitted to hospital after a suicide attempt. The patients were video-recorded during an interview with a psychiatrist followed by coding of their facial behaviour using FACS. The facial behaviour of patients who had a repeated suicide attempt was compared with non-repeaters. Repeaters showed more activity in the upper face, particularly frowning, and had more activity of the mouth when not speaking, possibly due to efforts to control emotion. They also looked downwards significantly more often than non-repeaters. Being alert to these non-verbal indicators could provide important information to therapists relating to the risk of suicide in patients.

In summary, since its inception in 1978, FACS has been utilised in a wide range of studies concerning facial expressions and has been established as a reliable and objective system for measuring facial behaviour.

## Automated Facial Recognition Systems

As manual systems such as FACS require extensive training and are labour intensive, requiring viewing of videotapes in slow motion in order to code facial displays, automated facial recognition systems are being developed and are beginning to show promise in their accuracy. Lewinski et al. (2014) reported that their FaceReader software had a ratio index agreement with FACS of 0.67. This is almost at the level of human coders. Cohn, Zlochower, Lien, and Kanade (1999) have also developed an automated face analysis method that discriminates FACS AUs. In an analysis of 872 video samples of 15 AUs or AU combinations, Cohn and colleagues reported their software demonstrated high concurrent validity with manual coding of AUs FACS coding. Average recognition accuracy ranged between 81-91 %, depending on which AUs were coded.

The above section has outlined the ways in which facial expressions can be most objectively measured. Prior to the development of FACS, research on facial expressions focused on inferences made by observers without closely examining what was occurring on the face itself. FACS provides a more bias-free, objective examination of facial behaviour than observational approaches, using a standardized measurement system. This has enabled identification of the differences in facial behaviour that occur when one simulates an emotion expression or masks a feeling with an unfelt emotion compared with authentic, spontaneously occurring emotional expressions. For this reason, FACS has been used in the present study to objectively categorise facial emotion expressions and to differentiate genuine from masked emotion expressions.

## TRAINING IN EMOTION RECOGNITION

Individuals who are skilled in emotion recognition presumably are more able to discern a person's emotional state, and from that expression, possibly predict future behaviours. This skill has been associated with enhancements in workplace effectiveness in business executives, improved leadership skills in school principals, better outcomes for therapy clients, higher ratings from patients in medical settings, and improved academic performance (Elfenbein et al., 2007). More recently, there has been a shift from viewing ERA as a personality trait to a tangible skill, leading to an interest in whether or not training can enhance the skills of emotion recognition.

As facial expressions of emotion do not always occur in a uniform, prototypical way every time an emotion is elicited, leakage of concealed emotions can occur as a brief micro-expression, or a partial expression. Many of the studies of training in ERA have focused on improving the ability to detect brief micro-expressions or subtle emotion expressions.

In this section, the literature concerning training to enhance ERA is outlined, beginning with studies involving non-clinical populations, followed by a review of some of the studies using clinical populations.

## Non-Clinical Populations

Endres and Laidlaw (2009) explored the abilities of first-year medical students (*N*=24) to detect emotional micro-expressions before and after training with the Micro-Expression Training Tool (METT) developed by Paul Ekman (http://www.paulekman.com). This training involved participants watching videos that show commonly confused emotion pairs (anger/disgust, contempt/happiness, fear/surprise, fear/sadness) in slow motion with attention drawn to important features of each expression. During practice with feedback provided, micro-expressions lasting 1/25 of a second were then displayed and participants required to label them. Feedback concerning accuracy was given (*right* or *wrong* displayed on-screen along with the cumulative percentage correct). In this study, the students who scored high on a pre-test of communication skills improved their ability to detect micro-expressions after training compared to students who scored low on the communication skills test. Although this study looked at medical rather than psychology students, it is relevant to the current study as benefit from training for some students was demonstrated, in particular, those who had previously been assessed as good communicators.

Matsumoto and Hwang (2011) examined the effect of training on 81 department store employees. Using a training group versus a no-training comparison group, it was found that training significantly improved staff members' ability to recognise micro-expressions compared with the untrained group. The micro-expression training was provided using the Microexpression Recognition Training Tool (MiX) (http://www.humintell.com). Two weeks after training, outcome measures using third party ratings found training had also improved social and communication skills. In a second study, Matsumoto and Hwang investigated training in a group of practicing trial consultants (*N*=25), mainly psychologists and attorneys. Training improved both the accuracy and response time in reading micro-expressions, and this improved ability was still evident two to three weeks after training when compared with a group (*N*=30) who received no training.

## Clinical Populations

In clinical populations, training in emotion recognition has been studied to remedy deficits that can occur in various disorders such as schizophrenia, intellectual disability, traumatic brain injury, and autism spectrum disorder.

As schizophrenia can result in impaired recognition of facial expressions, which in turn contributes to difficulties with interpersonal communication and social functioning, Russell et al. (2006) trained 20 outpatients with the METT. All patients had been diagnosed with schizophrenia by a consultant psychiatrist, and all except one were on stable doses of atypical antipsychotic medication. Pre- and post-training scores were compared with those of age-, education- and gender-matched healthy controls. The authors concluded that training with the METT improved the patients' abilities in emotion recognition to a level indistinguishable from the untrained healthy controls. However, no follow-up assessment was undertaken to determine whether the effect of training was retained.

Marsh et al. (2010) replicated an earlier study by Russell, Green, Simpson, and Coltheart (2008) using the METT to train people diagnosed with schizophrenia (n=31) or schizoaffective disorder (n=8). As was found in Russell et al.'s study, Marsh and colleagues found an overall improvement in recognition of micro-expressions after training, which was maintained at one-month follow-up. Interestingly, improved recognition of more difficult face stimuli not used in training was not evident immediately after training, but recognition had improved at the one-month follow up, suggesting opportunities to practice newly learned skills enhanced recognition of subtle and dynamic emotional expressions. Although training was found to improve emotion recognition ability in these studies, further research is required to determine if this training contributes to increased social competence.

A review of four published studies on the effects of emotion recognition training for people with intellectual disabilities was undertaken by Wood and Kroese (2007). The training methods used in the studies ranged from labeling affect displayed in photographs (Ekman & Friesen, 1975, 1976), viewing soap opera video clips, and labeling affect on line drawings. It was reported that, although improvements in skills were made and maintained for as long as 8-9 months, a generalisation of skills to improved social functioning was not always evident.

Overall, in both clinical and non-clinical populations, studies have found that training in emotion recognition improves ERA which has also been linked to better social and communication skills. With regard to some of the clinical populations, training appeared to remedy deficits in ERA that had occurred due to mental illness, such as in schizophrenia (e.g., Russell et al., 2006). In one of the studies (Marsh et al., 2010), it was also found that training generalised to improved recognition of subtle and dynamic expressions.

## SUMMARY OF EMOTION RESEARCH

The face contains a complex musculature that interacts directly with the brain during affective states to produce specific patterns of facial behaviour linked to discrete emotions. Researchers have argued that this ability has evolved in order to communicate these emotions to others (e.g., Darwin, 1872/1955; Ekman et al., 1980; Tomkins, 1962). There has also been considerable support for the notion of universality of facial expressions (Ekman & Friesen, 1971; Ekman et al., 1969; Elfenbein & Ambady, 2002; Frank & Stennett, 2001; Matsumoto & Willingham, 2009) compared with only a small number who claim expressions are cultural artefacts (Landis, 1924; Lutz & White, 1986; Russell, 1994). The ability to recognise emotions in others has been theorised as being a fundamental component of Emotional Intelligence (EI), which in turn has been positively correlated with enhanced social, occupational, and psychological wellbeing. Given emotions in therapy contexts are likely to signal significant events and changes, the ability of a therapist to accurately make inferences about a client's emotional state may provide guidance for timing of application of therapeutic techniques, and hence contribute to improved outcomes and rapport.

The development of objective systems for measuring facial behaviour such as EMG, FACS, and more recently, automated systems, has improved the methodological rigor of studies of facial expression.  These are based on anatomical analysis of expressions, and are not reliant on subjective inferences made by observers, though both FACS and EMG require subjective interpretation of the data and are therefore prone to error.  A number of studies have demonstrated the validity of FACS, the system used in the current study, to consistently categorise the facial expressions presented to participants.

Given the importance attributed to accurate emotion recognition, a number of studies have investigated whether this skill can be developed by training.  Many of the studies focusing on both clinical (e.g., Marsh et al., 2012; Russell et al., 2006; Russell et al., 2008) and non-clinical (e.g., Endres & Laidlaw, 2009; Matsumoto & Hwang, 2011) populations have demonstrated the efficacy of training, reporting that emotion recognition can be improved with targeted training that draws attention to features of facial expressions of emotion.

Research into ERA also needs to be mindful of the impact of age and experience as it has been identified that recognition of some emotions does decline with increasing age (Calder et al., 2003; Demenescu et al., 2014; Suzuki et al., 2007).  To a lesser extent, sex differences have also been investigated with inconsistent findings reported.  At best, studies have demonstrated only a slim female advantage at decoding emotion expressions.

# CHAPTER TWO: DECEPTION

"False face must hide what the false heart doth know."

*Shakespeare (Macbeth)*

Deception is ubiquitous in human interactions and people admit to lying an average of two times a day (DePaulo, Kirkendol, Kashy, Wyer, & Epstein, 1996). Therefore, it is not surprising that deception will also occur within the context of psychotherapy. This can be psychologically costly, hinder treatment outcomes, and interfere with the therapeutic process (Curtis & Hart, 2015). The advantages of being able to recognise people's emotional states from their facial expressions have been outlined in the previous chapter. However, control over facial muscles can be used to manipulate observers in social interactions (Mehu, Mortillaro, Banziger, & Scherer, 2012). In different contexts, people conceal their true emotions for a range of reasons. Being able to distinguish between genuine and posed expressions and able to detect when emotions are being concealed is important in a number of contexts. In a therapeutic context, such concealment can alert the therapist to a client experiencing internal conflict. Clients may be emotionally deceptive for a variety of reasons. They may be avoiding a painful emotion; wishing to keep their inner experiences private; or displaying a more 'appropriate' emotion out of courtesy.

This chapter begins by describing the neurobiological basis for both voluntary and involuntary facial expressions of emotion. Darwin's *Inhibition Hypothesis* is introduced followed by a brief summary of research supporting it. Elements of facial behaviour that can occur during deception are described, such as felt and false smiles, and spontaneous and masked emotional expressions. Darwin's hypothesis is developed further by Ekman who proposes a theory of *reliable facial muscles* that can distinguish between posed and spontaneous expressions. Research to support this theory is then described. Linking to the inhibition hypothesis are subtle expressions and micro-expressions. Studies investigating the occurrence of these expressions are presented and how these relate to deception through leakage of underlying emotion expressions occurring. Next, a brief summary of the literature pertaining to verbal and emotional deception is considered. The chapter concludes with a survey of individual factors that may contribute to accuracy in deception detection.

## Neurobiological Basis of Emotional Expressions

A neurobiological basis has been established for differences between voluntary and involuntary facial expressions. As described by Rinn (1984), impulses for voluntarily induced facial movements have been shown to emanate from the cortical motor strip and course through the pyramidal tract to the facial nucleus. For facial motor movements arising from emotional experience, the impulses are active in a phylogenetically older motor system, known as the extrapyramidal motor system. Evidence for this distinction comes from patients who have experienced lesions of the cortical motor strip and who are no longer able to contract facial muscles on command, such as a smile, but are able to smile bilaterally in response to amusement, even though the same muscles are involved. Conversely, patients with brain lesions which compromise functioning of parts of the extrapyramidal motor system, such as the basal ganglia, can activate their facial muscles on command but no longer express emotions spontaneously, for example, in Parkinson's disease (Rinn, 1984).

Additionally, studies using EMG (e.g., Dimberg, 1997; Dimberg & Thunberg, 1998; Dimberg et al., 2000) have consistently found that responses to emotional stimuli occur rapidly and automatically. Dimberg, Thunberg, and Grunedal (2002) found that when presented with emotionally evoking stimuli, people could not avoid producing emotional facial expressions, even when instructed not to react, thus supporting the theory of involuntary control of some facial responses.

The neurobiological basis of facial expressions is relevant to studying leakage of emotions, particularly if masking or concealment is attempted during the experience of strong emotion. It is likely, in accordance with Darwin's inhibition hypothesis, that those facial muscles under least voluntary control will be most likely to leak an expression of the emotion being experienced. Conversely, activation of muscles to fake an emotion without the underlying emotional experience can be difficult, particularly in the upper face region. For example, an expression of surprise in the upper face is likely to occur during attempts to express falsified sadness, as the "grief muscles" of the forehead have reduced cortical connectivity and therefore are under less voluntary control (ten Brinke, Porter, & Baker, 2012).

## The Inhibition Hypothesis

Although Darwin did not specifically focus on deception, he hypothesised that some facial expressions of emotion cannot be consciously inhibited, particularly when the experience of the emotion is strong (Ekman, 2003). Darwin (1872/1955) described how Duchenne, the French anatomist, illustrated that certain facial muscles can be consciously controlled, whilst others cannot be. Discussing Duchenne's work Darwin stated, "No one has more carefully studied the contraction of each separate muscle, and the consequent furrows produced on the skin. He has also, and this is a very important service, shown which muscles are least under the separate control of the will" (1872/1955, p. 5). Although inspirational, Duchenne's observations were limited to just a small sample of participants; therefore, external validity is limited (Gosselin, Perron, & Beaupre, 2010). Darwin posited that muscles that cannot be consciously inhibited also cannot be intentionally activated without genuine emotion. These two propositions collectively form the *inhibition hypothesis* (Ekman, 2003).

## Inconsistent and Leaked Expressions

There is potential for the face to reveal secretly held emotions through leakage of the underlying emotion occurring, despite attempts to suppress or conceal it. Also, through covering an expression by fabricating an unfelt emotion, for instance, attempting to smile whilst feeling sad, there can be evidence of activity of facial muscles that are inconsistent with a genuine expression.

The presence of inconsistent emotional expressions was examined by Porter and ten Brinke (2008). Participants were video-recorded whilst viewing emotionally evocative images and were asked to respond with either a genuine or deceptive expression. An analysis of the video recordings (697 expressions in total) found that inconsistent emotional leakage occurred in all of the participants at least once during deception, and occurred more frequently in masked than in genuine expressions, varying according to the falsified emotion. However, no difference was found between genuine neutral expressions and intentionally neutralised expressions of felt emotion, therefore only partially supporting the inhibition hypothesis. Masking an emotion with another, unfelt, expression is far more complex than simply neutralising an emotion, as it involves generating an artificial expression whilst simultaneously concealing a felt emotion. Negative expressions were more difficult to suppress during masked expressions compared with the positive expression of happiness. The authors propose this is due to people having more experience with simulating happy expressions in daily life. It may also be due to negative emotions involving muscles that are under reduced volitional control.

The inconsistencies in expressions as outlined above have also been reported by Porter et al. (2011) who investigated the nature of facial expressions in genuine and deceptive displays of emotion. They found that inconsistent emotional leakage was ubiquitous, occurring in 98.3% of participants at least once and was more likely to occur in the upper rather than lower facial regions. Similar to Porter and ten Brinke (2008), inconsistent expressions occurred more often in masked emotions, and negative emotions were more difficult to falsify than happiness. Inconsistencies were found most often in fearful expressions, and least often for happiness. Overall, it was reported that masking high intensity emotions resulted in more leakage, particularly in the upper face. Porter and ten Brinke also demonstrated that it was more difficult to mask an emotion than to neutralise one. However, leakage did occur during neutralisation of emotion, and in high intensity neutralisation inconsistencies occurred more frequently and for a longer duration than in low-intensity neutralisation, thereby providing strong support for the inhibition hypothesis.

## Felt and False Smiles

A number of studies have reported a physiognomic distinction between non-enjoyment smiles and smiles associated with a genuine positive emotion. Non-enjoyment smiles are marked by the action of the *zygomaticus major* (muscle between the cheekbones and lips) that pulls the lip corners outwards and upwards into a characteristic 'smile' expression. An enjoyment smile activates both the *zygomaticus major* and the *orbicularis oculi pars lateralis* (muscles orbiting the eye) which when contracted, cause wrinkling at the outer corners of the eyes (a 'crow's feet' appearance) (Ekman, Davidson, & Friesen, 1990; Ekman et al., 1980; Ekman & Friesen, 1982; Ekman & O'Sullivan, 2006). This was first discovered by French anatomist Duchenne de Boulogne and is referred to in the literature as the 'Duchenne Smile'. Duchenne observed the differences in the appearance of a genuine, spontaneous smile (one which includes activation of both zygomaticus major and orbicularis oculi) and one that resulted from electrical stimulation of just the zygomaticus major muscle. Duchenne and Non-Duchenne smiles are shown in Figure 2.

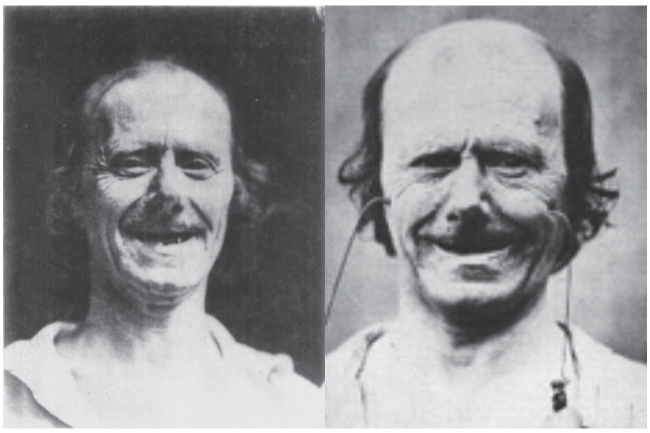


*Figure 2.* Left: a spontaneous, enjoyment smile (Duchenne Smile) activating both the zygomaticus major and the obicularis oculi muscles. Right: (non-Duchenne Smile) has the zygomaticus major muscle electrically stimulated to produce the smiling mouth. However, there is no activity of the obicularis oculi (Duchenne, 1862/1990, printed in Ekman et al., 1990, p. 343).

Regarding his observations, Duchenne wrote:

> "The emotion of frank joy is expressed on the face by the combined contraction of the zygomaticus major muscle and the orbicularis oculi. The first obeys the will but the second is only put in play by the sweet emotions of the soul; the … fake joy, the deceitful laugh, cannot provoke the contraction of this latter muscle … The muscle around the eye does not obey the will; it is only brought into play by a true feeling, by an agreeable emotion. Its inertia, in smiling, unmasks a false friend" (Duchenne, 1862/1990, as cited in Ekman et al., 1990, p. 342).

Duchenne's advice can assist an observer to distinguish between an involuntary smile of enjoyment, and a voluntary, false smile, expressed for another purpose. A non-Duchenne, or simulated, smile sometimes occurs to achieve other goals, such as concealment, deception, and social politeness (Bonanno et al., 2002). A false smile is also the most commonly used expression to conceal or mask an emotion (Ekman, 2009). In a study by Ekman et al. (1990), participants watched both pleasant and unpleasant films whilst being video recorded. Analysis of the videos revealed that Duchenne smiles occurred more frequently during the pleasant films and were also related to subjective reports of enjoyment. Non-Duchenne smiles did not differ in frequency during the viewing of pleasant and unpleasant films, which supports the proposition that non-Duchenne smiles are unrelated to emotional valence.

Some previous studies of deception have mistakenly concluded that smiling is related to untruthfulness, but these have failed to distinguish between Duchenne and non-Duchenne smiles. When Duchenne smiles are separated from all other types of smiles, it is possible to identify whether someone is experiencing genuine enjoyment or whether they are masking another emotion with a smile (Ekman, 2003). Being able to distinguish the Duchenne smile from other smile types can provide useful information for a therapist. For example, it has been found that clients showing improvements during psychotherapy also display increased Duchenne smiling, and that depressed patients show more Duchenne smiling during the discharge interview compared with during the intake interview (Ekman et al., 1990).

## Spontaneous and Masked Emotional Expressions

The terms 'felt' or 'spontaneous' emotional facial expressions are used to refer to expressions that are congruent with an underlying emotional state. Matsumoto et al. (2009, p. 214) define a spontaneous expression as "a naturally occurring expression that is not produced by direct, external requests by another person". However, they acknowledge that spontaneous expressions meeting this definition are difficult to obtain for research and most research to date has not used such naturally occurring spontaneous expressions. 'Feigned', 'fake', or 'deliberate' expressions refer to those intentionally posed by the sender for a particular purpose. These can be expressions posed in the absence of an underlying emotional state, or a posed expression can be used to 'mask' a felt expression, being incongruent with the emotion experienced (e.g., smiling when one is feeling sad). Ekman and O'Sullivan (2006) describe how felt expressions differ from feigned expressions in morphology, timing, symmetry, and cohesion. Feigned expressions vary in appearance and shape, and an absence of reliable movements can suggest a voluntary, rather than involuntary, expression. The presence of reliable facial movements should suggest a genuine, felt emotion is being expressed. The timing of expressions also varies, with onset being more abrupt and lasting .5 to 5 secs for genuine expressions, although this may vary with circumstances. Briefer or longer expressions may indicate a deliberate expression. With the exception of contempt, genuine expressions are generally more symmetrical than deliberate expressions. Although this difference is difficult to detect in real-time, it has been observed during FACS coding. Stronger asymmetry is noted on the left side of the face whereby the expression is more intense compared with the right side, when smiling is deliberate (Ekman et al., 1981). Cohesion refers to the congruence between the expression and the spoken word and other behaviours, with authentic expressions occurring more smoothly.

## Reliable Facial Muscles

Expanding on Duchenne's and Darwin's ideas, Ekman (2003) proposed that there are some facial muscles not generally under voluntary control. He refers to these as *reliable facial muscles*, which he claimed communicate specific emotions when activated. He contended that the reliable facial muscles can only be voluntarily activated with great effort and difficulty, and also that spontaneous activation of these muscles is very difficult to suppress. Learning and practicing how to activate the AUs, as occurs during training in FACS coding, can lead to some improved control of these facial muscles.

Several studies (e.g., Gosselin et al., 2010; Mehu et al., 2012) have tested Ekman's theory of reliable facial muscles and found that there are several AUs that are difficult to produce voluntarily, with the most reliable AUs able to be performed by only a small percentage of people. For example, in a study involving 20 adults, Gosselin et al. (2010) found that only one person was able to voluntarily activate the upper lip raiser, a component of disgust and anger expressions, without simultaneously co-activating other non-target AUs. A similar result was reported for activation of the outer brow raiser, associated with fear and surprise, with only one person successfully achieving this in at least one out of five trials, without co-activating other non-target AUs. The success in activating the nasolabial furrow deepener, a component of sadness, was also low even when co-activating other non-target AUs. The presence of inconsistent expressions may indicate feigned expressions; for example, most participants were able to activate the lip stretcher (a component of a fear expression), but also simultaneously activated the chin raiser, which is not involved in fear. However, conclusions about the link between inconsistent expressions occurring during feigned emotion expressions cannot be made from this study as participants were only asked to activate a single AU, and were not requested to feign emotions. It is therefore possible that feigning an emotion, rather than just activating a single AU, may yield different results.

Mehu et al. (2012) used the term *reliable AUs* to refer to Ekman's (2003) *reliable facial muscles* and labelled the AUs that are more easily activated, *versatile AUs.* They analysed the controllability of the range of AUs listed in the Facial Action Coding System using an online survey of certified FACs coders. The coders were asked how easy they had found it to perform each AU when they obtained FACS certification, responding on a scale ranging from 0 (not easy at all) to 100 (extremely easy). The responses for each AU were classified from the least to the most controllable. The reliable AUs, as a group, were significantly less controllable than the versatile AUs. Mehu and colleagues' survey provided a conservative assessment of controllability as during FACS training, coders learn and practice how to control AUs. Therefore, they were likely to be better at controlling their facial behaviour than others in the general population. Although these results were consistent with those reported by Gosselin et al. (2010) who had used inexperienced participants, the FACS coders were relying on their recall of task difficulty from when they had completed certification, not an evaluation of difficulty during actual performance.

In a second study using actors to portray emotion expressions, Mehu et al. (2012) compared the general effect of reliable AUs and versatile AUs on observers' judgements of emotion, perceived authenticity, and perceived intensity of the expressed emotion. Reliable AUs were significantly associated with perceived authenticity and intensity compared with versatile AUs (71% for reliable AUs, 14% for versatile), and to a lesser extent, with accuracy of emotion recognition, corroborating the reliable facial muscle hypothesis. Although actors were used instead of an emotion induction procedure, the actors were encouraged to feel the emotion requested. However, there was no examination of the actors' subjective feelings of emotions and the emotion expressed. This study also relied on judges' perceptions of emotion expressions without an independent check of what emotions, if any, were experienced by the actors during portrayals. Therefore, accuracy was based on judges' attributions of affective state matching what the actors were asked to express, rather than the emotion they actually felt.

Ekman (1993) concluded that there are no 'reliable' muscles for the expressions of disgust or surprise. This suggests that these two expressions would be relatively easy to falsify. However, Gosselin et al. (2010) found that some of the AUs in disgust and surprise were difficult to activate voluntarily, making it hard to fake these expressions.

## Micro-Expressions

Ekman and Friesen (1969) discovered brief 'micro' expressions when viewing films of psychiatric patients who had concealed their suicidal intentions during a clinical interview. Ekman (2009) proposed that micro-expressions occur when an attempt is made to mask or neutralise an emotion, and 'leakage' of the suppressed emotion occurs as a full face 'micro-expression' lasting about 200 ms. Micro-expressions were barely perceptible in real time, but when film was shown in slow motion, they conveyed emotional information. Some expert observers could even read the emotional information displayed by micro-expressions in real time (Ekman & Friesen, 1969).

When attempting to deceive, a simulated expression is used more often to mask a felt emotion than a neutralising expression. Ekman and Friesen (1969) suggested three reasons for this. Firstly, the omission of an expression needs to be filled to avoid deception being conspicuous by a void in communication. Secondly, to provide a barrier against breakthrough of the inhibited expression (a micro-expression) as when the felt emotion is strong, a neutral face is unlikely to succeed in preventing leakage - the simulation of a mutually incompatible expression provides more protection against signs of the felt emotion escaping. And, finally, a false expression may be required to achieve the goal of deception. Although Ekman and Friesen believe masking is more protective against leakage of an underlying emotion, Porter and ten Brinke (2008) found less leakage occurred during neutralised relative to masked expressions. However, Porter et al. (2011) found emotional inconsistencies were more likely during neutralisation of high emotion.

These different outcomes may have been due to differences in the intensity of the emotion used in the two studies.

In Porter and ten Brinke's (2008) examination of inconsistent facial expressions, they also looked for micro-expressions. However, the leakage that occurred often lasted longer than a micro-expression and was not always a full-face expression. In another analysis of videotaped expressions of genuine and deceptive individuals involving high-stakes interpersonal deception, ten Brinke and Porter (2012) found that micro-expressions rarely occurred. Furthermore, they were equally present in both sincere and deceptive emotional displays. Also, in an investigation of high- and low-intensity emotion expressions where participants deliberately masked or neutralised expressions, Porter et al. (2011) analysed video recordings of 1,711 expressions (256,650 frames in total) and found no complete micro-expressions, and only 18 partial micro-expressions. These did not occur simultaneously in both the upper and lower halves of the face, and only six out of 10 micro-expressions revealed the deceiver's underlying emotion. Thus, Ekman's notion of micro-expressions occurring during emotional leakage was not supported. Most instances of emotional leakage lasted longer than the 200 ms proposed by Ekman, with most lasting closer to a full second. The authors suggested that "subtle" deceptive emotional expressions occur relatively frequently, particularly during high intensity emotional displays. These expressions occur more often in the upper face, which is harder to control, and last long enough to be more readily detected by an observer than "micro-expressions".

## Subtle Expressions

Many of the studies of facial expression recognition have focused on high-intensity, full-face prototypical expressions of emotion, usually involving static displays. However, in real-life, what are more likely to occur are partial or blended expressions of emotion. Subtle expressions can occur when an emotion is just beginning to be experienced, or when individuals are attempting to regulate their emotional expression, resulting in a very slight, or partial, expression. Ekman (2007) believes subtle expressions occur more frequently than full-face and intense expressions and often provide more information as they provide clues about what the person is feeling but is choosing not to disclose verbally. Blends can occur when more than one emotion is being experienced, for example, feeling angry and afraid at the same time, and, according to Ekman, blends occur more often than pure emotions. Matsumoto and Hwang (2014, p. 349) offer an operational definition of "subtle" facial expressions as being "emotional expressions that involve relatively low-intensity and/or few appearance changes in the face". When FACS coding is used, subtle expressions are those defined as AU C-level intensity or lower. In addition to low-intensity versions of full-face configurations, subtle expressions may involve full-face configurations or only single components ("partial" expressions).

Although studies have largely focused on full-face, high-intensity expressions of emotion, there is a small number that have investigated judgements of facial expressions of emotion of varying intensity (Hess, Blairy, & Kleck, 1997; Matsumoto, Consolacion, Yamada, & Suzuki, 2002; Matsumoto & Hwang, 2014). Unsurprisingly, these studies found that judgements of low-intensity expressions are less accurate than high-intensity ones. Ekman (2007, p. 242) states that recognising subtle expressions is difficult and people are less accurate because they do not "appear to utilise the information contained in the more subtle expressions." Russell, Bachorowski, and Fernandez-Dols (2003) also found that recognition rates of spontaneous emotional expressions were lower than recognition of posed expressions, although this may have been due to the use of static images of expressions. It has been suggested by Ekman and Friesen (1982) that dynamic displays of expressions provide important temporal information that is not available in static images.

In an investigation of the effects of movement on recognition of subtle emotion expressions, Ambadar, Schooler, and Conn (2005) found that with the exception of happiness, participants were significantly more accurate at perceiving subtle expressions that were presented dynamically compared with static images. Some expressions that were too subtle to be identified in static displays became apparent during dynamic displays. Bould, Morris, and Wink (2008) and Bould and Morris

(2008), examined the effects of temporal dynamics in the perception of subtle expressions. They found that recognition of emotion expressions was enhanced by movement, particularly when individuals observed subtle rather than intense facial emotion expressions. A review conducted by Krumhuber, Kappas, and Manstead (2013) found strong support for the use of dynamic displays, particularly for subtle expressions, to enhance emotion judgements, as well as influencing accuracy in recognising authenticity of expressions.

Given that in everyday social interactions, subtle, and sometimes blended, expressions of emotion are more likely to occur than full-face expressions, the ecological validity of emotion recognition studies would be enhanced by using subtle dynamic expressions, rather than posed, static, and full-face expressions. As well, the inclusion of blended expressions of emotion is necessary to more accurately reflect social interactions when people experience more than one emotion. The present study utilises single and blended dynamic expressions of emotion in the investigation of ERA. Furthermore, as masked expressions frequently occur in social interactions in order to deceive or to conceal an underlying emotion, these, as well as genuine expressions, are also investigated.

## Verbal Deception

The field of research into deception detection is replete with interest in deceptive behaviour largely consisting of detecting the telling of untruths (e.g., DePaulo et al., 2003; Pete Blair, Levine, & Vasquez, 2015; Vrij, Granhag, Mann, & Leal, 2011; Wright Whelan, Wagstaff, & Wheatcroft, 2015) as deception occurs frequently in social situations. DePaulo et al. (1996) found in a survey of students that they lied, on average, once in every three social interactions, and that general community members lied once in every five social encounters. On analysis of lie content, it was found the lies were most frequently about feelings ($M$ = 37.4%), with faking positive feelings accounting for approximately 25% of all lies. The lies in DePaulo and colleagues' study were limited to small "everyday lies" conducted for the purpose of impression management, social support, and altruism. Even though these lies were considered inconsequential, the data obtained were from the perspective of the liar and therefore subject to self-serving bias. The study did not include larger, more consequential lies.

O'Sullivan and Ekman (2004) claim to have found 29 out of 12,000 professionals that can detect high-stakes lies at "wizardry" levels, defined by performance accuracy of between 80-90%. However, this is disputed by Bond Jr and Uysal (2007) who state O'Sullivan and Ekman's tests for wizardry have no psychometric and statistical criteria and that standardised testing procedures were not used. A meta-analysis by Aamondt and Custer (2006) investigated the deception detection abilities of a range of professional and other groups including teachers, social workers, psychologists, criminals, judges, and various law enforcement officers, finding that psychologists had mean accuracy rates of 61.6% across four studies. The groups that performed better than psychologists were teachers (70%), social workers (66.3%), criminals (65.4%), and secret service agents (64.1%). However, these data were only obtained from one study for each group; therefore, the results are tenuous and more research is required. Those whose accuracy was less than psychologists were judges (59.0%), police officers (55.3%), customs officers (55.3%), federal officers (54.4%), students (54.2%), and detectives (50.8%). The least accurate were parole officers (40.42%) but these data were also obtained from only one study. They also found that individual differences such as sex, age, education, confidence, or law enforcement experience, were not related to deception detection abilities. There was a very small ($r = .14$) correlation between self-monitoring and deception detection. The authors noted that most, if not all, of the studies in the meta-analysis lacked ecological validity as they did not include "real world" conditions related to deception detection such as high-stakes consequences for being caught, no behavioural baseline provided for comparison, and lack of spontaneity. As well, in "real-life", judgements can be made based on factors such as inconsistencies with previous statements or statements of others, indirect responses given to questions, and illogical stories.

In another meta-analysis, Bond Jr and DePaulo (2006) examined 206 studies of deception detection. Studies included in the analysis required that individuals received no training or instructions in deception detection and in which participants made judgements about both lies and truths. The meta-analysis found the average accuracy in discriminating liars from truth-tellers was 54% and that rates of deception detection varied little between studies. A possible explanation for low accuracy rates is that many studies use college students as observers; students do not routinely engage in deception detection (Vrij & Mann, 2001b). Another reason suggested was the use of low-stakes lies as stimuli; these may not elicit prominent clues to deception (Vrij & Mann, 2001b; Wright Whelan et al., 2015).

Wright Whelan et al. (2015) suggested that increasing the stakes might exacerbate the factors that lead to the production of deceptive cues, resulting in more prominent features of deception being displayed by the deceiver. They tested this theory using real-life, high-stakes video footage of family members appealing for information concerning missing or murdered loved-ones, half of whom were appealing honestly, the other half being deceptive (later found to be involved in the death or disappearance of their relative). Being strongly motivated to succeed in the lie, lies relevant to identity such as being scrutinised or evaluated, and lies about transgressions, are all moderating factors affecting the strength of cues to deception (DePaulo et al., 2003). Given that the deceivers in the videos would have been highly motivated to succeed at their lie, to protect their self-identity and to conceal their own transgressions, the videos provided ecologically valid stimuli of high-stakes liars. Using these high-stakes videos, Wright Whelan and colleagues investigated the accuracy of deception detection of 70 police officers, compared with 37 undergraduate students. Overall, all groups of participants produced accuracy rates above the level of chance. Mean accuracy was 70.75% ($SD$ = 8.68). Variations in individual accuracy ranged from 45% to 91%, with all police officers achieving accuracy rates of 74% or higher, compared with only 27% of non-police participants achieving 74% or higher. Although police officers were more accurate at detecting deception than students, there was no difference between groups in accuracy of detecting honest appeals. In attempting to account for the higher accuracy for deception detection amongst police compared with students, the authors propose an interesting explanation. During training, police officers are encouraged to adopt a more liberal criterion for the classification of deception, and a more cautious criterion when classifying honesty. This would result in a negative bias, leading them to attend more to dishonest, rather than honest, cues. There was no relationship between number of years of experience as a police officer and accuracy at detecting deception. As this study used stimuli from real-life, high stakes video footage, it demonstrated that accuracy rates improve when the stimulus materials better reflect what occurs in real-life situations where detecting deception is important.

A study similar to Wright Whelan et al.'s (2015), but yielding contrasting results with regard to truth detection, was conducted by Vrij and Mann (2001a). Police officers (N=65) viewed video-tapes of fragments of an interview with a murder suspect telling either lies or truths. The veracity of the suspect's statements was confirmed by collateral evidence, and truths and lies were also distinguished by behavioural differences such as gaze aversion and slower speech. This study did not examine any differences in facial expressions. Accuracy for truths was greater than for lies (70% vs 57%), and overall accuracy was significantly above the level of chance (64%). There was no correlation between accuracy for veracity and age or professional experience. This contradicts findings by Ekman and O'Sullivan (1991) who reported that age correlated negatively with deception accuracy amongst some of the law-enforcement personnel they examined. This is possibly due to a wider age range in their sample compared with Vrij and Mann whose participants had a relatively low and restricted age range. In the Ekman and O'Sullivan study, the deception detection ability of law-enforcement personnel, and other groups was evaluated. The law enforcement staff included police officers and staff from the U.S. Secret Service, Federal polygraphers from the Central Intelligence Agency (CIA), Federal Bureau of Investigation (FBI), and the National Security Agency (NSA). Other groups included judges, psychiatrists, a special interest group of various professionals, and students. Participants observed video-tapes of women giving either honest or deceptive accounts about a film they were viewing. Half of the women viewed a pleasant nature film, whereas the other half viewed a film with gruesome content.

All of the women had to answer questions about what they were seeing and claim to have positive feelings. For accuracy in deception detection, the Secret Service group differed significantly from the other groups; however, none of the other groups differed significantly from each other. No Secret Service staff scored below chance and nearly one-third achieved accuracy scores of 80% or more. The next closest group was the psychiatrists, followed by judges. Across all groups, no relationship was found between accuracy in deception detection and age, sex, or job experience. However, when looking at individual groups, age was negatively correlated with accuracy amongst the Secret Service and the polygrapher groups. The authors speculated that the Secret Service agents performed better than other groups because their roles require them to focus on nonverbal threat cues. They also interrogate people who are mostly telling the truth; dealing with a lower base rate of lying may develop skill in focusing on signs of deceit. The other law enforcement groups, who face a high base rate of lying, may focus more on obtaining evidence rather than detecting deceit.

Ekman, O'Sullivan, and Frank (1999) compared the deception detection abilities of groups of law-enforcement officers with clinical psychologists and academic psychologists. Participants included four law-enforcement groups. Two groups comprised individuals with a special interest or expertise in deception and included 23 federal officers from the CIA, and 43 county sheriffs. The other two groups had less interest or experience in deception detection and included 36 law-enforcement personnel and 86 federal judges. Three groups made up the psychologist participants: the first two groups were clinical psychologists with one group comprised of those with a strong interest in deception (*n*=107), and the remainder with no special interest (*n*=209). The third group was academic psychologists (*n*=125) who did not perform clinical work. The participants viewed video-tapes of 10 people, half of whom lied about their opinion, the other half were truthful. FACS was used to analyse the facial muscle movements of the videotapes to verify that there were significant differences in facial behaviours between the liars and truth tellers. The results showed overall mean accuracy was greatest for the federal officers group (73.0%), followed by clinical psychologists with an interest in deception (67.5%), sheriffs (66.7%), other clinical psychologists (62.1%), federal judges (62.0%), academic psychologists (57.7%), and other law-enforcement officers (50.8%). Clinical psychologists with a special interest in deception were more accurate at deception detection than the other two groups of psychologists. However, the deception-interested psychologists had attended a two-day workshop on "Lying, Deception and Malingering" which may have contributed to their greater accuracy rates, although the report did not state the temporal proximity of the workshop to the main study. For all groups, neither age nor sex was related to accuracy. For most groups, accuracy for detecting lies was greater than for truths. The authors propose that signs of lying are more easily identified than truth signals.

In an examination of age-related differences in both emotion recognition and deception detection, Ruffman, Murray, Halberstadt, and Vater (2012) found that older adults ($M$ = 71.03 years) were less accurate in recognising both lies and truths, compared with younger adults ($M$ = 20.96 years). They also found that deception detection accuracy was mediated by emotion recognition, with scores on the emotion task significantly predicting those on the lie detection task, leading to the authors concluding that emotion recognition assists lie detection.

## Emotional Deception

Although much has been published on deception involving verbal lies, there has been less published on emotional deception. Rather than deceiving through verbal accounts, emotional deception involves attempting to deceive an observer about one's emotional state by concealing felt emotions or falsely displaying an unfelt emotion. Ekman and O'Sullivan (2006) suggest that the face may provide important clues to deception such as visible traces of felt emotions occurring despite concealment efforts.

A pioneering article on the psychology of deception was published by Ekman and Friesen (1969). They proposed that nonverbal signals of deception occur more readily in the body than the face. They theorised that this is so because more attention is paid to the face than to the body, and, secondly, that sociocultural reinforcement of controlling facial behaviour leads to greater awareness of ongoing facial activity, whereas there is little feedback to body behaviour.  This results in people disguising their feelings through masking a felt emotion, or through simulating emotions not felt.  This theory was tested by Ekman and Friesen (1974) and it was found that observers made more accurate judgements about deception when judgements were based on the body rather than the face.  However, this only occurred when prior to the judgement task, observers viewed a sample of non-deceptive behaviour.  When no prior sample was available for comparison, there was no difference in accuracy between the body and the face when observers were unfamiliar with the person's nonverbal repertoire.  A similar finding was made by Brandt and Miller (2009) who reported that observers were more accurate in judgements of truthfulness or lying when they were familiar with the communicator's truthful behaviour.

The idea that deceptive nonverbal expressions of emotion are less easily disguised than deceptive verbal accounts was raised as far back as Darwin (Ekman & Friesen, 1974). Until fairly recently, there has been less interest in emotional deception compared with telling untruthful messages.  A few studies began to appear about 30 years ago (e.g., Ekman & Friesen, 1982; Waxer, 1983) followed by a lapse in interest until the last decade.  The earlier studies tended to focus on a dimensional model of emotion, using continua such as "positive-negative" or "pleasant-unpleasant" rather than specific emotion categories such as fear and anger.  Waxer's (1983) study used actors conveying genuine or simulated displays of high anxiety finding that observers attending to nonverbal cues achieved significantly greater accuracy compared to lexical and paralinguistic cues, suggesting nonverbal cues provide more potent clues to deception.  However, Waxer's study did not attend specifically to facial expressions; other non-verbal clues from hands, eyes, mouth, and torso were also included.

Warren, Schertler, and Bull (2009) examined the correlation between a deception detection task and scores on the Micro-Expression Training Tool (METT) and the

Subtle Expression Training Tool (SETT), as well as the use of verbal and non-verbal clues. Participants' scores on the METT and SETT were compared with their performance in recognising deception or truthfulness when observing video footage of encoders responding either truthfully or deceptively to emotionally evocative or unemotional video scenes. They found that, although overall mean performance in detecting deception was little better than chance, accuracy for detecting emotional lies was significantly above chance (64.34%), a finding supported by subsequent research which found greater accuracy when observing high-stakes lies (e.g., Wright Whelan et al., 2015). The SETT was developed to improve attention to subtle features of emotion expressions, and performance on the SETT was positively correlated with emotional lie detection. This association suggests that attention to subtle facial clues aids detection of emotional lies. However, SETT performance was negatively correlated with performance in detecting unemotional lies. Participants' use of facial expressions in their decision-making was also significantly positively correlated with emotional lie detection. This lends support to the nonverbal leakage theory (Ekman & Friesen, 1969, 1974) that suggests lies based on highly arousing emotions will contain more clues to deception. There was no correlation between METT scores and emotional lie detection. However, there were only three micro-expressions visible in the emotional lies, hardly sufficient to draw a firm conclusion.

Porter et al. (2011) explored whether observers could successfully judge the veracity of genuine and masked emotion expressions. Using a procedure similar to the one adopted in the current study, participants were observed while watching emotionally evocative images from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 1997) while responding with genuine, simulated, masked, or neutralised emotion expressions. Overall accuracy by observers at recognising emotional deception was not significantly above chance. However, when each emotion type was analysed separately, judges were able to detect deception in happy expressions at above the level of chance, but not for expressions of sadness, disgust, or fear. The observers' superior performance in detecting falsified happy expressions may be due to a familiarity bias, as happiness is the more commonly experienced expression (Somerville & Whalen, 2006). Also, the authors reported that in high intensity emotion, more leakage occurred during masking of happy and sad expressions, compared with masking of fear or disgust, which may have provided more clues.

## Facial Behaviour in High-Stakes Deception

Studies report that the success of the untrained observer at detecting deceptive expressions is no better than chance (e.g., Ekman, Friesen, & O'Sullivan, 1988); but this outcome may be due to research being conducted with deception involving low stakes. Wright Whelan et al. (2015) demonstrate that using real-life, high-stakes situations can result in deception detection levels above chance. Telling lies involving high-stakes is generally more difficult than telling low-stakes lies as they can be complex and stressful, and often accompanied by powerful emotions that need to be inhibited or falsified. Lying can also be cognitively demanding as the liar needs to construct an alibi that is plausible, that is consistent with known facts, and which avoids implicating the liar in the crime under investigation (ten Brinke & Porter, 2012).

An investigation into the behavioural consequences of high-stakes deception was conducted by ten Brinke and Porter (2012). This involved examining videotapes of a large international sample of individuals (*N*=78) making public appeals for the return of a missing relative. Approximately half of the pleaders were deceptive (*n*=35) and later convicted of involvement in the crime by overwhelming evidence such as DNA, forensic evidence, or security camera footage. Each video was coded for behavioural (e.g., speech, body language, indicators of cognitive load), and facial signals of deception (universal emotional expressions). Facial coding was performed by coders blind to veracity and trained in a highly reliable procedure based on FACS and developed for Porter and ten Brinke's (2008) study. Relative to genuine pleaders, deceptive pleaders expressed more disgust in lower facial regions (40.0% vs. 16.3%), and were less likely to express sadness/distress in the upper face (31.4% vs. 55.8%) or lower face (20.0% vs.48.8%) than truthful pleaders. These results are consistent with Gosselin et al.'s (2010) findings that only about 20% of people could voluntarily activate the nasolabial furrow deepener, as expressed in genuine sadness. Although the presence of a facial expression does not reveal its source, the authors suggest that expressions of disgust in deceptive pleaders could be a reaction to their own murderous actions, feelings of moral disgust, or a lingering revulsion for the victim.

In addition to facial expressions, ten Brinke and Porter (2012) found deceptive pleaders blinked nearly twice as often as individuals making genuine appeals, possibly due to increased arousal associated with masking emotions. Porter and ten Brinke (2008) also found blink rates increased during masking of emotions. Other cues included the use of fewer words overall and more tentative words amongst deceivers.

The cues to deception reported by ten Brinke and Porter (2012) were included in a comprehensive programme developed to train mental health and legal professionals to detect high-stakes lies (Shaw, Porter, & ten Brinke, 2013). An evaluation of a one-day workshop presented to 42 legal and mental health professionals found participants' mean accuracy in discriminating between honest and deceptive pleaders increased from *M* = 46.4% at baseline, to *M* = 80.9% after training. Components of the training included attending to body language, facial expressions, and analyses of verbal statements. Although promising results were reported, the authors stated it was not possible to determine which elements of the training contributed most to the gains in accuracy observed.

## Individual Factors Contributing to Deception Accuracy

In an early study of individual differences amongst people participating in a deception detection task, Lavrakas and Maier (1979) found that, although overall mean accuracy was only just above chance (54.45%), 22% of the variance was accounted for by individual factors. Accuracy for veracity of verbal accounts was positively correlated with scores on a test of social intelligence, and negatively correlated with time spent watching television and by the number of siblings a participant had.

Porter, Campbell, Stapleton, and Birt (2002) also examined individual differences contributing to accuracy in deception detection. Participants judged the honesty of reports of emotional events presented in audio-visual or audio-only formats. The mean accuracy of all participants was 57.6%; however, one quarter of judges demonstrated accuracy at rates of 75% or higher, with two people achieving 100% accuracy. Individual factors were examined and it was found that accuracy rates were higher in left-handed people, when the target being judged was less attractive, and the target was of the opposite sex. Accuracy was associated with the use of a higher number of facial cues compared with verbal content cues, which were negatively correlated with detection accuracy. Accuracy was unrelated to personality traits.

A study of a clinical population conducted by Etcoff, Ekman, Magee, and Frank (2000) found that people with aphasia following left cerebral hemisphere damage (LHD) performed better than people with right hemispheric damage (RHD) and healthy controls at detecting emotional deception from facial expression clues. The success in interpreting lying from facial expression cues for the LHD group was 73% compared with 50% for those with RHD, 57% for healthy matched controls, and 50% for undergraduate students. All but one of the participants with aphasia was tested more than one year post injury. The one who was tested in less than a year only scored at chance level. Etcoff et al. concluded that an absence of language skills may contribute to the development of more sensitive recognition of non-verbal behaviour, leading to aphasics with LHD becoming more sensitive to subtle emotional clues which may assist in detecting deception. A right-hemispheric advantage was also reported by Malcolm and Keenan (2005) who found that people were significantly more accurate at detecting deception when true and false verbal statements were presented to their left ear, compared with their right. Although Malcolm and Keenan found no handedness differences, unlike Porter et al. (2002), they acknowledged that their method was not particularly sensitive to hand effects and the low number of left-handers in their sample may have also affected this comparison.

Although there is only a small number of studies investigating individual differences that contribute to deception detection accuracy, it may be that social intelligence, prioritising attention to non-verbal over verbal clues, and possibly a right-hemispheric advantage, all contribute to enhancing deception detection.

## SUMMARY OF DECEPTION RESEARCH

Many of the studies outlined on emotion deception have been based on Darwin's inhibition hypothesis where it has been theorised that insufficient suppression of emotional facial expressions results in leakage of underlying emotions providing valuable clues to deception. This leakage is exhibited by fragments of repressed affective expressions which, due to reliable groups of facial muscles not under volitional control, escape unbidden by the deceiver. Although repression of emotions is learned from a young age and shaped by cultural display rules, leakage nonetheless occurs and a number of studies have found support for Darwin's inhibition hypothesis (e.g., Dimberg et al., 2002; Porter & Brinke, 2008; Porter et al., 2011).

Expanding on Darwin's theory, Ekman (2003) suggested the term *reliable facial muscles* to refer to those muscles that are difficult to activate voluntarily. Support has been found for Ekman's suggestion with a number of AUs being identified that are difficult to produce without the corresponding emotion (Gosselin et al., 2010). Also, ten Brinke and Porter (2012) studied videotapes of genuine and deceptive pleaders and found differences in the facial behaviour which further supports the *reliable facial muscle* theory.

Both micro-expressions and subtle emotion expressions appear to provide clues to suppressed emotions. However, less support has been found for the presence of micro-expressions and it seems that although they do occur, they are rare and occur during both genuine and falsified emotion expressions (Porter & Brinke, 2008; ten Brinke & Porter, 2012). Furthermore, the brevity of these expressions can make them difficult to detect. Given that subtle expressions occur more frequently and are more representative of social interactions, it is of interest to investigate whether skill in perceiving subtle expressions is related to accuracy for emotional deception detection. Warren et al. (2009) tested this theory and found a significant positive correlation between scores on the SETT and accuracy in emotional deception detection.

Other deception studies have reported mixed results with many finding only chance ability (e.g., Bond Jr & DePaulo, 2006; Porter et al., 2011). However, it has been found that accuracy for deception detection improves when high-stakes lies are told, or when high intensity emotions are evoked, compared with unemotional lies (e.g., Ekman & O'Sullivan, 1991; ten Brinke & Porter, 2012; Warren et al., 2009; Wright Whelan et al., 2015) due to the leakage of suppressed emotions. Using this information, Porter et al. (2010) and Shaw et al. (2013) developed training programmes which successfully increased the accuracy of deception judgements amongst trainees.

Individual factors contributing to accuracy in deception detection have been investigated and include handedness (Porter et al., 2002), a right-hemispheric advantage (Etcoff et al., 2000; Malcolm & Keenan, 2005), comprehensive training and feedback (Porter, Woodworth, & Birt, 2000; Shaw et al., 2013), and professional experience (Ekman et al., 1999; Wright Whelan et al., 2015).

Although there has been less interest in emotional deception compared with verbal deception, the recent studies that have emerged have shown some promise of identifying reliable nonverbal clues to deception such as movements of hands, eyes, and torso, as well as consistent findings of leakage in facial emotional expressions during deception. The current study explores whether training in recognition of subtle expressions improves accuracy in detection of emotional deception from facial expressions. The training focuses specifically on facial emotion expressions and does not include body language cues. The video clips used for testing recognition accuracy provide only head and shoulders views of the people expressing emotions and there is no sound or verbal content to provide additional cues.

# CHAPTER THREE: THE CURRENT INVESTIGATION

## AIMS AND OBJECTIVES

As outlined above, many of the studies involved in training emotion recognition have focused on micro-expressions and have used the Micro-expression Training Tool (METT) http://www.paulekman.com) or similar, such as the MiX (e.g., Matsumoto & Hwang, 2011). This is despite studies reporting that micro-expressions are rare. Other studies have utilised prototypical pictures of facial affect which have less ecological validity when encounters with others in real-life involve dynamic and often brief facial expressions of emotion.

There is wide agreement in the literature that there are approximately seven basic emotions which are demonstrated by unique facial expressions (Ekman, 2007; Ekman & Friesen, 1975). As emotions can be fleeting, these may occur very briefly, or be presented very subtly due to low intensity of emotion or attempts by people to suppress or mask their expression.

The aim of the current study was to investigate whether a brief, on-line training programme teaching recognition of subtle expressions of emotion would enhance this skill in psychologists. Also of interest was whether the experience level of psychologists impacted on this ability. Finally, the study was also interested in psychologists' skill at detecting deceptive expressions of emotion.

Key findings from previous studies show that brief interventions such as the METT have improved accuracy in recognising facial expressions of emotion in both clinical (e.g., people with schizophrenia: Russell et al., 2008), and non-clinical populations, such as medical students (Endres & Laidlaw, 2009), and sales personnel (Matsumoto & Hwang, 2011). It has also been reported that scores on the SETT are positively correlated with detecting deception (Warren et al., 2009). Research into the effects of age on emotion recognition has also found age-related decline occurs in the recognition of some, but not all, emotions. Finally, a small number of studies have found that deception-interested psychologists, and others, are able to detect deception at levels above chance when the stimuli used involve either high-stakes lies or emotional deception.

## RATIONALE

Emotions are strong motivating states that can triumph over other powerful drives such as hunger, sex, and the will to live (Ekman, 2007). Therefore, awareness of the emotional state of others is paramount in a therapy relationship. Being able to "read" others is advantageous for therapists as it improves communication and can lead to more informed judgements. For example, the facial behaviour of patients who subsequently attempted suicide has been found to differ from those who did not (Archinard et al., 2000). Although there has been considerable interest in assessing empathy in therapists, there has been remarkably little research to date on therapists' abilities in recognising emotional expressions or how this skill may link to therapy outcomes. This is surprising given that accurate emotion expression perception is considered a specific ability related to empathy and the construct of Emotional Intelligence. Additionally, research on training in emotion recognition has been limited largely to clinical populations with only a handful of studies located concerning non-clinical populations. Of these, very few have focused on therapists' abilities; instead, the populations of interest have been medical students, salespeople, and trial attorneys. The supposed importance of therapists accurately recognising emotional expressions suggests a need for further investigation.

Training for therapists has a strong emphasis on what is said and how to respond. However, the use of language to represent emotions is a cultural product, whereas facial expression of emotion is believed by many to have developed through evolution, transcending culture. Attending to non-verbal communication such as facial expressions can lead to enhanced understanding and the development of rapport in a therapeutic relationship between therapist and client, irrespective of their cultural backgrounds.

The major limitation of most previous studies of emotion recognition is that they have used stimuli presenting posed or simulated expressions. Of interest in the current study is not only whether therapists can accurately recognise emotion expressions, but also whether deliberately posed expressions intended to mask an underlying emotional expression can be discerned from genuine, spontaneously generated expressions of emotion. Additionally, given the importance of this ability when working with people, the present research investigated whether this skill can be enhanced through a brief training intervention. Ekman (2007) suggested that ERA can be enhanced through attentive interactions with others, and through training in the subtleties of facial emotion expressions.

Another limitation of many previous studies is that they have used full-face, high-intensity expressions. However, in a therapy setting, it is likely that clients will display more subtle, or partial, emotion expressions, or expressions containing a blend of more than one emotion. Therefore, a study that investigates skill in recognising more subtle emotional expressions, as well as blended expressions, has greater ecological validity.

As the focus of the current study was on subtle expressions of emotion, which previous research has demonstrated are more difficult to detect, dynamic, rather than static expressions have been utilised. It has been found that the temporal characteristics of dynamic expressions enhance the recognition of subtle emotion expressions.

## HYPOTHESES

The following hypotheses are based on findings in the existing literature.

### Hypothesis 1

Although there is no research on the effectiveness of emotion recognition training with therapists, research with other populations has demonstrated it is possible to increase facial emotion recognition skills with an intervention. Based on the theory of universal expressions unique to discrete emotions, it was expected that training which focuses on the specific facial behaviour associated with each emotion would increase trainees' abilities to accurately identify a specific emotion expression.

### Hypothesis 2

In line with the existing research which reports that accuracy of emotion recognition declines with age, it was expected that less experienced, younger therapists would demonstrate greater accuracy in recognising emotion expressions compared with older, more experienced therapists.

### Hypothesis 3

In accordance with the extant literature demonstrating that mean accuracy rates for therapists detecting deception are about 60%, it was hypothesised that deceptive expressions would be identified at levels above chance before training. It was also expected that training in recognising subtle emotion expressions would improve deception detection skills.

# CHAPTER FOUR: METHOD

*The study involved a three-stage process; the first stage involved creating the facial expression stimuli to be used in Stage 2, using an emotionally evocative process. In the second stage, participants viewed the stimuli, and made decisions about whether the expressions were genuine or masked, and what emotion expression was displayed. This was followed by half the participants completing training in facial emotion recognition. All participants then repeated the emotion recognition task. Stage 3 was a replication of Stage 2, using an updated version of the facial emotion recognition training. This chapter outlines the methods employed, including details of the participant sample for each stage, and the recruitment process. The method for creating the stimuli, the emotion recognition training used, and the procedure used to collect the data are described.*

## AIM

The aim of the study was two-fold: first, to assess therapists' degree of accuracy in recognising facial expressions of emotion, and second, their ability to detect deception. As it has been found that leakage of an underlying emotion can occur in a facial expression when an individual attempts to mask an emotion with another expression, it was also of interest to study whether deception detection or emotion recognition could be enhanced by a training tool developed specifically to improve skill in recognising subtle expressions of emotion.

## STAGE 1: CREATION OF EMOTION AND DECEPTION STIMULI

### Participants

To create the stimulus set for Stage 2 of the study, 19 Massey University students (nine male and 10 female) aged between 19 and 28 years (*M* = 22.3 years; *SD* = 5.6) volunteered to participate in response to a poster advertisement (refer Appendix A) placed on noticeboards in student common areas around the university. An *Information Sheet* (refer Appendix B) was emailed to all potential participants who responded to the advertisement. Self-reported ethnicity of participants was eight NZ European, one NZ European/Maori, five Chinese, two American, one American/Finnish, one Indian, and one Turkish. Participants received $15.00 as compensation for time and travel costs.

## Ethics

A low risk notification was approved by the Massey University Human Ethics Committee (refer to Appendix L for approval letter).

## Stimuli

The study utilised a selection of emotionally evocative images from the International Affective Picture System (IAPS; Lang et al., 1997). The IAPS is a set of 956 emotionally evocative colour photographs that have male, female, and combined norms for ratings of valence (ranging from unpleasant to pleasant), arousal (ranging from calm to excitement) and dominance (ranging from controlled to influential). The IAPS has been developed for use in experimental investigations of emotion and attention. Thirty-six pictures (refer Appendix F) were selected for the study representing a range of low to high arousal (*range* = 1.72-7.26) and valence (*range* = 1.45-8.34) based on the combined means for males and females provided with the IAPS norms. These norms were based on a 9-point rating scale for both arousal and valence. Images selected were high valence, low valence, or neutral. As the IAPS offers dimensional rather than discrete emotion categories, the images were also selected according to the emotion categories provided by Mikels et al. (2005). In their study, 60 participants rated negatively valenced pictures on a 7-point scale according to the degree they experienced a particular emotion, choosing from fear, sadness, disgust, and anger. Ratings could be given for more than one emotion. Means for each category rating were calculated and each picture categorised as either a *single* emotion category (if the mean for a single emotion was higher than the combined means of the other three emotions); a *blended* emotion category (if two or three means were higher than the rest); and as *undifferentiated* if all means overlapped. A similar process was carried out with positively valenced images used to create a happiness category. In the present study, six images from each of the following single emotional categories defined by Mikels et al. were selected: disgust, sadness, and happiness. Disgust was chosen as this expression can sometimes be confused with anger, particularly amongst younger people. Moreover, disgust can be reliably and ethically elicited in a laboratory setting. Additionally, according to the categories defined by Mikel et al. there were no images rated as purely anger evoking. All of the anger images contained in the IAPS elicited a blend of negative emotions such as anger and disgust, anger and sadness, anger and fear, or a blend of all three. Sadness was selected due to its relevance to psychological therapy, given that clients often present with various symptoms of distress. Sad expressions may also be masked, particularly by men, due to cultural expectations where it is less acceptable for men to show distress publicly (Ekman & Friesen, 1975). Happiness was used as although this is an easily recognisable expression, it is also most commonly used to mask another emotion (Ekman, 2009). Being able to distinguish between Duchenne and non-Duchenne smiles is also helpful for therapists as increased Duchenne smiling has been linked to client improvement during therapy (Ekman et al., 1990). Fear was excluded as according to Mikels et al., there were only 12 images in the IAPS that elicited purely fear. These images largely contained pictures of snakes or dogs, which in a New Zealand context is unlikely to evoke a strong fear response in the participants. Furthermore, a large number of the participants who volunteered to participate in the creation of stimuli, were veterinary science students, and therefore, even had this category been included, given their area of training, it is unlikely they would have been sufficiently aroused by the fear-evoking pictures. There were no categories reported by Mikels for evoking Surprise or Contempt, so these categories were not used in the current study.

A further eleven images of neutral valence and low arousal, and seven highly valenced images were also used. Although Mikels et al. (2005) provide discrete categories for sadness and disgust, they state that as happiness is a broader construct, ‘happy’ images were categorised as ‘awe’, ‘amusement’, ‘contentment’, and ‘excitement’, with many images generating a blend of these emotions. The ‘happy’ images selected for the current study included the full range of these categories, including blends. A sample of the images selected is displayed in Figure 3.

*Figure 3.* Sample of images from the Powerpoint slideshow showing three emotion categories: Disgust (left), sadness (middle), and happiness (right).

Screening of the IAPS images was accompanied by music intended to increase the intensity of the negative emotions. Music can trigger emotional responses and influence emotions (Juslin & Västfjäll, 2008) and Western classical-style music has been used in a number of studies interested in emotion induction (Kreutz, Ott, Teichmann, Osawa, & Vaitl, 2008). The music track selected was Earth, composed by Hans Zimmer for the movie *Gladiator*. This music was composed to mediate powerful emotional cues and was categorised by Eerola and Vuoskoski (2011) as inducing negatively valenced, high tension emotions, and contained no lyrics. Although this may have reduced the impact of the positively valenced images presented, additional spontaneous happy expressions were elicited following the slideshow using a series of jokes. The jokes selected were humorous statements made by New Zealand rugby players that had been published in the media (see Appendix G). The rugby player statements were also used in addition to the IAPS images as although the slideshow contained high valence pictures intended to evoke genuine happy expressions, the happy images may not have been sufficient to evoke

Duchenne smiles of sufficient intensity. Furthermore, whilst reading the jokes, participants gave humour ratings which confirmed their genuine enjoyment of the statements.

## Apparatus

A slideshow was compiled with the IAPS images and music using Microsoft Powerpoint. Each slideshow lasted for 4 min, 8 s. The slideshow was displayed on an Apple MacBook Pro, 13” screen laptop computer.

### *Slideshow*

For each slideshow, facial expressions were elicited for two conditions (spontaneous and masked), for all three emotion categories. For each category, participants were required to express disgust, sadness, or happiness. This created one genuine expression (congruent with the emotion category set shown), and two masked emotional expressions. The categories and emotional expressions requested are shown in Table 1.

Table 1

*Emotion Categories and Expressions Requested*

| Category of Images Viewed | Emotional Expression Requested |
|---|---|
| Disgust | Disgust (genuine)<br>Sadness (masked)<br>Happiness (masked) |
| Sadness | Sadness (genuine)<br>Disgust (masked)<br>Happiness (masked) |
| Happiness | Happiness (genuine)<br>Disgust (masked)<br>Sadness (masked) |

## Procedure

Each participant attended individually and the procedure was carried out in one of the School of Psychology's laboratories. Prior to video-recording, a check was made that the participant had read the information sheet, and a *Consent Form* (refer Appendix C) was signed and a *Participant Questionnaire* (refer Appendix D) completed. The instructions for viewing the slideshow were read to the participant (refer Appendix E). Participants were shown prototype photographs of the three emotions that they would be asked to express. One prototype each for disgust, sadness, and happiness from Ekman and Friesen's Pictures of Facial Affect (Ekman & Friesen, 1976) was shown (see Figure 4).

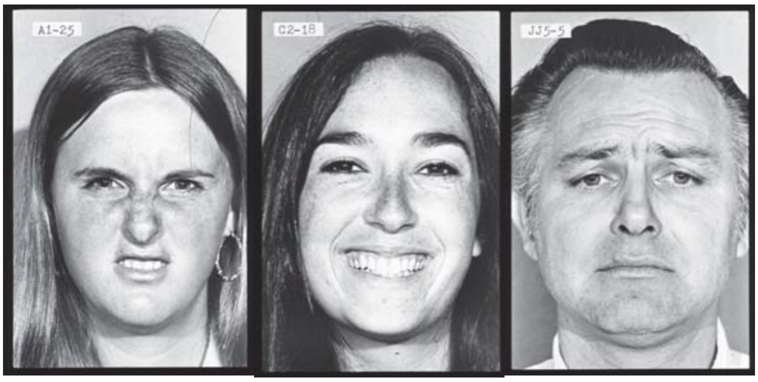


*Figure 4.* Prototypical Expressions: *Disgust (left), happiness (middle), and sadness (right).*

A procedure similar to the one used by Porter and ten Brinke (2008) was followed, although in their study the sets of images were defined by the expressions that participants were instructed to exhibit and each set contained one happy, one sad, and one neutral image. In the current study, a set was defined by the emotion it was intended to elicit, not the facial expression requested.

Two elicitation procedures were used for all participants. First, a slideshow of IAPS images to evoke happiness, disgust and sadness, and secondly, to elicit authentic happiness, participants read a series of jokes presented to them on the computer screen.

Participants were seated at a table directly in front of the Macbook Pro laptop computer on which the timed slideshow was presented. A full-frontal close-up, full-face view of the participant was recorded in colour by the computer's in-built camera. Room lighting was standard and a white screen was placed behind each participant to create a neutral background.

While being video recorded, each participant viewed the timed slideshow. Three emotion categories, each containing six images, were included in each Powerpoint presentation – disgust, happiness, and sadness. Each emotion set was presented in counter-balanced order to create a total of six slideshows. Three participants viewed each slideshow. Porter and ten Brinke (2008) used a similar procedure and analysed the impact of counter-balancing and found order effects were minimal.

The slideshow commenced with three neutrally valenced images of low arousal to neutralise any emotion the participant was feeling about participating and being video-recorded, and to orient them to the slideshow. The neutral images were followed by the first of an emotionally arousing pair of images. The first image of a pair was displayed to act as a 'primer' for the emotion that was to be elicited, followed by an instruction to express an emotion whilst viewing the next slide, which was from the same emotion category set. After each pair of emotionally arousing slides was presented, a neutral slide, such as a picture of clothes pegs, was presented for 5 s to allow the participants to return to a neutral face. Each image appeared for 5 s and each instruction slide appeared for only 1 s. This was considered long enough to read the two-word instruction and with such a short presentation, it is unlikely that it interfered with the emotional impact of the primer slide.

***Genuine Expressions***

To elicit a genuine emotion expression, a primer image was presented to evoke the desired emotion, followed by a request to express the emotion congruent with the slide category. For example, a disgust image was presented, and a request was made to express disgust, followed immediately by screening the second disgust image. This was then followed by a neutral slide.

***Masked Expressions***

A similar procedure was followed to elicit masked emotional expressions, except the expression requested was incongruent with the emotion evoked. For example, a request to express happiness when viewing sad or disgusting images.

All of the slideshows ended with a number of highly valenced images. As there were a number of negatively valenced images presented, the purpose of this was to restore the participants to a positive affective state.

Participants then viewed a slideshow of jokes. As humour can be very individual and culturally dependent, as a qualitative check, participants were asked to give a humour rating for each joke on a scale of 0-10. Ratings given ranged from 0-10 (mean rating = 5.3; *SD* = 1.3). The purpose of this was to record additional spontaneous happy expressions and also to enable participants to finish the session positively after earlier seeing some gruesome images.

There was a technical failure with one of the video recordings requiring a new replacement participant who viewed the same slideshow as the one abandoned.

***Video Editing***

Following the video creation, each participant's video recording was viewed while simultaneously viewing the Powerpoint slides that they had been watching at the time. Notations were made of the time on the recording and the frame numbers where they viewed each emotion-evoking slide; the expression instruction was also noted. Each video clip was categorised by emotion type and as genuine or fake, depending on whether the expresser had been expressing an emotion congruent with the image observed.

Each video was then dissected to create short (5-8 s long) clips, each relating to the discrete phase of emotion expression as noted above. These were edited with iMovie and each clip was assigned a unique number corresponding to the participant and the emotion expression. Video clips were then randomly ordered to create a Quicktime movie containing 153 emotion clips and lasting 29 min, 36 s.

***Field-Testing***

The Quicktime movie was field-tested by a group of four post-graduate psychology students and one senior clinical psychologist. The group watched the QuickTime movie and for each clip selected which emotion term they believed best matched the emotion they were viewing, and selected whether they believed the emotion was genuine or fake.

Each participant's record was then scored and the results entered into a database. The items where the group collectively scored 90% or greater accuracy were discarded to avoid ceiling effects. A total of 49 items were discarded leaving a set of 104.

***Coding of Facial Expression Stimuli***

The remaining 104 video clips were sent to a certified FACS coder for a preliminary analysis. The FACS coder previewed each clip and recorded the emotion expressed, and the veracity of the emotion as either genuine or fake (masked). These were then compared with the researcher's notations made about the veracity and emotion type that were expressed during the viewing of the PowerPoint slideshow. Only those coded by the FACS coder that were consistent with the researcher's record of emotion and veracity were used. Even though the expressions elicited were intended to be either genuine or fake, based on the manipulation, this does not guarantee that the participant actually felt a genuine emotion at the time, and therefore their expression may not be authentic. Although the IAPS images are categorised for each emotion, the emotion is based on the subjective experience of most viewers. For example, a person may not feel genuine happiness at cute kittens whereas most other people might.

From these, 57 clips were selected and each one edited to be 5 s in length, with the apex of the expression being retained in the mid-section of the clip. These were then coded by the FACS coder, blind to both veracity of display and intended emotion expression. Each frame of the video-recording was analysed for the presence of universal emotional expressions. After coding, a final set of 50 video clips was retained, 25 expressed genuine expressions and included nine happiness, ten disgust, and six sad expressions. The remaining 25 expressed 'fake' or masked expressions and included eight happiness, eight disgust, and nine sad expressions. There were 17 clips that included traces of a leaked emotion. Of these, leakage occurred in two clips where a genuine emotion was expressed with a blend of disgust and sadness occurring. Although only images from discrete emotion categories were used, it is possible that both of these emotions occurred simultaneously for the expresser, possibly the result of a personal reaction to the image. The remaining 15 clips exhibiting leakage were masked expressions where the underlying, felt emotion leaked out. Veracity of expression was determined by the presence or absence of reliable AUs, for example, a genuine expression of happiness included AU6 (cheek raiser and lid compressor) and AU12 (contraction of the zygomaticus major, the lip puller). An additional three clips were used for the practice trial section. A 40% sample (20 clips) was coded by a second, certified FACS coder. Inter-rater reliability was assessed by calculating the ratio of agreement for each clip by totalling the number of facial AUs agreed by the two coders, multiplied by two, and divided by the total number of AUs scored by each coder. The mean ratio of agreement was 0.88, based on the agreement index developed by Wexler (1972, cited in Ekman et al., 1985).

A programme which became known as the Deception Detection and Emotion Recognition Task (DDERT) was developed by a computer programmer based at Massey University's School of Psychology for use in Stage 2 of the study. This was a Windows-based programme which presented the 50 video clips in a different randomised order every time the programme was run. Prior to the presentation of the 50 video clips, a practice trial of three clips was presented before the main test. After each presentation of a video clip, a screen appeared with a veracity option (Genuine or Masked). For each veracity option, users chose Definitely/Probably/Possibly, depending on the strength of their opinion about whether the expression was genuine or masked. They also selected from a list of seven emotions, which emotion was expressed. A maximum of two emotions could be selected.

A sample of still images taken from the video clips is shown below in Figure 5.

FACS codes at apex of expression: Image being viewed whilst instructed to express disgust
4D+7C+9C+17A+20E

Genuine disgust. Disgust is seen as a nose wrinkling (AU9) or upper lip raise (AU10). The AU9 adds to the lowering of the brows. The combination of AU7 (lid tightening) and AU20 (lip stretch) give the impression that the participant was not only genuinely disgusted, but quite taken aback by what they saw.

FACS codes at apex of expression: Image being viewed whilst instructed to express sadness
1E+4E+11D+17D

Genuine sadness. Very difficult to both create the stimulus to elicit a genuine sadness, and very difficult to fake good sadness. Perfect / \ eyebrow arrangement with a maximum intensity inner eyebrow raise (AU1) and brow lowering (AU4). There are clear signs of the deepening of the nasolabial furrow (AU11), and the raising of the chin boss (AU17). Very genuine.

*Figure 5* cont. over

FACS codes at apex of expression: 4C+12A+53B

Image viewed whilst instructed to express disgust.

<u>Fake disgust, hiding happiness.</u> The frame shows a flicker of a smile that disappears quickly, leaving a trace of a lip corner puller (AU12), slightly more visible on the participant's right side. The brow lowering (AU4) and head raise (AU53) could be part of the fake disgust; however there was no indication of the nose wrinkler (AU9), upper lip raiser (AU10) of chin boss (AU17) that would be normal in genuine disgust. Frame has the participant do a lip stretch (AU20) quickly (a "yuck" look), which disappears quickly.

*Figure 5.* Examples of the stimuli used in this study.

## STAGE 2: DECEPTION DETECTION AND EMOTION RECOGNITION TASK AND FACIAL EMOTION RECOGNITION TRAINING – STUDY 1

### Participants

Psychologists who had not previously completed any training in facial emotion recognition were invited to participate in Stage 2 of the study via advertising to members of the New Zealand Psychological Society and the New Zealand College of Clinical Psychologists (refer Appendix H). They were either currently practicing or had been practicing within the last 12 months to allow for people who may have been on parental or study leave. Sixty-six people responded to the advertisements and were emailed an *Information Sheet* (refer Appendix I). Of these, nineteen withdrew due to work or other commitments, two experienced technical difficulties due to incompatibility with the CD-ROM and their desktop computer, and two did not return both data files for analysis so their data were not included in the study. Of the remaining 43 who participated, there were 30 females and 13 males aged between 27 and 68 years (*M* = 42.2 years; *SD* = 11.3). Self-reported nationality of participants was 37 New Zealand, one United Kingdom, one Netherlands, three Australian, and one Chinese.  All were practising psychologists with experience ranging between 1 and 40 years (*M* = 11.80 years; *SD* = 10.66).

## Stimuli

### ***Deception Detection and Emotion Recognition Task (DDERT)***

The DDERT programme developed from the emotion expression stimuli created in Stage 1 was copied to individual CD-ROMS, one for each participant.  The DDERT contained the deception detection and emotion recognition tasks for assessing each participant's deception detection and emotion recognition abilities.

### ***The Subtle Expression Training Tool (SETT)***

Ekman's SETT was used to train participants in the training group in facial emotion recognition (see www.paulekman.com).  The SETT teaches recognition of very subtle signs of an emotional expression, often appearing in only one region of the face such as the brows, eyelids, cheeks, nose, or lips.  The training has been developed as these "mini" expressions can leak out when a person is trying to conceal a strong emotion; therefore, it is suitable for training detection of masked emotional expressions.  The SETT has two sections:  *Learn* and *Practice*. The *Learn* section provides specific information about the subtle signs of each of the seven universal emotions:  anger, fear, sadness, disgust, happiness, surprise, and contempt. The expressions can be viewed at three speeds (slower; normal; faster); however, in the present study, participants were instructed to use the "normal" setting.  After viewing each emotion in the *Learn* section, participants can test their skill in the *Practice* section.  In this section, an expression is flashed on screen and users select the corresponding emotion button on the right of the screen.  If users give an incorrect response, they can replay the emotion once more before making another selection.  If an incorrect selection is made the second time, the correct answer is given.   There are 37 expressions presented in the *Practice* section.  At the end of the practice section, a percentage correct score is displayed.

## Procedure

Participants were assigned a number in order of receipt of expression of interest in participating in the study. Fifty numbers were randomly assigned to two groups using an on-line random sequence generator (www.random.org). Initially the number of participants who registered an interest in the study exceeded 50 and so the excess above 50 were placed on a waitlist. As some participants withdrew due to work commitments, they were replaced by the next person on the waitlist. This system enabled recruitment to continue even after data collection had commenced.

Each participant was mailed a package containing a CD-ROM with the DDERT, a Consent Form (refer Appendix J), the Information Sheet, a demographic questionnaire, and a set of instructions. There were two sets of instructions, depending on whether the participant was assigned to the training or no-training (control) group (refer Appendix K(a) and K(b)). The instruction pertaining to saving the data file was individualized for each participant. The first 34 people to participate in the study were also sent a 55-item questionnaire which was used by a Massey University Honours student, who was investigating emotional competencies of therapists. As not all participants completed the questionnaire and the purpose of it was for use in another study, which was made clear to participants, no data from the questionnaire were analysed for the current study.

Each participant was contacted individually via telephone or email and a time was scheduled for them to participate in the study. The researcher was available to answer any questions by telephone and email at all times whilst the participant completed the study, carefully following the instruction set.

Initially the researcher intended visiting all the participants at their worksite to collect the data. However, as participants from a broad geographical region expressed an interest in participating, this became impractical. The participants all completed the tasks either in their home office or workplace setting on their own Windows-based computer. After completing the consent form and questionnaires, participants completed the DDERT contained on the CD-ROM. Step-by-step instructions were provided which outlined the process in detail. At the commencement of the main trial, participants were instructed to save a file to their desktop which they later emailed to the researcher. The file name contained the participant's number with the suffix (a).

Completing the DDERT involved viewing the 50 video clips created in Stage 1 above, and making a decision about whether each expression was a genuine or masked expression. The instructions stated "You will view a short (5 second) video clip of an emotional expression. This expression may be a genuine expression or a masked expression (a fake expression covering a felt emotion). You may notice more than one emotion in the video clip".

A practice trial containing three clips was presented to orient the participant to the process and the instructions. After each clip was presented, a selection box appeared (refer Figure 6). The first field contained the selection concerning veracity of expression. Participants were asked to select whether they believed the expression was genuine or masked, and to what degree they believed this to be true out of Definitely/Probably/ Possibly. They then selected an emotion term that best matched the emotion they believed they saw expressed. A maximum of two emotions could be selected. After the practice trial, the main trial began and 50 video clips were presented in randomised order. Following each clip, the selection screen appeared.

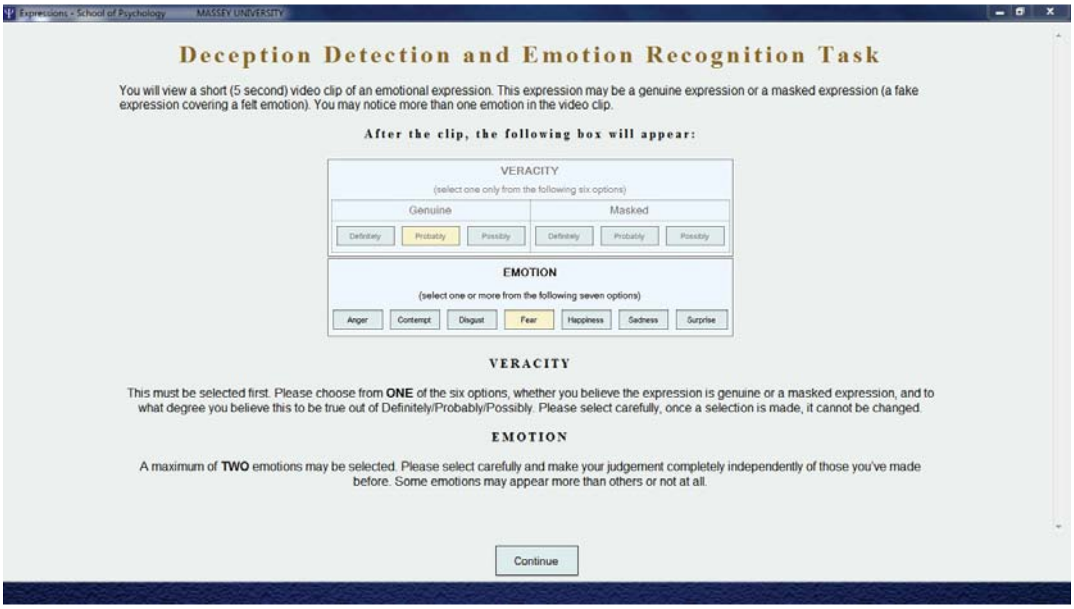



*Figure 6.* Screenshot of selection screen.

### *Training Group Participants*

Twenty two participants were randomly assigned to the Training Group. There were 17 women and five men, aged between 27 and 60 years (mean age = 40 years, *SD* = 9.8). Years of experience as a therapist ranged from one to 31 years (mean experience = 11.0 years; *SD* = 10.0). After completing the first trial of the DDERT, participants were emailed instructions for accessing the SETT. They were instructed to set the speed to "normal" and to complete the "Learn" section for all emotions, followed by the "Practice" section, then repeat both sections. The training procedure took approximately 90 mins.

### *Control Group (No Training)*

Participants assigned to the "No Training" group included 13 women and 8 men aged between 27 and 68 years (mean age = 44.48 years; *SD* = 12.55)[2]. Years of experience practicing as a psychologist ranged from one to 39 years (mean = 13.2 years; *SD* = 11.7). After completing the first trial of the DDERT, participants were instructed to continue "business as usual", for example, complete work administration tasks for 90 min, with no face-to-face client contact during this period.

[2] One participant did not disclose age; the group mean age was substituted for the missing value.

After completing the SETT training and administration tasks (control group), both groups again completed the DDERT. A second data file was saved to the participant's desktop, following the same naming convention as above but with the suffix (b). At the conclusion of all the tasks, these two data files were emailed to the researcher for scoring. Participants in the no-training group were then provided free access to the SETT training to view at their leisure as a gratuity for participating in the study.

This completed the data collection.

# CHAPTER FIVE: DATA ANALYSIS

*Two separate analyses were conducted: Firstly, Hypotheses 1 and 2, the effect of training and experience on emotion recognition, were analysed using mixed within-between analysis of variance (ANOVA). This was followed by analysis of Hypothesis 3, that training would increase accuracy in detecting deception. Data obtained to test the third hypothesis were analysed using signal detection theory measures of performance and ANOVA.*

## OVERVIEW OF THE DATA ANALYSIS

Unless otherwise stated, the following assumptions underlying the use of parametric tests were met:

**Sampling** – scores were obtained from a sample of volunteers from the population of interest, randomised to two groups.

**Independence of observations** – all data were obtained from participants independently of one another.

**Normal distributions** – unless otherwise stated, the data were approximately normally distributed.

**Homogeneity of variance** – unless otherwise stated, the variability of the data in the two groups was homogeneous.

Data were analysed using SPSS for Mac (version 22.0.0.0). The *F* values and associated statistics are reported in the text. The outcome measures used were statistical significance ($p$ = .05), and effect size, which was estimated using partial eta squared ($\eta_p^2$). The following general rules of thumb suggested by Cohen, Cohen, West, and Aiken (2003) for interpreting the effect sizes were adopted: 0.01 (small), 0.09 (medium) and 0.25 (large). Using these values, the study was underpowered having a power of 46% for the sample size used for $p = .05$ (Faul, Erdfelder, Lang, & Buchner, 2007). Where independent samples *t*-tests have been conducted, the results reported are two-tailed. ANOVA tables can be found in Appendix M.

## AGE AND EXPERIENCE LEVEL

As the literature has shown age-related differences in ERA (e.g., Calder et al., 2003; Demenescu et al., 2014; Ruffman et al., 2008) as well as differences in level of experience (Machado et al., 1999), it was important to investigate the effects of Age and Experience Level on emotion recognition and deception detection.

As both Age and Experience are continuous variables that may be influencing the scores on the dependent variables, it would have been preferable to control for these using ANCOVA (Analysis of Covariance). However, a number of the assumptions underlying the use of ANCOVA were not met. (E.g., lack of normality and homogeneity of regression slopes).

To investigate whether Age or Experience level impacted on results, the relationship between these two variables was established using Pearson's product-moment correlation coefficient. There was a strong positive correlation between the two variables ($r = .85$, $N = 43$, $p < .01$). However, this correlation indicates that only about 72% of the variance is shared between these two variables, leaving 28% unexplained. For this reason, ANOVAs with the dependent variables from the Signal Detection analysis were performed for both Age and Experience Level variables. The results revealed that level of experience had only a slightly stronger effect on the dependent variables than did age. Therefore, ANOVAS were conducted for Experience Level only.

An independent samples *t*-test was conducted to compare years of experience and age for the two experimental groups. There was no significant difference ($t(41)$ =

-.49, *p > .05)* for years of experience between the Training group (*M* = 11.02, *SD* = 9.96) and the No-Training group (*M* = 12.62, *SD* = 11.53). There was also no significant difference between the groups for Age ($t(41) = -.1.29$, $p > .05$; Training group *M* = 40.05, *SD* = 9.81, No Training group *M* = 44.48, *SD* = 12.55). The descriptive statistics for the two groups are displayed in Table 2.

Table 2

*Descriptive Statistics Across Experimental Groups*

| | Female | Male | Age range | Age (Years) *M* (*SD*) | Years of Experience *M* (*SD*) |
|---|---|---|---|---|---|
| Training (*n*=22) | 17 | 5 | 27-60 | 40.05 (9.81) | 11.02 (9.96) |
| No Training (*n*=21) | 13 | 8 | 27-68 | 44.48 (12.55) | 12.62 (11.53) |

***Experience Level***

A median split was performed using SPSS to divide the sample into two groups creating a categorical variable - Experience Level, according to years of experience: Low Experience (*M* = 3.16 years, *SD* = 2.42); High Experience (*M* = 20.86, *SD* = 8.01). An independent samples *t*-test showed that the difference between these two means was significant, $t(41) = 9.71$, $p < .001$. There was also a significant difference between the mean age of the two groups, Low Experience *M* = 34.23 years, *SD* = 6.82, High Experience *M* = 50.57 years, *SD* = 8.77, $t(41) = -6.84$, $p < .001$. The descriptive statistics are displayed in Table 3.

Table 3

*Descriptive Statistics for Split by Level of Experience*

| | Female | Male | Age range | Age (Years) *M* (*SD*) | Years of Experience *M* (*SD*) |
|---|---|---|---|---|---|
| Low Experience (*n*=22) | 17 | 5 | 27-53 | 34.23 (6.82) | 3.16 (2.42) |
| High Experience (*n*=21) | 13 | 8 | 36-68 | 50.57 (8.77) | 20.86 (8.01) |

The following analyses included ANOVAs investigating the effects of both Training and Experience Level on the dependent variables of interest, effectively creating four groups (Low Experience – Training; Low Experience – No Training; High Experience – Training; High Experience – No Training).

The descriptive statistics for the groups split by experience level across experimental conditions are shown in Table 4.

Table 4

*Age and Years of Experience for Low and High Experience Groups Across Experimental Conditions*

| | Age range | Age (years) *M* (*SD*) | Years of Experience *M* (*SD*) |
|---|---|---|---|
| Low Experience | | | |
| Training (*n*=12) | 27-46 | 33.67 (6.31) | 3.29 (2.49) |
| No Training (*n*=10) | 27-53 | 34.90 (7.67) | 3.00 (2.45) |
| High Experience | | | |
| Training (*n*=10) | 36-60 | 47.70 (7.51) | 20.30 (6.97) |
| No Training (*n*=11) | 37-68 | 53.18 (9.34) | 21.36 (9.17) |

To check whether there were any significant differences between the years of experience split between the experimental groups, independent samples *t*-tests were conducted. There was no significant difference ($t(20) = .28$, $p > .05$) between years of experience for the Low Experience participants assigned to the Training group, $M = 3.29$, $SD = 2.49$, and the Low Experience participants assigned to the No-Training group, $M = 3.00$, $SD = 2.45$.  Neither was there a significant difference ($t(19) = -.30$, $p > .05$) for years of experience between the High Experience participants assigned to the Training group, $M = 20.30$, $SD = 6.97$, and the High Experience participants assigned to the No-Training group, $M = 21.36$, $SD = 9.17$.

## ANALYSIS OF EMOTION RECOGNITION

The design employed was a 2 x 2 x 2 mixed ANOVA. The first two factors were the between-subjects factors of Training (Training; No Training) and Experience Level (High; Low). The third factor was the within-subjects factor of Time (Time1; Time2). The dependent variable was the percentage of correct judgements made by the participants. The hypothesis tested was that the effects of training would increase the percentage of correct judgements of emotion type made by the participants who received the SETT training compared to those who did not.

Three separate analyses were performed for three types of emotion expression presented. The first emotion type was *Single* expression, where only one emotion expression was displayed in the video clip. The second emotion type was a *Masked* expression. This occurred where the video clip featured an emotion expression that was masking an underlying emotion. For example, the person was displaying fake happiness whilst feeling sad. Therefore, the masked expression was the fake Happy expression. The third emotion type was a *Leaked* expression. Here, the emotion expression was masked by another expression, but leakage of the underlying emotion occurred. For example, the person was displaying fake disgust whilst experiencing sadness and traces of the sad expression could be observed.

The analysis of emotion recognition was to be performed using Emotion as a factor with three levels: Happy, Sad, Disgust. However, a number of the dependent measure scores were not normally distributed, for example, 44.2% of participants scored 100% for correctly identifying a Happy emotion at Time 1, and 48.8% scored 100% at Time 2, creating ceiling effects. This resulted in the distributions being highly negatively skewed. For this reason, each emotion has been analysed in separate ANOVAs.

## ANALYSIS OF DECEPTION DETECTION

Signal detection theory (SDT) (Green & Swets, 1966; McNicol, 1972) was used to analyse the results of the deception detection task as it provides a useful way to measure decision making behaviour occurring under conditions of uncertainty. SDT yields four combinations of stimulus and response events for a set of stimuli consisting of 'signal' and 'noise' items. In the present study, a *signal* item is a falsified facial expression of emotion; a *noise* item is a genuine emotion expression. The four possible response events are 'hits' where a signal is correctly identified as a signal (the participant says it is a fake expression when it is a fake expression) or a 'miss', where a signal is identified as noise (the participant says it is genuine when it is fake). A 'correct rejection' occurs when a genuine expression is correctly identified as genuine, and a 'false alarm' occurs when a genuine expression is incorrectly identified as a fake expression. The total number of hits, divided by the number of signal (fake) items presented, is referred to as the hit rate (HR). The total number of false alarms, divided by the number of noise (genuine) items presented is the false alarm rate (FAR).

There are a number of measures of decision-making performance available with SDT. Firstly, it provides measures of sensitivity; that is, for the present data, how well each participant was able to correctly discern deceptive facial expressions from genuine expressions. The parametric measure of sensitivity, *d´ (d-prime),* is calculated by transforming the HR and FAR to *z* (standard) scores, then subtracting *z*(FAR) from *z*(HR).

As *d´* assumes normal distributions of signal and noise with equal variances for the standard SDT model, a non-parametric measure, *P(A),* was also calculated as an alternative measure of sensitivity. *P(A)* is obtained by plotting the receiver-operating characteristic (ROC) curve, with HR values on the y-axis, and FAR values on the x-axis. (These HR and FAR values were obtained from a rating scale task; see the Method section). The area that falls under the curve defines *P(A)* and can range from 0.5 (50%) where the participant cannot discriminate between signal and noise, to 1.0 (100%) where the participant makes no false alarms and achieves a 100% hit rate (Green & Swets, 1966; McNicol, 1972).

SDT also provides a measure of *bias.* This provides information about whether the participant has a tendency to make judgements conservatively, that is, to respond "fake" only when they are certain a fake expression is presented and they are more likely to judge an item as being genuine. Participants who make more fake than genuine responses have a liberal bias. SDT assumes that bias is independent of sensitivity. The measure for bias is called *c* (for criterion) and is defined as

> $c = -\frac{1}{2}[z(\text{HR})+z(\text{FAR})]$. Values of $c < 0$ represent a liberal bias, and $c > 0$ a conservative bias (McNicol, 1972).

When the *z* of the HR and FAR are plotted, SDT predicts the ROC curve will be a straight line with a slope of 1. This implies that holding a strict criterion provides the same sensitivity as a lax criterion. For example, when holding a strict criterion, the observer will obtain a low HR and low FAR rate. As the criterion becomes more lax, HR increases, but so does the FAR.

As a discriminability index, *d´*, does not depend on the criterion adopted, providing a bias-free measure of sensitivity. However, in the present study, *d´*values were not equal across the ratings; in one case the *d´* value obtained for the last rating was more than 100% larger than the *d´* value for the first rating. Therefore, *d´* was not independent of bias. Furthermore, as a parametric measure, *d´* assumes a model with normal distributions and equal variances, and both of these assumptions were likely violated. For this reason, the non-parametric measure, *P(A),* was used when analysing the results.

# CHAPTER SIX: RESULTS FOR STUDY 1

## EMOTION RECOGNITION

*The statistical results for the emotion recognition task are presented in this section. Firstly, the results for correctly identifying the emotion type are presented. These results are presented separately for each expression type: Single expression, Masked expression, and Leaked expression.*

Mean percentage scores for both groups were calculated for Time 1 and Time 2 for each of the expression types, and for each emotion category. Table 5 presents the means and standard deviations for each of the three categories of emotion expression.

Table 5

*Mean Percentage Scores and Standard Deviations (in Parentheses) for All Emotion Expressions for Time 1 and Time 2 for Each Group (Training vs No-Training) and Experience Level (High vs Low)*

| | Training | | | | No Training | | | |
|---|---|---|---|---|---|---|---|---|
| Category | Low experience *M* (*SD*) | | High experience *M* (*SD*) | | Low experience *M* (*SD*) | | High experience *M*(*SD*) | |
| | T1 | T2 | T1 | T2 | T1 | T2 | T1 | T2 |
| Single Expression | | | | | | | | |
| Happy | .94 (.08) | .94 (.05) | .89 (.14) | .97 (.04) | .94 (.07) | .97 (.07) | .88 (.15) | .87 (.16) |
| Sad | .55 (.22) | .72 (.17) | .59 (.22) | .77 (.12) | .53 (.27) | .65 (.17) | .40 (.17) | .50 (.25) |
| Disgust | .68 (.16) | .63 (.15) | .60 (.14) | .67 (.22) | .65 (.18) | .67 (.20) | .62 (.17) | .54 (.16) |
| Masked Expression | | | | | | | | |
| Happy | .83 (.25) | .87 (.31) | .95 (.16) | .95 (.16) | .90 (.21) | 1.0 (.00) | .77 (.34) | .90 (.20) |
| Sad | .31 (.23) | .47 (.22) | .50 (.25) | .55 (.26) | .35 (.29) | .38 (.25) | .26 (.27) | .38 (.27) |
| Disgust | .64 (.16) | .64 (.15) | .60 (.11) | .65 (.28) | .69 (.14) | .73 (.13) | .44 (.24) | .39 (.15) |
| Leaked Expression | | | | | | | | |
| Happy | .09 (.09) | .10 (.10) | .08 (.09) | .09 (.06) | .09 (.06) | .08 (.07) | .06 (.09) | .10 (.11) |
| Sad | .14 (.10) | .17 (.10) | .22 (.19) | .15 (.09) | .12 (.11) | .14 (.07) | .11 (.11) | .14 (.18) |
| Disgust | .50 (.37) | .42 (.36) | .50 (.33) | .45 (.37) | .55 (.27) | .50 (.41) | .55 (.27) | .50 (.45) |

## Single Emotion Expressions

As described earlier, a single emotion was an expression coded by FACS as displaying only the action units (AUs) associated with one specific emotion.

The dependent variable was participants' total accuracy scores (%) for correctly identifying the emotion expressed. As the Deception Detection and Emotion Recognition Task (DDERT) allowed for participants to select up to two emotion terms, this provided an increased opportunity to select the correct expression by selecting two emotion terms when only a single emotion was displayed. A preliminary analysis was conducted on the frequency of participants selecting two emotion terms when only a single emotion was presented. The total number of dual selections was divided by 33 (the number of clips displaying only one expression) to obtain a percentage score. All participants selected two emotion terms at least once when viewing a single emotion. The number of dual selections was calculated for each group revealing there was no difference between groups in the occurrence of dual selection at either Time 1 or Time 2 (for both groups, $M$ = .21 and $M$ = .20, respectively).

***Happy***

The scores for correctly identifying a Happy expression were not normally distributed as can be seen in the histograms in Figure 7. This was likely due to the task of correctly identifying a Happy expression being too easy, despite pilot work suggesting otherwise. Several different types of transformation were carried out with the data, but all of them left the distributions in a highly skewed state. Nevertheless, an ANOVA was still conducted using the percentage of correct decisions for the Happy emotion. The results of this analysis are reported below but caution is required when interpreting the results.

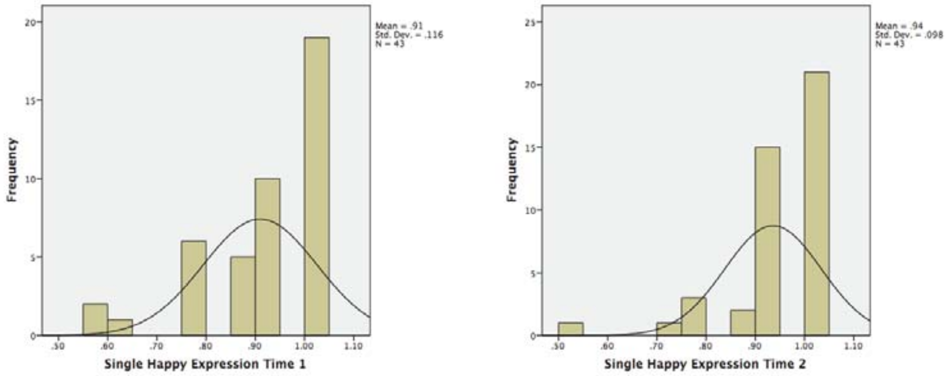



*Figure 7.* Distribution of scores for correct identification of a single Happy expression at Time 1 and Time 2 for all participants. Note that the frequency ranges on the ordinates of the two plots are different.

The mean scores for each group are shown in Table 5 and graphically presented in Figure 8, showing that for the participants with Low Experience, Training had no effect, with scores remaining static from Time 1 to Time 2 ($M$ = .94). The Low Experience (No Training) group's scores increased slightly from Time 1 ($M$ = .94) to Time 2 ($M$ = .97). However, for participants with High Experience, Training increased scores by 8 percentage points from Time 1 ($M$ = .89) to Time 2 ($M$ = .97). The scores for the High Experience (No Training) group remained much the same from Time 1 ($M$ = .88) to Time 2 ($M$ = .87).

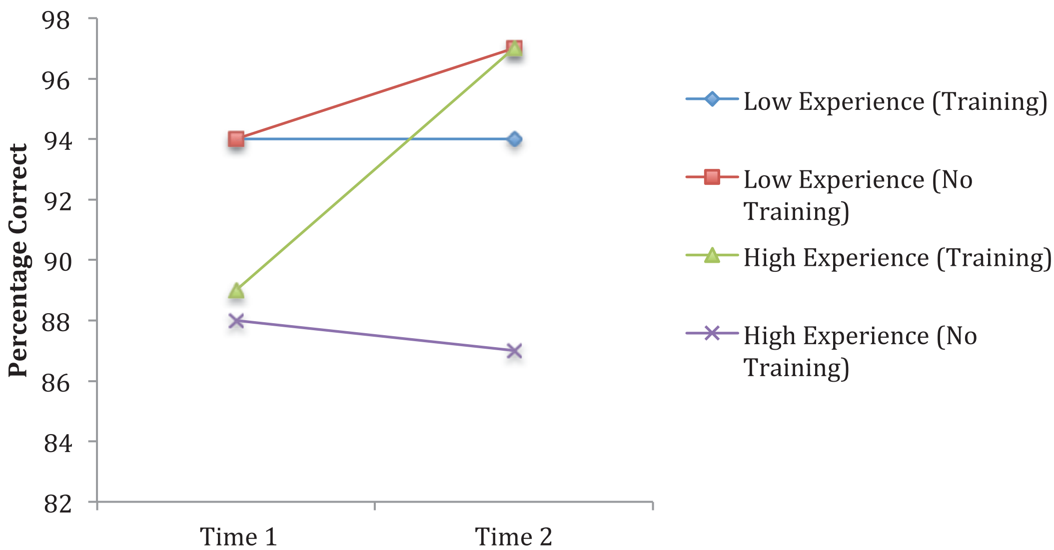



*Figure 8.* Change in scores for identifying a single Happy expression from Time 1 to Time 2 for all groups.

The main effect of Time on accurate recognition of a single Happy expression was not statistically significant, but there was a moderate effect size, $F(1,39) = 3.53$, $p = .07$, $\eta_p^2 = .08$, reflected by the change in scores from Time 1 to Time 2 for the High Experience Training group, and the untrained Low Experience group, showing a practice effect for these two groups.

There was no interaction between Time and Training, or between Time and Experience Level, both *F*s < 1. Therefore, Training alone did not affect any change in scores. However, there was a near significant three-way interaction between Time, Training, and Experience Level, with a moderate effect size, $F(1,39) = 3.60$, $p = .07$, $\eta_p^2 = .09$. Figure 8 shows that the main component of this interaction was the effect of training on the high experience participants.

***Sad***

Figure 9 reveals that Training increased emotion recognition for Sadness, for both Experience Level groups. The Low Experience (Training) group's mean score increased by 17% from Time 1 ($M = .55$) to Time 2 ($M = .72$). The High Experience (Training) group's mean score increased by 18% from Time 1 ($M = .59$) to Time 2 ($M = .77$). The Low Experience and High Experience (No Training) groups' mean scores increased by 12% and 10%, respectively, from Time 1 to Time 2, suggesting a small practice effect.

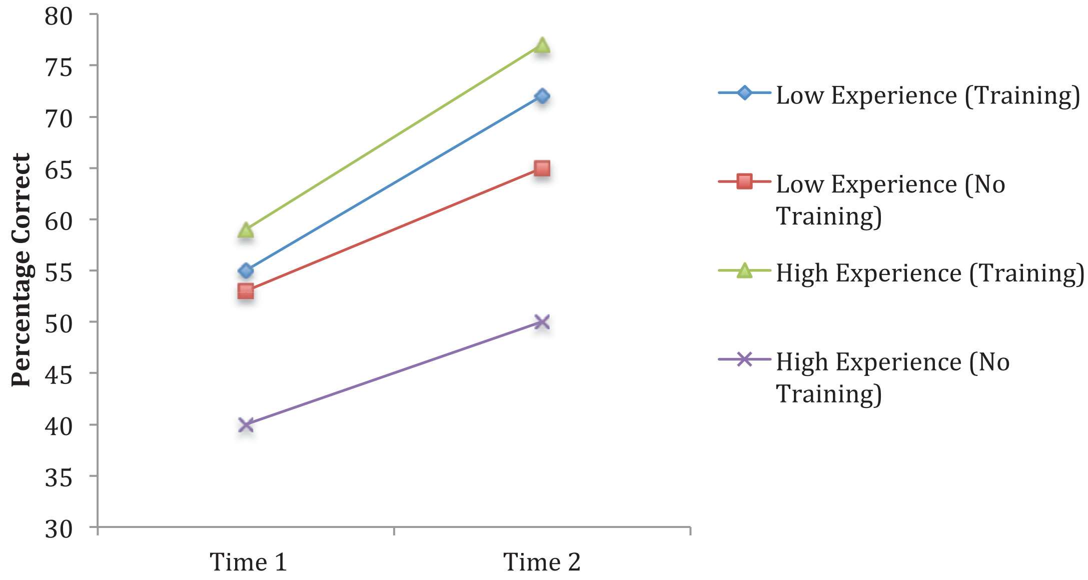



*Figure 9.* Change in scores for identifying a single Sad expression from Time 1 to Time 2 for all groups.

The main effect of Time on accurate recognition of a Sad expression was statistically significant, $F(1,39) = 27.84$, $p < .001$, $\eta_p^2 = .42$, a large effect size. There was a significant main effect of Training, $F(1,39) = 5.95$, $p = .02$, $\eta_p^2 = .13$ with a medium-large effect size. Therefore, while training had some effect, a sizeable amount of the change from Time 1 to Time 2 was likely due to a practice effect. There was no main effect of Experience Level, and no interaction between Time and Experience Level, or between Time, Training and Experience Level, all *F*s < 1.

This is similar to the results found for Happy expressions where it appears that Training had some effect, depending on Experience Level. However, unlike the results for Happy expressions where Training only improved the performance of participants with High Experience, Training was effective for increasing recognition of Sad expressions for both Low and High Experience groups. There was no effect of Experience Level, $F < 1$.

***Disgust***

The results for Disgust are displayed graphically in Figure 10 showing that Training increased the scores for the High Experience group but had little effect on the scores of participants in the Low Experience group. The High Experience (Training) group had an increase in scores from Time 1 (*M* = .60) to Time 2 (*M* = .67) but those that did not receive training had a decrease in scores from Time 1 (*M* = .62) to Time 2 (*M* = .54). Training appeared to have a negative effect on the Low Experience group with scores declining from Time 1 (*M* = .68) to Time 2 (*M* = .63) whereas those in the Low Experience (No Training) group showed little change from Time 1 to Time 2 (*M* = .65 and *M* = .67, respectively). It would seem, therefore, that with regards to detecting disgust, training is most beneficial for participants with a high level of experience.

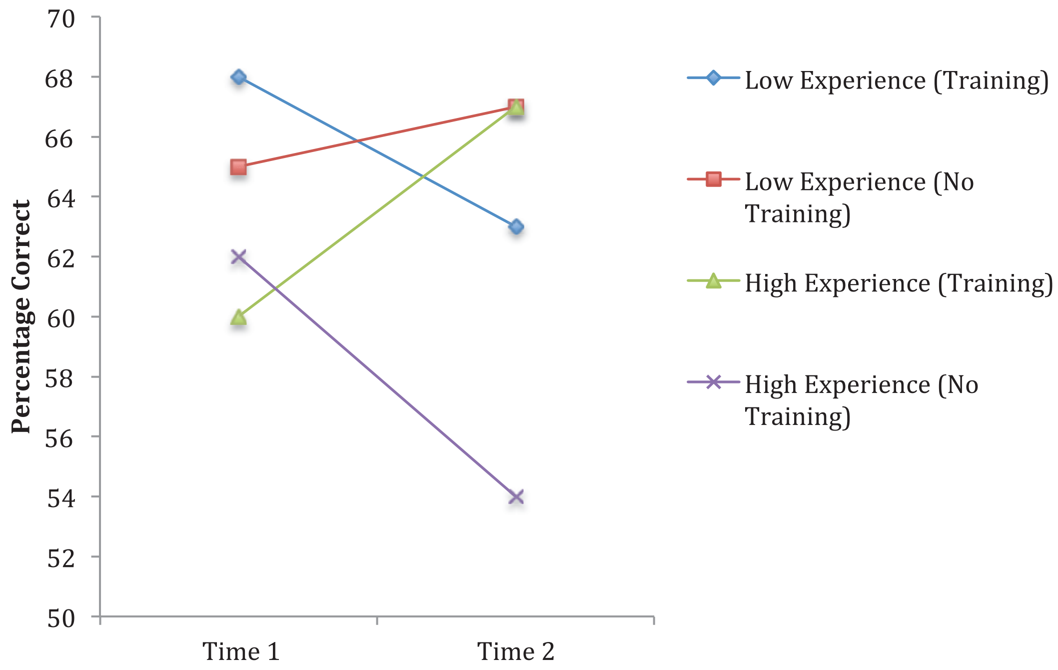



*Figure 10.* Change in scores for identifying a single Disgust expression from Time 1 to Time 2 for all groups.

The main effect of Time on accurate recognition of a Disgust emotion was not significant, $F < 1$, and there was no interaction between Time and Training or Time by Experience Level, both *F*s < 1.

There was a marginal three-way interaction between Time, Training, and Experience Level, $F(1,39) = 3.25$, $p = .08$, $\eta_p^2 = .08$, with a moderate effect size. However, Experience Level had no effect, $F(1,39) = 1.36$, $p = .25$, $\eta_p^2 = .03$. Figure 10 clearly shows that most of the variance in the 3-way interaction was due to the effect of training on the high experience people. So the main effect of Experience was qualified by the 3-way interaction.

## Masked emotion expressions

### *Happy*

As with the single Happy expressions, the scores for correctly identifying a masked Happy expression were not normally distributed as can be seen in the histograms in Figure 11. Despite this, ANOVAs for masked Happy expressions were still conducted, but caution is required when interpreting results.

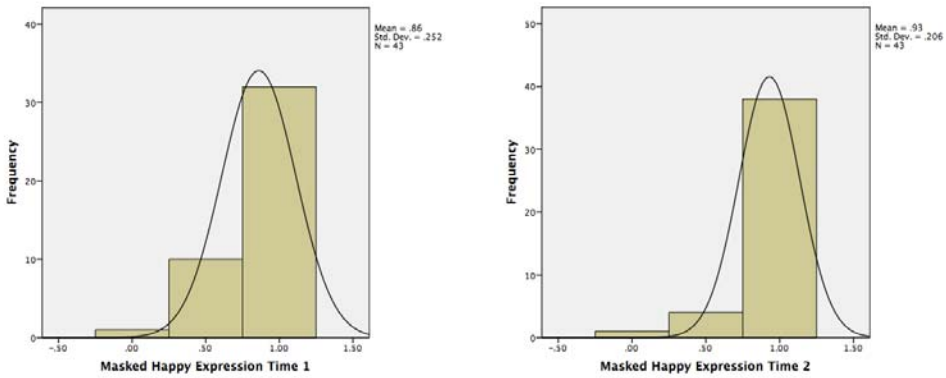



*Figure 11.* Distribution of scores for correct identification of a Masked Happy expression at Time 1 and Time 2 for all participants. Note that the frequency ranges on the ordinates of the two plots are different.

The mean scores for accurately detecting a masked Happy expression for each group are shown in Table 5 and graphically in Figure 12. For participants with Low Experience, Training had only a small effect, increasing scores by 4% from Time 1 ($M$ = .83) to Time 2 ($M$ = .87). For those with Low Experience who did not receive training, their scores showed a larger increase (10%) from Time 1 ($M$ = .90) to Time 2 ($M$ = 1.00). The standard deviation at Time 2 (0.00) was attenuated by a ceiling effect as all participants in this group scored 100%. For the High Experience (Training) group, scores remained constant from Time 1 to Time 2 ($M$ = .95), and for those who did not receive Training, the greatest score increase was observed (13%) from Time 1 ($M$ = .77) to Time 2 ($M$ = .90). However, the Training group's score at

Time 1 was *M* = .95, therefore a ceiling effect may have limited the scope of improvement.

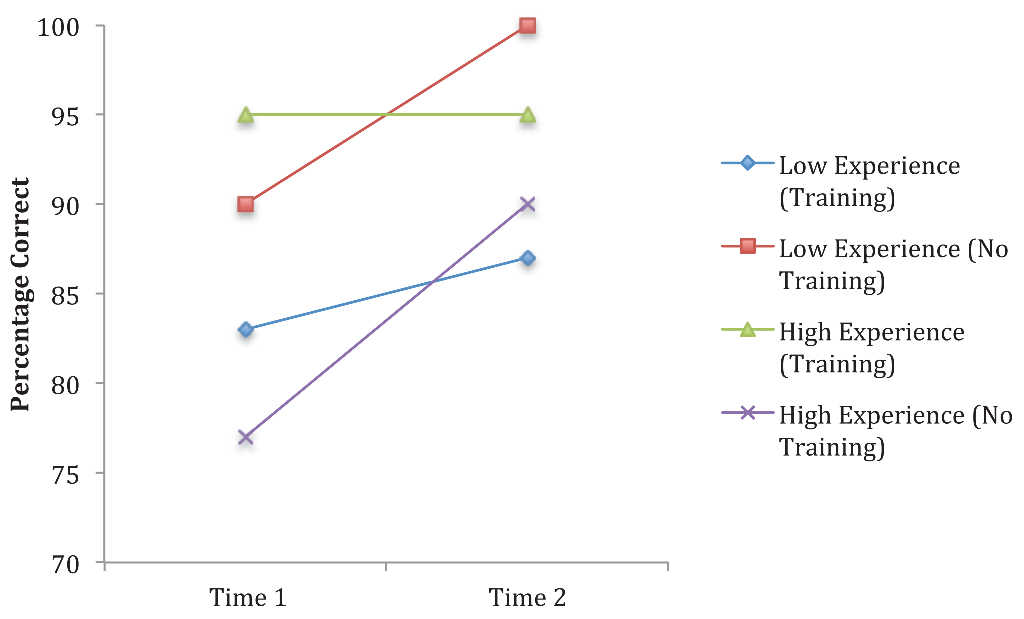



*Figure 12.* Change in scores for identifying a masked Happy expression from Time 1 to Time 2 for all groups.

However, the increase in scores from Time 1 to Time 2 was greater for both groups that did not receive training. This suggests that although there was an increase in scores from Time 1 to Time 2, this was a spurious difference, by chance, rather than an effect of the intervention. No interaction was noted between Time and Experience Level, or between Time, Training and Experience Level.

There was no main effect of Training or Experience Level on accurately identifying a masked Happy expression; however, there was a near significant interaction between Training and Experience Level, $F(1,39) = 3.76$, $p = .06$, $\eta_p^2 = .09$. This suggests that Training has a moderate effect, conditional upon level of experience. It appears that for masked Happy emotions, the Low Experience group received the most benefit from training.

***Sad***

Figure 13 shows that scores increased for all groups for accurately recognising a masked Sad expression. The Low Experience (Training) group's mean score increased by 16% from Time 1 (*M* = .31) to Time 2 (*M* = .47). The High Experience (Training) group's mean score increased by 5% from Time 1 (*M* = .50) to Time 2 (*M* = .55). The Low Experience and High Experience (No-Training) groups' mean scores increased by 3% and 12%, respectively, from Time 1 to Time 2, suggesting a small practice effect. Although the High Experience (Training) group showed a smaller increase at Time 2 compared with the High Experience group who did not receive training, an independent samples *t*-test showed that the difference between their scores at Time 1 was significant ($t(19) = 2.12$, $p$=.05). This may have contributed to the smaller increase in the Training group's scores as there was less scope for an increase in performance.

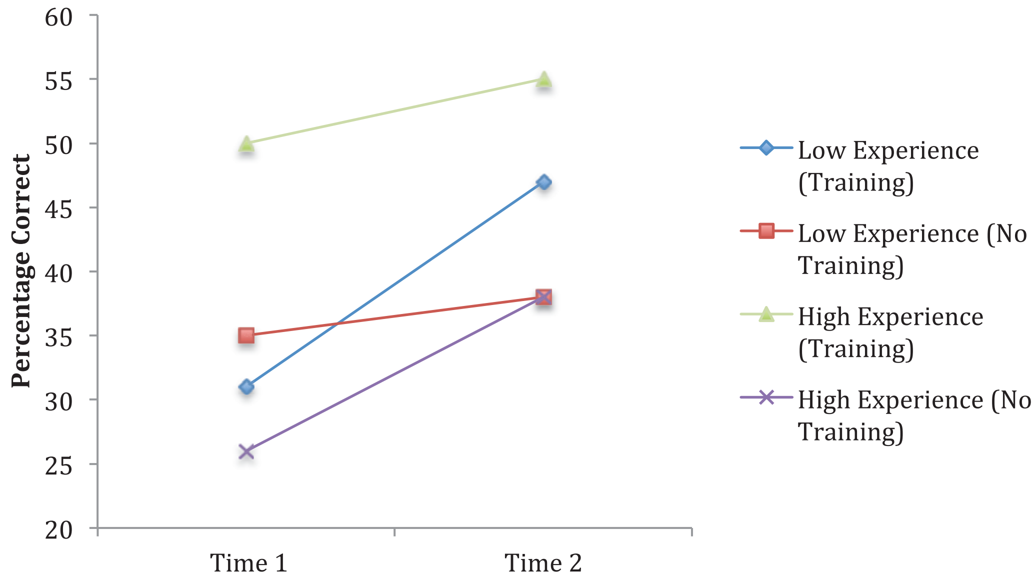



*Figure 13.* Change in scores for identifying a masked Sad expression from Time 1 to Time 2 for all groups.

The main effect of Time on accurate recognition of a masked Sad expression was statistically significant, $F(1,39) = 4.77$, $p < .04$, $\eta_p^2 = .11$ with a medium effect size. However, there was no interaction between Time and Training, or Time and Experience Level, both *F*s < 1.

There was no main effect of Experience Level, *F* < 1.

***Disgust***

The results for masked Disgust (see Figure 14) show that Training effected the greatest increase in scores for the High Experience group, Time 1 (*M* = .60), Time 2 (*M* = .65), but had no effect on the scores of participants in the Low Experience group which remained static from Time 1 to Time 2, (*M* = .64). The scores for the High Experience (No Training) group decreased by 5% from Time 1 (*M* = .44) to Time 2 (*M* = .39), whereas scores increased by 4% from Time 1 (*M* = .69) to Time 2 (*M* = .73) for the Low Experience group that did not receive training.

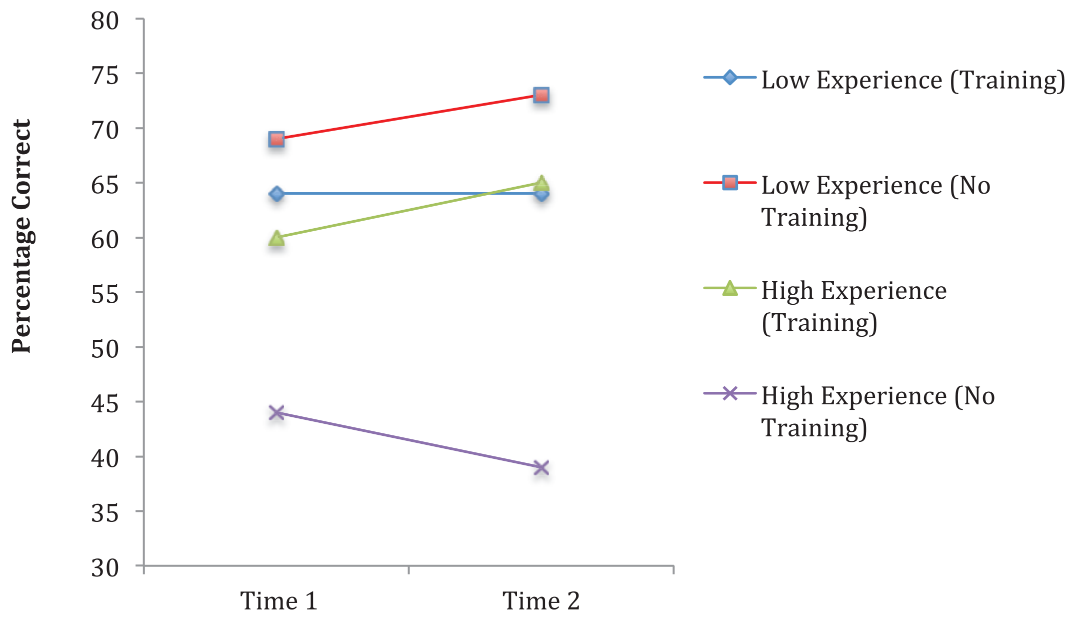



*Figure 14.* Change in scores for identifying a masked Disgust expression from Time 1 to Time 2 for all groups.

There was no main effect of Time, and no interaction between Time and Training, or Time and Experience Level, with all *F*s < 1.

There was a significant main effect of Experience Level, $F(1,39) = 12.42$, $p < .01$, $\eta_p^2 = .24$, yielding a large effect. However, a significant three-way interaction between

Time, Training, and Experience Level, $F(1,39) = 10.93$, $p < .01$, $\eta_p^2 = .22$, qualified this result. It would seem, therefore, that similar to recognition of a single Disgust expression, training is most beneficial for participants with a high level of experience[3].

## Leaked Emotion Expressions

### *Happy*

The mean scores for accurately detecting a leaked Happy expression for each group are shown in Table 5 and are visually displayed in Figure 15. They reveal only a small effect of Training (1% increase) from Time 1 (*M* = .09, *M* = .08) to Time 2 (*M* = .10, *M* = .09), for both Low and High Experience groups. The Low Experience (No Training) group showed a 1% decrease from Time 1 (*M* = .09) to Time 2 (*M* = .08). The High Experience (No Training) group had the greatest increase (4%) from Time 1 (*M* = .06) to Time 2 (*M* = .10).

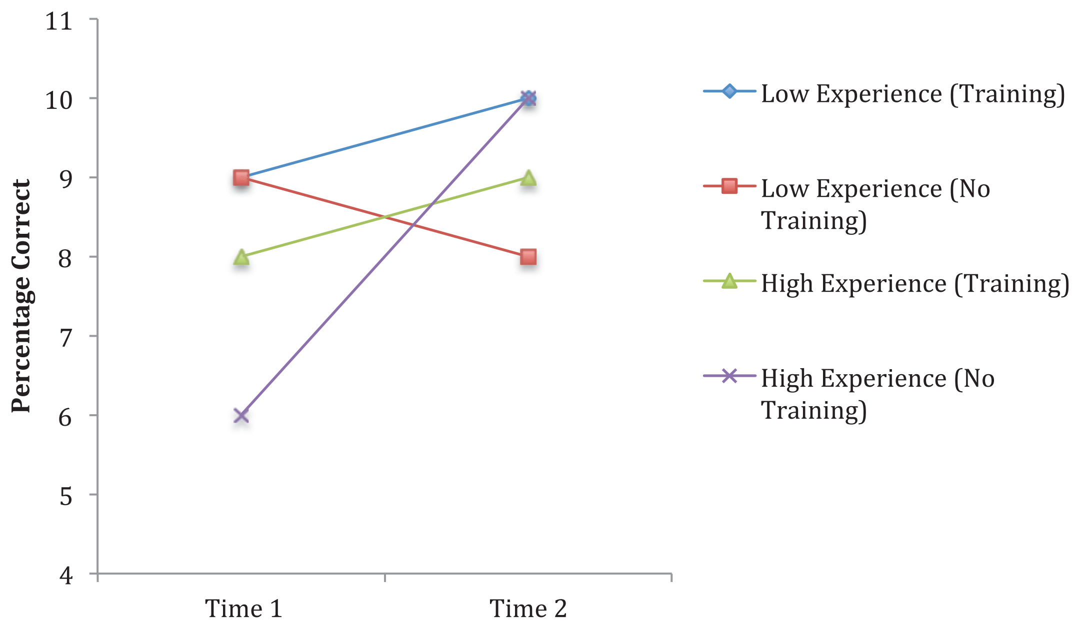



*Figure 15.* Change in scores for identifying a leaked Happy expression from Time 1 to Time 2 for all groups.

---

[3] It can be noted that the High Experience, No Training group at Time 1 performed considerably poorer than the other three groups for recognition of single sadness, and masked disgust. Since participants were randomly allocated, this is a chance effect.

There was no main effect of Time on accurate recognition of a leaked Happy expression, no interaction between Time and Training, no interaction between Time, Training and Experience Level, all *Fs* < 1.

There were no main effects of Training or Experience Level, and no interaction between Training and Experience Level, all *F*s < 1.

***Sad***

The assumptions of normality and equal variance were met for scores for detecting leaked Sad expressions except the Training group's scores showed a slight positive skew at Time 1, 1.63, *SE* = .49, and a kurtosis value of 4.87, *SE* = .95. The No-Training group's scores showed some kurtosis at Time 2, 1.12, *SE* = .97. The Levene's test of homogeneity of variance was also significant at Time 2, suggesting that training reduced the variability amongst participants.

Figure 16 shows that Training increased emotion recognition ability for leaked Sadness, only for those with Low Experience, from Time 1 (*M* = .14) to Time 2 (*M* = .17). For the High Experience group, Training reduced their scores from Time 1 (*M* = .22) to Time 2 (*M* = .15). Both Low and High Experience Level groups which did not receive Training also increased their scores from Time 1 (*M* = .12, *M* = .11) to Time 2 (*M* = .14, *M* = .14).

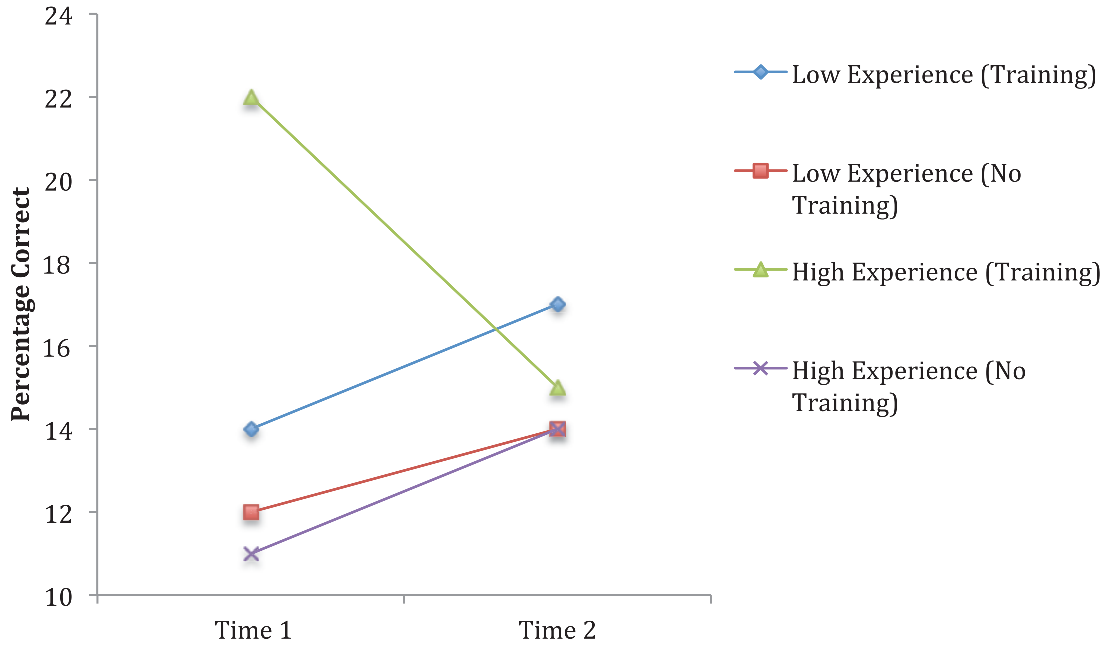



*Figure 16.* Change in scores for identifying a leaked Sad expression from Time 1 to Time 2 for all groups.

There was no main effect of Time on accurate recognition of a leaked Sad expression, and no interactions between Time and Training, or Time and Experience Level, and no three-way interaction between Time, Training, and Experience Level, all Fs < 1.

There was no main effect of Experience Level, and no interaction between Training and Experience Level, both *F*s < 1.

***Disgust***

Figure 17 displays the change in mean scores for leaked disgust for all groups from Time 1 to Time 2. All of the scores decreased, with the greatest decrease occurring for the Low Experience (Training) group (8% decrease). All other groups demonstrated a 5% decrease from Time 1 to Time 2. This suggests Training has no effect on improving accuracy for recognising leaked Disgust.

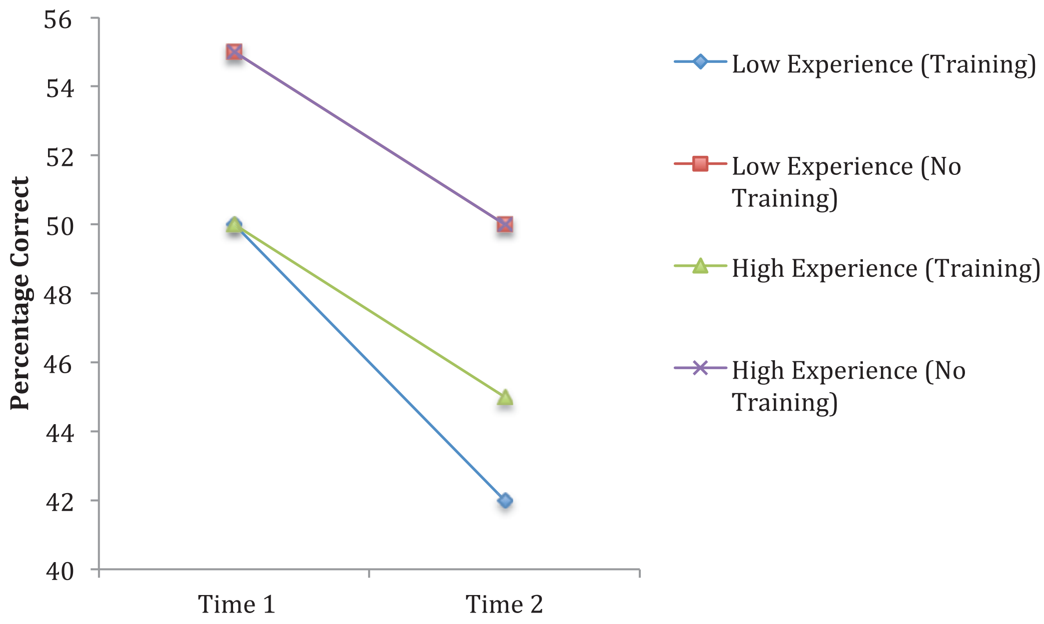



*Figure 17.* Change in scores for identifying a leaked Disgust expression from Time 1 to Time 2 for all groups. Note, the scores for the untrained High and Low Experience groups are equal.

The ANOVA analysis revealed no main effects of Time, and no interactions between Time and Training, or Time and Experience Level, and no three-way interaction between Time, Training, and Experience Level, with all *F*s < 1.

There were no main effects of Training or Experience Level, and no interaction between Training and Experience Level.

These results confirm that Training and Experience Level have no effect on recognition of a leaked Disgust expression. However, it is noted that the scores for recognising leaked Disgust were considerably higher than those for recognising leaked Happiness or Sadness. This would suggest that even without training, leaked Disgust is easier to detect than the other two emotions.

## SUMMARY OF RESULTS FOR EMOTION RECOGNITION

Training was found to be effective for the recognition of single emotion expressions, as predicted by Hypothesis 1. The effect was particularly salient for enhancing recognition of sad expressions. However, these results were largely influenced by experience level and in most cases training was most effective for those with a high level of experience, with the exception of sadness. For this latter emotion, both groups benefitted from training for a single sad expression, and for masked sadness, less experienced participants demonstrated greater improvement following training. A smaller effect was found for improving the recognition of happy expressions, but it is possible that this was limited by a ceiling effect. With regard to expressions of disgust, for all three expression types (Single, Masked, Leaked), although training did not significantly increase recognition, those who did not receive training demonstrated a decline in disgust recognition at Time 2. This could possibly be a chance effect.

With regard to masked and leaked emotions, no effect of training was evident.

Mixed support was found for Hypothesis 2, dependent on emotion type. For single happy expressions, the High Experience group scored lower than the Low Experience group prior to training, but the two groups showed no difference following training. For single sad expressions, the High Experience group that received training achieved similar scores to the Low Experience groups; however, the untrained High Experience group achieved the lowest scores overall. A similar result was found for single disgust expressions, whereby the High Experience group performed similarly to the Low Experience group after training, whereas the High Experience untrained group performed the worst.

For masked expressions, all groups performed similarly for masked happy expressions. For sadness, the High Experience group achieved the highest scores after training compared to the Low Experience group. Both untrained groups achieved the lowest scores, but there was some evidence of a practice effect for the High Experience group. For masked disgust, all groups performed similarly, except for the High Experience untrained group, which achieved the lowest scores.

With regard to leaked expressions, there was little difference between groups for recognising leaked happiness or disgust. For sadness, the High Experience group performed best prior to training, but the Low Experience group scored higher after training.

Overall, training was most effective for improving recognition of single emotion expressions, particularly sadness. With the exception of sadness, training was also more effective for participants with a higher level of experience. However, training did not increase the performance of the high experience group to a level superior to those with low experience, but rather appeared to improve their performance to a level similar to that of their less experienced counterparts.

## DECEPTION DETECTION

*The statistical results for the deception detection task are presented in this section. Firstly, results for the measure of sensitivity P(A) is presented, representing participants' accuracy in detection deception. This is followed by an examination of the results for Hit Rates (HR), False Alarms (FAR), and Response Bias (c).*

### Signal Detection

A Group (Training vs No-Training) x Experience Level (Low Experience vs High Experience) x Time (Time 1, Time 2) ANOVA was carried out for each variable obtained from the Signal Detection analysis. This included the sensitivity measure *P(A*); Hit Rate (HR); False Alarm Rate (FAR); and Response Bias (*c*). All assumptions of ANOVA were met unless otherwise indicated. The descriptive results are presented in Table 6.

Table 6

*Mean Group Values for Time 1 and Time 2 for P(A), HR, FAR, and c, for Low Experience and High Experience across Experimental Groups*

| | Training | | | | No Training | | | |
|---|---|---|---|---|---|---|---|---|
| | Low experience *M* (*SD*) | | High experience *M* (*SD*) | | Low experience *M* (*SD*) | | High experience *M* (*SD*) | |
| | T1 | T2 | T1 | T2 | T1 | T2 | T1 | T2 |
| *P(A)* | .64 (.06) | .63 (.09) | .66 (.07) | .69 (.09) | .60 (.09) | .58 (.10) | .62 (.09) | .65 (.06) |
| HR | .43 (.17) | .41 (.11) | .45 (.12) | .46 (.09) | .47 (.20) | .46 (.14) | .49 (.14) | .53 (.18) |
| FAR | .26 (.13) | .26 (.14) | .27 (.12) | .26 (.14) | .37 (.12) | .33 (.14) | .33 (.15) | .32 (.15) |
| *C* | .45 (.46) | .45 (.25) | .39 (.30) | .41 (.28) | .22 (.37) | .29 (.34) | .25 (.32) | .21 (.50) |

*Note.* Low experience (training) *n*=12; High experience (training) *n*=10; Low experience (no-training) *n*=10; High experience (no-training) *n*=11.

## Sensitivity - P(A)

As shown in Figure 18, the scores for both High Experience groups have increased by 3% from Time 1 (*M* = .66, *M* = .62) to Time 2 (*M* = .69, *M* = .65) for Training and No-training, respectively. However, for the Low Experience groups, scores showed a slight decrease from Time 1 (*M* = .64, *M* = .60) to Time 2 (*M* = .63, *M* = .58), irrespective of whether Training occurred or not. What appears to have a (small) impact on scores at Time 2 is Experience Level rather than Training. That is, the High Experience participants improved in sensitivity from Time 1 to Time 2, irrespective of Training, and Low Experience participants' sensitivity declined from Time 1 to Time 2, irrespective of Training.

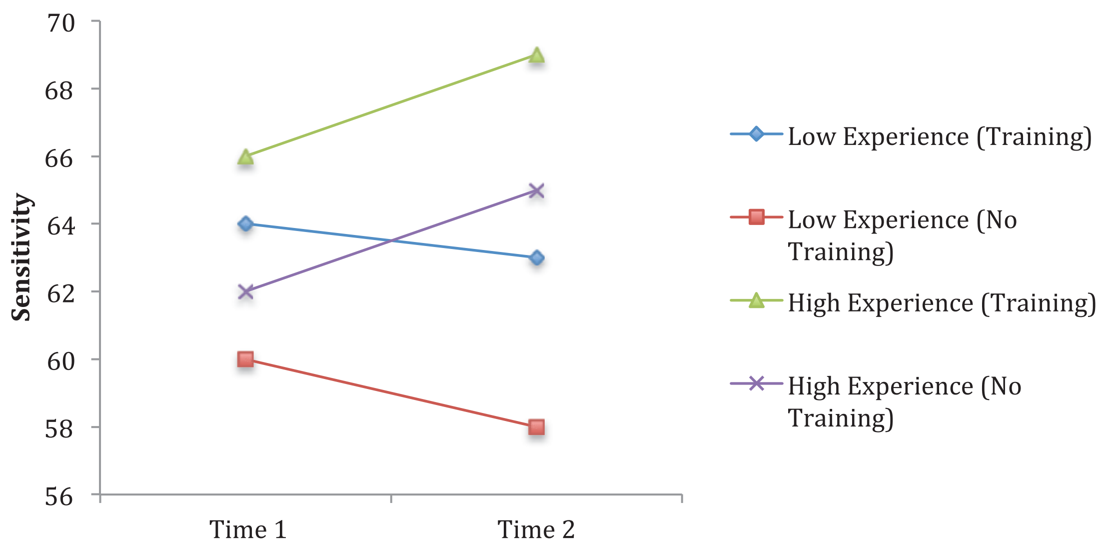



*Figure 18.* Indices of sensitivity as measured by *P(A)* for all groups for Time 1 and Time 2.

There was no main effect of Time, and no interaction between Time and Training, both *F*s < 1. Therefore, Training alone did not effect any change in scores. There was no three-way interaction between Time, Training, and Experience Level, $F < 1$.

There was a near significant main effect for Training, $F(1,39) = 3.73$, $p = .06$, $\eta_p^2 = .09$, indicating a medium effect for the difference between groups. However, as there was no interaction between Time and Training, this most likely represents a chance difference between the groups, not an effect of the intervention. There was also a near significant main effect for Experience Level, $F(1,39) = 3.91$, $p = .06$, $\eta_p^2 = .09$, representing a medium effect of Experience. There was no interaction between Training and Experience, $F < 1$. It would therefore appear that it is experience practicing as a psychologist, not training, that contributes to the development of skills in deception detection.

A further analysis examined the numbers of individuals in each group who scored very high or very low. Table 7 shows four levels of accuracy scores for both Training and No-Training groups, ranging from below chance (<50), chance (50-59), above chance (60-69), and high accuracy (≥70).

Table 7

*Percentage of Participants in Each Group Achieving Different Deception Detection Accuracy Levels*

| | Training | | No Training | |
|---|---|---|---|---|
| | T1 | T2 | T1 | T2 |
| <50 | - | 4.5 | 14.3 | 9.5 |
| 50-59 | 22.7 | 22.8 | 33.3 | 28.6 |
| 60-69 | 54.6 | 27.2 | 52.4 | 42.9 |
| ≥70 | 22.7 | 45.5 | 14.3 | 19.0 |

As Table 7 shows, the group that received Training had a larger proportion of participants scoring in the high accuracy range at Time 2, 45.5% compared with 19% (No Training). To determine whether experience level impacted on the improvement of accuracy, the numbers in each group were examined and are shown in Table 8.

Table 8

*Percentage of Participants in Each Group by Experience Level Achieving Different Deception Detection Accuracy Levels*

| | Low Experience | | | | High Experience | | | |
|---|---|---|---|---|---|---|---|---|
| | Training | | No Training | | Training | | No Training | |
| | T1 | T2 | T1 | T2 | T1 | T2 | T1 | T2 |
| < 50 | 0 | 8.3 | 10 | 20 | 0 | 0 | 18.2 | 0 |
| 50-59 | 25 | 25 | 30 | 30 | 30 | 30 | 9.1 | 27.3 |
| 60-69 | 58.3 | 33.4 | 50 | 30 | 40 | 20 | 54.5 | 54.5 |
| ≥70 | 16.7 | 33.3 | 10 | 20 | 30 | 50 | 18.2 | 18.2 |

When comparing by experience level, it can be seen that in the Low Experience Training group, one-third (33.3%) of participants achieved a high level of accuracy after Training, compared with one-half (50%) of the High Experience participants.

What appears to have occurred is an improvement in scores following training for those who scored in the 60-69 range. The percentage of people scoring in this range reduced at Time 2, possibly due to them moving to the next accuracy bracket, whilst the numbers of those scoring in the 50-59 range showed little change. Caution needs to be exercised in interpreting these findings due to the low numbers in each group, resulting in apparently large shifts in percentage changes.

***Hit Rates (HR)***

As displayed in Table 6, there is little change in Hit Rates from Time 1 to Time 2. However, it is noted that for both the High Experience groups, Hit Rates increased slightly from Time 1 to Time 2, as would be expected with an increase in *P*(A).

There was no main effect of Time, no interaction between Time and Training, or Time and Experience Level, and no three-way interaction between Time, Training, and Experience Level, all *F*s < 1.

There was no interaction between Training and Experience Level, *F* < 1.

It appears that Experience Level, rather than Training, has the greatest effect as the Low Experience participants showed a decline in HR from Time 1 to Time 2, irrespective of Training. The High Experience participants who received training remained static in their scores from Time 1 to Time 2, whereas those who did not receive training increased by 4% from Time 1 (*M* = .49) to Time 2 (*M* = .53).

***False Alarm Rates (FAR)***

There was no main effect of Time, no interaction between Time and Training, or Time and Experience Level, and no three-way interaction between Time, Training, and Experience Level, all *Fs* < 1.

There was a significant main effect of Training, $F(1,39) = 4.42$, $p = .04$, $\eta_p^2 = .10$. However, the differences between the Training and No Training group only appear evident for the Low Experience (No Training) group whose FAR decreased by 4% from Time 1 (*M* = .37) to Time 2 (*M* = .33). Table 6 reveals that all three other groups remained fairly static in their FAR scores from Time 1 to Time 2.

***Response Bias (c)***

Figure 19 displays the differences in response bias between the groups. Both the Low Experience and High Experience (Training) groups exhibited a more conservative response bias compared with those that did not receive Training. However, these differences were apparent at Time 1 as well as Time 2, and as all participants were assigned randomly, have most likely occurred by chance.

The Low Experience (Training) group showed no change in bias from Time 1 to Time 2. The High Experience (Training) group showed a slightly more conservative bias at Time 2, whereas the Low Experience (No Training) group's bias changed by 7% in favour of a more conservative approach. That is, they were less likely to say fake at Time 2 than they were at Time 1. However, the High Experience (No Training) group displayed a tendency to be more liberal at Time 2.

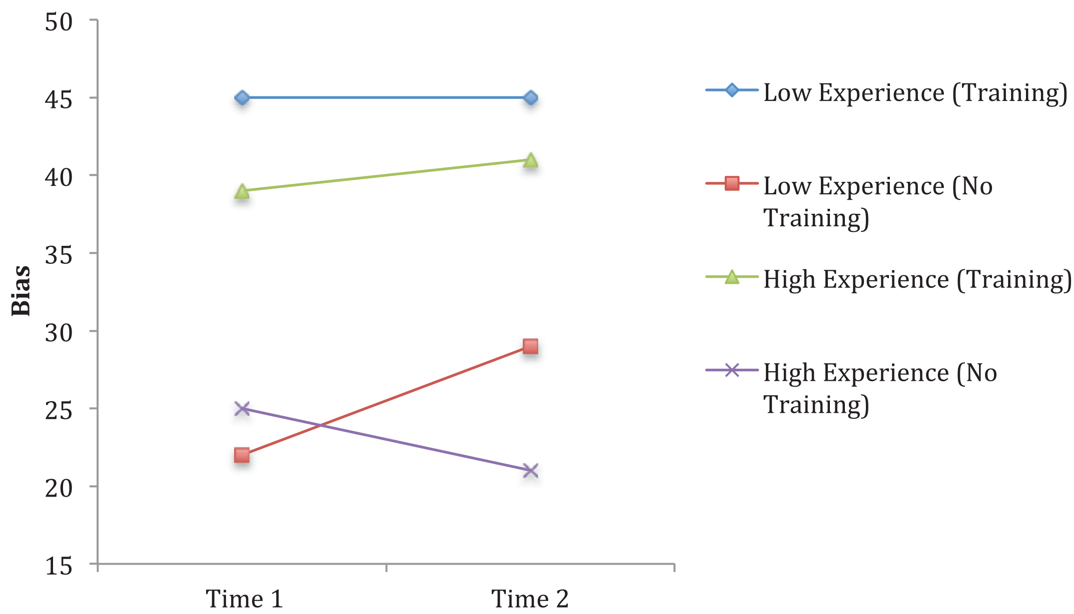



*Figure 19.* Change in response bias from Time 1 to Time 2 for all groups.

There was no main effect of Time, no interaction between Time and Training, or Time and Experience Level, and no three-way interaction between Time, Training, and Experience Level, all *Fs* < 1.

There was a near significant main effect for Training, $F(1,39) = 3.82$, $p = .06$, $\eta_p^2 = .09$, revealing a medium effect size. This occurred due to participants in the Training group being more conservative, irrespective of Experience Level, at both Time 1 and Time 2. Therefore, this appears to be a chance group difference, not an effect of the intervention. There was no main effect of Experience Level, and no interaction between Training and Experience Level, both *Fs* < 1.

Mean scores for all groups for *c* were > 0, indicating an overall conservative response bias. This means that participants had a slight bias in favour of responding *genuine*. The most conservative group was the Low Experience group who received Training. The Low Experience participants assigned to the No Training group were less conservative than their trained counterparts at both Time 1 ($M$ = .22) and Time 2 ($M$ = .29). Given there was little change in bias for all groups between Time 1 and Time 2, it would appear this difference between groups is an artifact of random assignment and is unrelated to the intervention. Besides, all the changes in the criterion across conditions and groups are very small and relatively unsystematic.

## SUMMARY OF RESULTS FOR DECEPTION DETECTION

As predicted, the psychologists in the current study were able to discriminate between falsified and genuine expressions at levels above chance, supporting Hypothesis 3. Although there was no significant effect of training on deception detection ability overall, it appears that training interacts with experience level and increases the accuracy of a greater proportion of participants to a level above 70%, for those with high experience.

# CHAPTER SEVEN: DISCUSSION - STUDY ONE

## SYNOPSIS OF STUDY ONE'S AIMS

The goal of this research was to study therapists' abilities in recognising emotional facial expressions, and whether training enhances this ability. It was hypothesized that training would increase emotion recognition skill. Also of interest was the impact of therapist level of experience on emotion recognition abilities. Research has established that recognition of most facial emotion expressions declines with increasing age, and in the current study, age was highly correlated with experience. It was therefore hypothesised that younger, albeit less experienced, clinicians would perform better than older, more experienced clinicians at recognising facial emotion expressions.

A further aim of the study was to explore whether therapists were able to detect deceptive expressions at levels above chance, and whether training in recognising subtle expressions of emotion improved deception detection. It was hypothesised that therapists would be able to detect deception at above chance levels. It was also of interest to find out if training in key features of subtle emotion expressions would improve deception detection.

In the current study, the broad focus is on therapists, incorporating clinicians from various professions such as counsellors, psychologists, social workers, and occupational therapists. However, only clinical psychologists were used as participants in the study as they have similar academic backgrounds and level of training compared to the wider population of therapists whose training incorporates various different pathways. This enables conclusions to be drawn based on age and experience level, without introducing additional confounds that would occur if a more eclectic group of therapists had been selected.

As experience level correlated highly with age (*r*=.85), and previous research has investigated the impact of age, not experience (with the exception of Machado et al.'s, 1999 study), age is used as a proxy for experience level when discussing the current study's results.

Previous research on age-related differences in emotion recognition has used samples of older adults with a mean age greater than the high experience (older) group used in the current study. For example, in Suzuki et al. (2007) the mean age of older adults was 69.7 years versus younger adults with a mean age of 20.6 years. The meta-analysis conducted by Ruffman et al. (2008) surveyed studies where the mean age of the younger groups was less than 45, and the older adult groups exceeded 55 years. In the current study, although there was a significant difference between the mean age of the highly experienced (*M* = 50.6) and the lesser experienced (*M* = 34.2), this age difference is not as extreme as Suzuki et al.'s and other studies. Therefore, the age-related differences observed in the current study might be expected to be less pronounced compared with previous research.

## FINDINGS FROM THE EMOTION RECOGNITION TASK

### Single Emotion Expressions

Training in recognising single expressions of happiness[4] was effective for the highly experienced group but had no effect on the scores of the group with low experience. Prior to training, more experienced participants scored lower than their less experienced counterparts. It was found that training did not result in the highly experienced individuals achieving superior scores to those with less experience; rather, training increased the scores of the highly experienced participants to a level

[4] Caution is needed when interpreting the happiness effect due to the data violating the normal distribution assumption.

equal to the less experienced participants who did not receive training. Training, therefore, appears to have had the remedial effect of restoring recognition of happiness that may have declined with age. Overall, high accuracy levels for detecting happiness were noted, aligning with previous findings of higher accuracy for recognition of happy faces (Hess et al., 1997; Kirita & Endo, 1995). Ceiling effects for recognition of happiness were found, which is consistent with many other studies (e.g., Isaacowitz et al., 2007). It is possible that the stimuli used were too easy and did not provide a sensitive test of differences between groups, despite subtle expressions being used. However, Hess et al. (1997) also found that recognition of joy expressions was close to 100%, even at very low levels of intensity. The current study adds to previous findings showing that overall accuracy is highest for happy expressions, even for very slight, subtle expressions.

Although differences in experience level influenced the ability to recognise happiness, after training, all groups were fairly equal, supporting Hypothesis 1 that training enhances emotion recognition for single happy expressions. Whilst recognition of happy expressions appears to remain intact over the lifespan in most studies, this is possibly an artefact of ceiling effects. However, despite this possibility, Isaacowitz et al. (2007) found that middle-aged adults scored lower for happiness recognition relative to both younger and older adults. With regard to recognition of happiness in the current study, without training, the group of more experienced clinicians scored lower, consistent with Isaacowitz et al.'s findings and supporting Hypothesis 2, that less experienced therapists would be more accurate.

For recognition of single sad expressions, training was effective for both the highly experienced and less experienced groups with increases in scores from Time 1 to Time 2 of 18 and 17 percentage points, respectively, supporting Hypothesis 1. There was also a small practice effect observed for the two groups which did not receive training. It is possible that the SETT training draws attention to reliable signs of sadness, which are quite distinct and once learned, are more readily recognised increasing the likelihood of discerning sadness from other expressions. For example, the training demonstrates how raising upwards of the inner corners of the eyebrow, even without any other change in expression, is a reliable indicator of sadness.

As discussed earlier, a number of studies (e.g., Calder et al., 2003; Suzuki et al., 2007) have found that recognition of sadness declines with age. However, the results of the current study did not support these earlier studies' finding. For recognition of sadness, there were no significant differences between the high and low experience groups prior to training. Therefore, Hypothesis 2 predicting that younger, less experienced clinicians would be more accurate at emotion recognition is not supported for this particular emotion expression. This finding is consistent with that of Sullivan, Ruffman, and Hutton (2007) who found no age-related decline in sadness recognition, but not with others (e.g., Calder et al., 2003; Malatesta, Izard, Culver, & Nicolich, 1987) who have found lower rates of sadness recognition amongst older adults.

A possible explanation for this finding is the way in which the use of dynamic stimuli influenced how the expressions were processed. Kirita and Endo (1995) suggested the happy face advantage over sadness recognition is eliminated when the stimuli are presented in a way that requires analytic processing (e.g., inverted images). It is possible that presenting the sad stimuli as video clips enhanced the recognisability which resulted in both age groups exhibiting similar recognition rates.

The results of the current study's findings that older, more experienced participants are equally proficient at recognising sadness when compared with their younger counterparts is inconsistent with the literature (with the exception of Sullivan et al., 2007). It may be that experience working as a therapist preserves recognition of sadness due to the frequency with which this emotion is encountered with clients. Lastly, it must be born in mind that this study's 'older' group was younger than the 'older' groups reported in the literature.

Mixed results emerged for the recognition of disgust expressions. Accuracy was improved by training but only for the highly experienced individuals; in fact, training adversely affected the performance of those with low experience, who exhibited a decline in disgust recognition after training. However, those with less experience who did not receive training, improved in recognition of disgust, whilst the untrained highly experienced individuals had the greatest decline in accuracy for disgust. Hypothesis 1 therefore received only partial support, with training improving disgust recognition for the highly experienced therapist group only. It is unclear why this group benefited from training, whereas training appeared to have a detrimental effect on those with less experience. It is possible that, as the literature has demonstrated an improvement of disgust recognition with increasing age, the older therapists were more receptive to training for this particular emotion, compared with the younger, less experienced, therapists.

Unlike other studies reporting that detection of disgust improves with age (e.g., Calder et al., 2003; Isaacowitz et al., 2007; Suzuki et al., 2007), this finding was not replicated in the current study. Both of the highly experienced groups had lower disgust scores at Time 1, compared with the low experience groups. This could be due to the smaller mean age difference between the groups, a 'younger' older group, and an 'older' younger group, compared with other studies. In two experiments conducted by Calder et al. (2003), the worst performing groups in both experiments at disgust recognition were the 41-50 year groups, mean ages ($M$ = 46.8; $M$ = 47.2) relative to the younger ($M$ = 24.3; $M$ = 23.9) and older ($M$ = 65.2; $M$ = 66.5) age groups. The best performing groups at detecting disgust were the two oldest groups in both experiments. The mean age of both of 41-50 year age groups in the Calder et al.'s studies is similar to the mean age of the current study's high experience group. Therefore, although it has been reported that disgust recognition increases with age, it seems from Calder and colleagues' studies that this ability does not follow a linear trajectory, but may decrease during middle-age and improve again with older age. Also, Suzuki et al.'s (2007) 'older' group had a mean age of 69.7 years, which is significantly older than the highly experienced group in the current study. Although both of the highly experienced groups recorded the lowest scores for recognising disgust at Time 1, after training, the highly experienced therapists' scores were equal to the low experience group that did not receive training. Also, it has been reported by Suzuki et al. that younger people have more difficulty discriminating between disgust and anger which may account for their lower scores at detecting disgust. In the current study, no anger expressions were displayed but participants still had the option of selecting anger if they thought that is what they had viewed.

## Masked Emotion Expressions

Similar to single happy expressions, a ceiling effect was found for recognition of masked happy expressions. A practice effect was also noted as both untrained groups demonstrated an increase in accuracy at Time 2. However, a limitation is that there were only three masked happiness expressions, providing few opportunities to demonstrate improvement. This occurred because in the generation of the video clips, happiness is the easiest emotion to fake and therefore there were very few video clips with happiness masking another expression available for selection. For this expression type, Hypothesis 1 was not supported as training did not enhance recognition of masked happiness. With regard to Hypothesis 2, there was no effect of experience level on accurate recognition of masked happiness. As a result, Hypothesis 2 was not supported for recognition of masked happiness.

For recognition of masked sad expressions, training had a greater impact on the accuracy of less experienced therapists compared with the more experienced ones. However, those with high experience who received training achieved significantly higher scores at Time 1 than did their untrained counterparts. This was a spurious difference which may have resulted in a ceiling effect restricting the effect of training for this group. So with respect to these expressions, Hypothesis 1 is only partially supported.

It is possible that the greater impact of training on recognising masked sad expressions for those with less experience occurred because this group, being younger, had less overall life experience with exposure to sadness. And whilst there was no difference between the high and low experience for single sad expressions, the more subtle masked sad expressions may have been detected more easily by more experienced individuals. According to Somerville and Whalen (2006), experience with sad expressions occurs less frequently than for happy or neutral expressions. However, their study only surveyed undergraduate students. Perhaps the experience of encountering emotion expressions is different for older people due to life stage differences and greater potential for losses occurring resulting in emotional trauma. As a result, as the participants in the high experience group were also older, they may have gathered more cumulative exposure to this particular emotion compared with their younger colleagues. Also, sadness is an emotion seen frequently by clinicians providing therapy. Given the superior accuracy of the highly experienced group for recognising masked sad expressions, Hypothesis 2 was not supported in relation to this expression.

As was the case for single disgust expressions, training was effective for masked disgust expressions for the highly experienced group but had no impact on those with lesser experience, lending only partial support to Hypothesis 1. Once again, training served to increase the scores of those with high experience to the same level as those with less experience. Thus, one may speculate that this difference is due to the performance of the high experience group being lower at Time 1, relative to the lesser experienced. The difference in mean scores between the high and low experience groups who did not receive training, were significant at both Time 1 ($M$ = .44; $M$ = .69) and Time 2 ($M$ = .39; $M$ = .73), respectively, supporting Hypothesis 2 that less experienced therapists would be more accurate in emotion recognition, compared with older, more experienced ones.

Once again, this finding is inconsistent with the literature reporting a disgust recognition advantage amongst older adults. As mentioned previously, the difference in mean age for the 'older' group in the current study is significantly younger than the 'older' groups that the literature has focused on, thus demonstrating the importance of studies using similar aged groups for comparison. However, the current findings are similar to Calder et al.'s (2003) 41-50 year group, who performed the worst at disgust recognition compared with younger and older age groups. It is possible that a decline in disgust recognition occurs through middle-age, but then improves again with increasing age. This may occur as with increasing age, people's health can become more fragile. Recognition of disgust therefore becomes more salient in order to protect older, more vulnerable adults from ingesting contaminated foods.

### Leaked Emotion Expressions

Accurate recognition of leaked happiness improved mostly as a function of experience, with both high experience groups improving over time, irrespective of training. A practice effect was also noted with the high experience (no training) group showing the greatest increase in scores at Time 2. For the less experienced, there was very little change in accuracy scores from Time 1 to Time 2, irrespective of training. At Time 1, both low experience groups demonstrated higher rates of accuracy compared with the more experienced groups.

For leaked sad and disgust expressions, there was no effect of training or experience level. So in relation to both of these leaked emotions, Hypotheses 1 and 2 were not supported. It was interesting to note, however, that accuracy for recognising leaked disgust was greater than recognition of other leaked emotions, suggesting that the signs of leaked disgust may be more readily observed compared with the other two emotions. Perhaps increased sensitivity to disgust is to protect against ingesting disease vectors. Haidt, McCauley, and Rozin (1994) concluded that disgust is an oral defense, acting as a guardian of the mouth and protecting against consumption of offensive or contaminated foods.

## SUMMARY OF EMOTION FINDINGS

In the current study, there were differences in ERA found between the low and high experience level groups, with the low experience groups achieving higher scores pre-training, except for sadness. Nevertheless, following training, the highly experienced groups' scores were similar to those with less experience. These findings are not consistent with those reported by Machado et al. (1999) who reported that experience level was positively correlated with superior accuracy in emotion recognition. Machado and colleagues' experienced group was relatively young (mean age = 34.8 years), which was more similar to the current study's low experience group (mean age = 34.2 years) and significantly younger than the mean age of the current study's high experience group (*M* = 50.6). As outlined by Machado et al., their findings could have been due to age or to training, as both groups differed in respect to these variables. Furthermore, their control group consisted of undergraduate students, so there would have been variability amongst the two groups for levels of academic achievement and ability. In the current study, all participants were fully-trained clinical psychologists with a minimum of one-year post-graduate practice. They would have been a more homogenous group of participants than the participants used by Machado et al. in terms of professional training.

Although much of the literature reports age-related declines in ERA, this is largely based on a cross section of community dwelling individuals from a range of occupations. It may be argued that the emotion recognition skill of practicing clinical psychologists will increase over time due to frequent interpersonal encounters, and with individual clients through the course of therapy as they gain familiarity. Although Elfenbein and Ambady (2002) state that acquaintance with the same cultural group can increase facial emotion recognition due to the subtle differences in emotion expressions that can occur, there is less evidence that familiarity with an individual enhances recognition, with people exhibiting similar accuracy for judging expressions of friends, strangers, and even their own expressions (Ansfield, DePaulo, & Bell, 1995). Although Zuckerman et al. (1975) reported that familiarity with another person increased accuracy amongst men but not women, they suggested that this had likely occurred due to a female advantage resulting in a ceiling effect. So in therapy, whilst membership of the same cultural group may enhance recognition of emotion, particularly for fear and disgust as found by Elfenbein and Ambady, increased familiarity with a client does not necessarily translate into improved emotion recognition.

The current study has identified three major obstacles to comparing the ERA results to the literature. Firstly, there is a large variation in age-bands used to define 'young' or 'old' age groups. This creates critical differences making it difficult to compare when age is a key independent variable. Secondly, many studies have used static high-intensity, full face expressions of emotions, many of which have been generated by actors portraying an emotion, rather than an expression resulting from a subjective emotional experience. As reported by Hess et al. (1997), increased intensity was related to increased accuracy in recognition. However, in real-life, interpersonal interactions involve low-level, subtle expressions. Finally, many of the studies reviewed by Ruffman et al. (2008) used stimuli which included only young and middle-aged adults. The use of a wider range of adults in stimuli sets may yield different outcomes.

In summary, these results show that, with the exception of sadness, training is most effective for clinicians with more experience. The reason for the greater effect being dependent on experience appears to be that the lower scores exhibited by the more experienced group prior to training provided greater scope for improvement. For happy expressions, training appears to have enhanced the performance of the highly experienced group, but has effected little change on the scores of those with less experience. Interestingly, both the low experience groups' scores were higher than both the high experience groups' scores at Time 1. At Time 2, the scores of those with greater experience and who received training, were similar to the untrained lesser experienced. It appears, then, that the decreasing recognition that occurs with age appears to be offset by training.

## FINDINGS FROM THE DECEPTION DETECTION TASK

### Sensitivity *P(A)*

Participants' accuracy at correctly discerning deceptive facial expressions from genuine expressions was measured using the non-parametric measure, *P(A).* Mean accuracy scores increased for both high experience groups, irrespective of training and the results of the ANOVA revealed a near significant main effect for experience level with a medium effect size. Further examination of the data also showed that the number of individuals achieving a score of 70% or above increased by 22.8% after training, whereas for those who did not receive training, there was only a modest increase of 4.7% at Time 2, suggesting a slight practice effect. Interestingly, in the more experienced group, half scored 70% or above after training, whereas only one-third of the less experienced therapists performed at this level following training. These findings conflict with those reported by Ruffman et al. (2012) who found that older adults were worse than younger adults at detecting deception. Ruffman et al. also found that emotion recognition significantly predicted lie detection. It was surprising that the current study found a negative association between emotion recognition and deception detection as it would appear these two skills may be related. Ekman and O'Sullivan (1991) reported a positive correlation between detecting micro-expressions and deception detection and it has been proposed that the right hemisphere is predominant for both ERA and deception detection, which would suggest these skills are possibly related. The difference between the current study's findings and those of Ruffman et al.'s study could have arisen due to methodological differences. Firstly, the stimuli used by Ruffman et al. for the emotion recognition task were static images and included six emotions. Secondly, their deception task included videos of people presenting an opinion (either honestly or dishonestly) and included verbal information. Also, the length of the videos ranged from 60 *s* to 113 s, compared with 5 s in the current study. Finally, there was a much larger difference between the mean ages of the groups ($M$ = 21.0 years; $M$ = 71.0 years) compared with the present study, which may also account for these different outcomes.

The results from the current study suggest therapists perform at above chance levels for detecting emotional deception. The mean scores for accurately identifying deception ranged from 58% for the lowest group (low experience, no training) to 69% for the highest group (high experience, trained). There were large variations in individual accuracy, which ranged from 44% to 82%. These results are very similar to Ekman et al.'s (1999), with mean group scores of the psychologists in their study ranging from 57.7% for Academic Psychologists to 67.5% for Clinical Psychologists who had demonstrated an interest in deception detection by registering for a two-day workshop on "Lying, Deception, and Malingering". The current study's results also align with those reported by Wright Whelan et al. (2015) who found accuracy for deception detection ranged from 67.8% for members of the general public, to 72.5% for CID officers. The authors suggested that the findings of above-chance accuracy were due to the methodology of using high-stakes lies, compared with other studies (mainly involving low-stakes lies) which have found that most people are no better than chance at detecting deception (e.g., Bond Jr & DePaulo, 2006). Contrasting with these results was Porter et al.'s (2011) study on emotional deception which used a similar method to the present study whereby deceivers expressed emotions incongruent with emotions evoked by images they were observing. It was found that with the exception of happiness, observers could not distinguish between real and false expressions above the level of chance. However, their observers were all university students with a mean age of 20.3 years, and so would be unlikely to have had professional practice in interpersonal deception. Accordingly, it is not surprising that the present study, using a similar method to create the emotion deception stimuli, found therapists were better at detecting deception, as the role depends on the ability to read the verbal and nonverbal cues of others and involves a large amount of face-to-face contact. The review by Aamondt and Custer (2006) also reported that teachers had a high deception detection rate of 70%, which is possibly due to the high level of face-to-face contact required by this role. The addition of music during the emotion induction procedure may have deepened the emotional experience, leading to more leakage of felt emotions.

In the current study, contrary to Ekman and O'Sullivan's (1991) findings, experience level was positively correlated with deception detection. They found for the Secret Service group, age and experience level were both negatively correlated with deception detection. In the Federal Polygrapher group, age, but not experience level, was also negatively correlated with accuracy in detecting deception. For the other groups examined in the Ekman and O'Sullivan study, there was no correlation between age or experience level and deception detection accuracy.  Others have also reported no correlation between years of experience and deception detection accuracy (e.g., Ekman et al., 1999; Porter et al., 2000; Wright Whelan et al., 2015).

### Bias (*c*)

All of the groups displayed a conservative bias, meaning they were more inclined to respond *genuine* rather than *fake.*  This is unsurprising as a foundation of clinical psychology training is to treat clients with unconditional, positive regard.  This tendency towards responding 'genuine' contrasts with Wright Whelan et al.'s (2015) proposal that  police officers were more accurate at detecting deception due to adopting a liberal bias.  However, their analysis did not utilise SDT; they used hit rate as their dependent measure ignoring false alarms. Thus, a person who consistently responded with one particular choice might score 100% correct.  Detectability and bias in favour of one or the other choices are confounded.

## SUMMARY OF DECEPTION DETECTION FINDINGS

Previous studies have found no positive correlation between either age or experience and deception detection accuracy, with some finding negative correlations between these variables.  However, the present study has found a positive correlation between age/experience level and accuracy in detecting deception.  Furthermore, it was found that for those with greater experience, training was more effective for increasing deception detection accuracy.  This is possibly due to the use of stimuli which included only an emotional expression with no verbal information provided whereas the previous studies mentioned, except for Porter et al. (2011), have used video recordings of people providing honest or dishonest verbal accounts.  Ekman and O'Sullivan (2006) reported that there are a number of differences between genuine and masked emotion expressions, including morphology, timing, and symmetry.  It is possible that by excluding verbal information, participants attended more to these clues to deception, and that prior experience as a therapist increased the ability to discern between genuine and faked emotion expressions.

In summary, the current findings that more experienced therapists detect emotional deception with greater accuracy compared with those with less experience contrasts with other results relating to age and experience, and correlations between emotion recognition and deception detection (e.g., Ekman & O'Sullivan, 1991; Ruffman et al., 2012; Wright Whelan et al., 2015). This may be because these studies, unlike the present one, used videotapes of people lying and truth telling which included verbal information. Moreover, their emotion recognition tasks included static images of six emotions. In Ruffman et al.'s study, the older participants were community-dwelling people who had not studied psychology and had a mean age significantly greater than those in the current study. It is possible that the current study's finding that older, more experienced therapists are more accurate at detecting emotional deception may be due to an accumulation of experience in working with people in an emotionally expressive dyad, despite their lower accuracy rates for ERA. The lack of association between ERA and deception detection found in the current study may be at least partially due to the use of only three emotions compared with six as in the aforementioned studies. Although the more experienced therapists were less accurate at identifying discrete emotion expressions, this did not appear to influence their ability to discern authentic from faked expressions. Older, more experienced therapists may also be higher in self-monitoring, a trait that Aamond and Custer (2006) reported is related to accuracy in deception detection.

With regard to studies examining deception detection, in addition to differences in stimuli used to test detection accuracy, a major difficulty in making comparisons is the use by some of percentage correct scores, whilst others have used SDT, which provides independent assessments of accuracy and response bias.

# CHAPTER EIGHT: STUDY TWO

*A second study was conducted to evaluate the effectiveness of a new version of the Subtle Emotion Training Tool used in Stage 2 of Study 1. The method, details of the participant sample, and procedure used to collect the data are described. This is followed by an analysis of the data and presentation of the results, following the same procedure used in Stage 2 of Study 1.*

## DECEPTION DETECTION AND EMOTION RECOGNITION TASK AND EMOTION RECOGNITION TRAINING – STUDY 2

Whilst the first study was being conducted, a new version of the Subtle Emotion Training Tool was released. Study 1 used the original version of the Subtle Emotion Training Tool (SETT). This presented information about the seven universal expressions: Happy, Sad, Anger, Fear, Surprise, Contempt, and Disgust, all displayed by the same young Caucasian woman. The rationale presented by the training programme was that using one face would eliminate the distractions that may occur when looking at different facial features of a number of faces.

More recently, an updated version of SETT has been released which has features superior to the original SETT such as a wider range of ethnicity of faces from both sexes, and an increased number of expressions (42 expressions compared with 37). It also provides more comprehensive commentary on each of the expressions in the *learn* phase, providing verbal descriptions of the reliable signals of each emotion expression. The commentary also describes the different ways that emotions can be signaled, including how to distinguish between Duchenne (genuine) and non-Duchenne (simulated) smiles by observing the activation of the muscles surrounding the eyes. It also describes how asymmetry, such as a smile being more intense on one side of the face or a tilting of the head, can provide clues that an expression may not be a genuine, spontaneous expression.

For this reason, a second study was conducted to provide a preliminary evaluation of the effectiveness of the new SETT training tool.

## Participants

Psychologists who had not previously completed any training in facial emotion recognition were invited to participate in Study 2 via email networks of members of the New Zealand Psychological Society and the New Zealand College of Clinical Psychologists, and through a presentation made to a group of Corrections Psychologists participating in a regional training day. All participants were currently practicing psychologists. Thirty people responded and were emailed an *Information Sheet* (refer Appendix I). Of these, five withdrew due to work or other commitments. Of the remaining 25 who participated, there were 17 females and 8 males aged between 24 and 64 years ($M$ = 38.6 years; $SD$ =9.73). Self-reported nationality of participants was 18 New Zealand, two South African, one NZ Maori, one Maldives, one British, one Irish, and one European. Twenty-one were practicing clinical psychologists, and four were practicing registered psychologists, with experience ranging between 1 and 39 years ($M$ = 6.90 years; $SD$ = 8.64). Fourteen participants worked in mental health settings, and eleven participants were employed by the Department of Corrections.

## Stimuli

### *Deception Detection and Emotion Recognition Task (DDERT)*

The same DDERT programme developed from the emotion expression stimuli created in Stage 1 and used in Stage 2 of Study 1 was utilised. This was copied to individual CD-ROMS, or USB drives, depending on participant preference, one for each participant.

***The Subtle Expression Training Tool (SETT)***

Ekman's SETT, released in April 2014, was used to train participants in the training group in facial emotion recognition (see www.paulekman.com). Similar to the original SETT used in Stage 2 of Study 1, the current version teaches recognition of very subtle signs of an emotional expression. The course description appearing on the Paul Ekman Group website states: "Subtle Expressions Training Tool is the premier training program for learning to recognize subtle expressions. Subtle expressions are the very "mini" expressions that often appear in just one region of your face, such as the brows, eyelids, cheeks, nose, or lips. They can occur when a person is trying to conceal a strong emotion — subtle expression can leak, exposing the true emotion in a tiny change in expression. These small movements may also occur when an emotion is just beginning, often before the person is aware of their emotional state." This version of the SETT has three sections: *Intro, Learn* and *Practice*. The *Intro* section provides an overview of the course and information regarding subtle emotion expressions. The *Learn* section provides specific information about the subtle signs of each of the seven universal emotions: anger, fear, sadness, disgust, happiness, surprise, and contempt. The expressions can be viewed at three speeds (slower; normal; faster); however, in the present study, participants were instructed to use the "normal" setting. Information is presented visually as well as via an audio commentary of the key features provided by Paul Ekman. After viewing each emotion in the *Learn* section, participants can test their skill in the *Practice* section. In this section, an expression is flashed on screen and users select the corresponding emotion button on the right of the screen. If users give an incorrect response, they can replay the emotion once more before making another selection. If an incorrect selection is made the second time, the correct answer is given. At the end of the practice section, a percentage correct score is displayed. The training in the *Learn* section may be repeated as often as required although it is recommended that the practice test is not taken more frequently than three-monthly to avoid increasing accuracy through memorising faces and expressions. Once a score of 80% or more is achieved, the trainee receives a certificate of course completion.

## Procedure

Participants were assigned a number in order of receipt of expression of interest in participating in the study. Thirty numbers were randomly assigned to two groups using an on-line random sequence generator (www.random.org). The procedure followed for testing emotion expression recognition and providing training to participants was identical to the procedure used in Stage 2 of Study 1 and previously described in Chapter Four.

## DATA ANALYSIS

The same procedure for data analysis as described in Chapter Five was undertaken for Study 2. ANOVA tables can be found in Appendix N.

In the results section, the results are presented and described for the analysis of emotion recognition and deception detection. However, the results from Studies 1 and 2 cannot be meaningfully compared for a number of reasons, such as the use of two different training programmes. The problems of comparison are covered in detail in the discussion chapter.

## AGE AND EXPERIENCE LEVEL

An independent samples *t*-test was conducted to compare years of experience and age for the two experimental groups. There was no significant difference ($t(23) = 1.24$, $p = .255$) for years of experience between the Training group ($M = 5.00$, $SD = 5.87$) and the No-Training group ($M = 9.21$, $SD = 10.65$). There was also no significant difference between the groups for Age ($t(23) = 2.40$, $p = .815$); Training group ($M = 34.46$, $SD = 8.20$), No Training group $M = 43.00$, $SD = 9.57$). The descriptive statistics for the two groups are displayed in Table 9.

Table 9

*Descriptive Statistics Across Experimental Groups*

| | Female | Male | Age range | Mean Age *M* (*SD*) | Years of Experience *M* (*SD*) |
|---|---|---|---|---|---|
| Training (*n*=13) | 9 | 4 | 24-50 | 34.46 (8.20) | 5.00 (5.87) |
| No Training (*n*=12) | 8 | 4 | 31-64 | 43.00 (9.57) | 9.21 (10.65) |

***Experience Level***

A median split was performed using SPSS to divide the sample into two groups creating a categorical variable - Experience Level, according to years of experience: Low Experience (*M* = 1.82 years, *SD* = 0.87); High Experience (*M* = 13.64, *SD* = 9.50). An independent samples *t*-test showed that the difference between these two means was significant, $t(23) = 4.66$, $p < .007$. Unlike in the first study, there was no significant difference between the mean age of the two groups, Low Experience *M* = 35.21 years, *SD* = 8.33, High Experience *M* = 42.82 years, *SD* = 10.06, $t(23) = 2.07$, $p = .75$. The descriptive statistics are displayed in Table 10.

Table 10

*Descriptive Statistics for Split by Level of Experience*

| | Female | Male | Age range | Mean Age *M* (*SD*) | Years of Experience *M* (*SD*) |
|---|---|---|---|---|---|
| Low Experience (*n*=14) | 12 | 2 | 24-53 | 35.21 ( 8.33) | 1.82 (0.87) |
| High Experience (*n*=11) | 5 | 6 | 29-64 | 42.82 (10.06) | 13.64 (9.50) |

The following analyses included ANOVAs investigating the effects of both Training and Experience Level on the dependent variables of interest, effectively creating four groups (Low Experience – Training; Low Experience – No Training; High Experience – Training; High Experience – No Training).

The descriptive statistics for the groups split by experience level across experimental conditions are shown in Table 11.

Table 11

*Age and Years of Experience for Low and High Experience Groups Across Experimental Conditions*

| | Age range | Age (years) *M* (*SD*) | Years of Experience *M* (*SD*) |
|---|---|---|---|
| Low Experience | | | |
| Training (*n*=9) | 24-48 | 32.78 ( 7.74) | 2.00 ( 1.00) |
| No Training (*n*=5) | 32-53 | 39.60 ( 8.26) | 1.50 ( 0.50) |
| High Experience | | | |
| Training (*n*=4) | 29-50 | 38.25 ( 9.03) | 11.75 ( 6.90) |
| No Training (*n*=7) | 31-64 | 45.43 (10.29) | 14.71 (11.09) |

To check whether there were any significant differences between the years of experience split between the experimental groups, independent samples *t*-tests were conducted. There was no significant difference ($t(12) = 1.25$, $p > .05$) between years of experience for the Low Experience participants assigned to the Training group, $M = 2.00$, $SD = 1.00$, and the Low Experience participants assigned to the No-Training group, $M = 1.50$, $SD = 0.50$. There was no significant difference ($t(9) = .48$, $p > .05$) for years of experience between the High Experience participants assigned to the Training group, $M = 11.75$, $SD = 6.90$, and the High Experience participants assigned to the No-Training group, $M = 14.71$, $SD = 11.09$.

# CHAPTER NINE: RESULTS FOR STUDY 2

## EMOTION RECOGNITION

*The first section of this chapter presents the statistical results for the emotion recognition task.*

As in Study 1, Stage 2, the design employed was a 2 x 2 x 2 mixed ANOVA. The first two factors were the between-subjects factors of Training (Training; No Training) and Experience Level (High; Low). The third factor was the within-subjects factor of Time (Time1; Time2). The dependent variable was the percentage of correct judgements made by the participants. The hypothesis tested was that the effects of training would increase the percentage of correct judgements of emotion type made by the participants who received the SETT training compared to those who did not.

Also of interest was whether level of experience of participants would influence judgement accuracy.

Mean percentage scores for both groups were calculated for Time 1 and Time 2 for each of the expression types, and for each emotion category. The descriptive results are presented in Table 12 showing the means and standard deviations for each of the three categories of emotion expression.

Table 12

*Mean Percentage Scores and Standard Deviations (in Parentheses) for All Emotion Expressions for Time 1 and Time 2 for Each Group (Training vs No-Training) and Experience Level (High vs Low)*

| | Training | | | | No Training | | | |
|---|---|---|---|---|---|---|---|---|
| Category | Low experience *M* (*SD*) | | High experience *M* (*SD*) | | Low experience *M* (*SD*) | | High experience *M* (*SD*) | |
| | T1 | T2 | T1 | T2 | T1 | T2 | T1 | T2 |
| Single Expression | | | | | | | | |
| Happy | .95 (.08) | .90 (.08) | .91 (.67) | .95 (.07) | .96 (.04) | .92 (.08) | .92 (.10) | .92 (.10) |
| Sad | .54 (.31) | .69 (.24) | .42 (.23) | .70 (.17) | .63 (.06) | .56 (.16) | .57 (.21) | .57 (.20) |
| Disgust | .70 (.13) | .81 (.08) | .55 (.17) | .70 (.16) | .64 (.15) | .72 (.08) | .60 (.19) | .81 (.17) |
| Masked Expression | | | | | | | | |
| Happy | .94 (.17) | .89 (.22) | .88 (.25) | .75 (.29) | 1.00 (.00) | 1.00 (.00) | .93 (.19) | .86 (.25) |
| Sad | .48 (.21) | .45 (.22) | .29 (.08) | .71 (.29) | .33 .20) | .43 (.34) | .29 (.16) | .45 (.23) |
| Disgust | .64 (.12) | .63 (.22) | .57 (.30) | .57 (.13) | .53 (.21) | .55 (.29) | .65 (.24) | .66 (.19) |
| Leaked Expression | | | | | | | | |
| Happy | .16 (.17) | .10 (.14) | .07 (.08) | .22 (.19) | .18 (.11) | .08 (.11) | .06 (.07) | .13 (.15) |
| Sad | .19 (.17) | .19 (.15) | .08 (.17) | .41 (.29) | .20 (.14) | .10 (.09) | .21 (.12) | .22 (.16) |
| Disgust | .78 (.26) | .67 (.25) | .50 (.00) | .75 (.50) | .40 (.42) | .60 (.22) | .43 (.34) | .71 (.39) |

## Single Emotion Expressions

### *Happy*

As in Study 1, the scores for correctly identifying a Happy expression were not normally distributed as can be seen in the histograms in Figure 20. Once again, this is likely a result of the ease of the task.

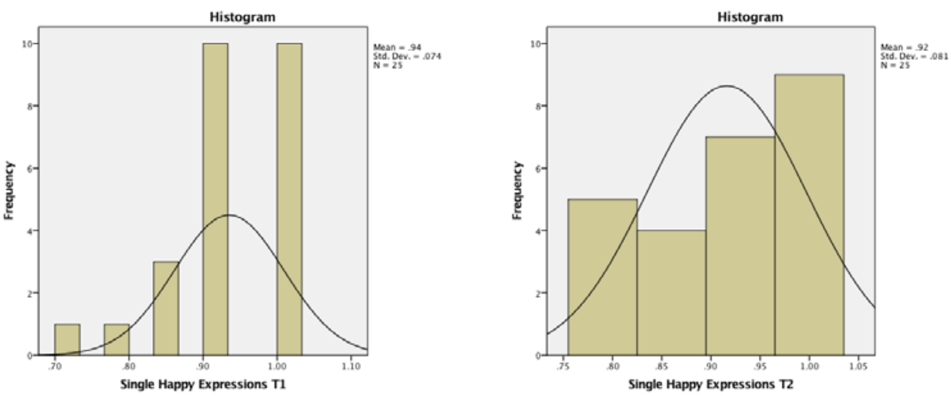



*Figure 20.* Distribution of scores for correct identification of a single Happy expression at Time 1 and Time 2 for all participants.

The mean scores for each group are shown in Table 12 and graphically presented in Figure 21, showing that for the participants with Low Experience, scores declined over time irrespective of whether training occurred or not. The Low Experience (Training) group's scores decreased from Time 1 (*M* = .95) to Time 2 (*M* = .90), and for the Low Experience (No Training) group, scores decreased from Time 1 (*M* = .96) to Time 2 (*M* = .92). However, for participants with High Experience, Training increased scores from Time 1 (*M* = .91) to Time 2 (*M* = .95). The scores for the High Experience (No Training) group remained the same from Time 1 (*M* = .92) to Time 2 (*M* = .92).

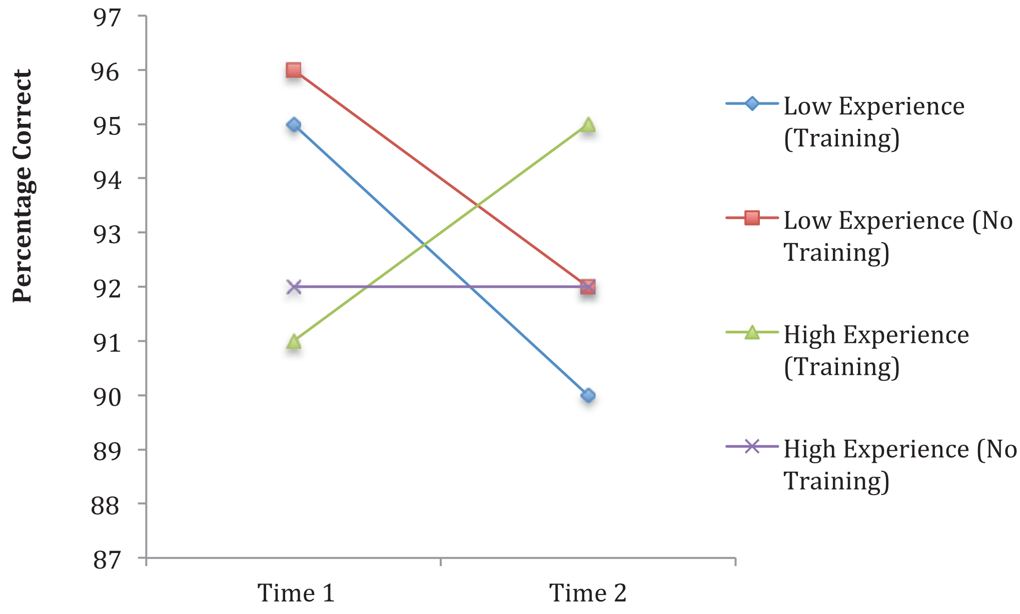



*Figure 21.* Change in scores for identifying a single Happy expression from Time 1 to Time 2 for all groups.

There was no main effect on accurate recognition of a single Happy expression of Time, and no interaction between Time and Training, or Time, Training, and Experience Level, all *F*s < 1.

There were no main effects of Experience Level or Training, and no interaction between Training and Experience Level, all *F*s < 1.

***Sad***

Figure 22 reveals that Training increased emotion recognition for Sadness, for both Experience Level groups. The Low Experience (Training) group's mean score increased by 15% from Time 1 (*M* = .54) to Time 2 (*M* = .69). The High Experience (Training) group's mean score increased by 28% from Time 1 (*M* = .42) to Time 2 (*M* = .70). The Low Experience (No Training) group's mean score decreased by 7%, and the High Experience (No Training) group's mean score remained the same from Time 1 to Time 2.

The results for both training groups are similar to those found in Study 1, with training improving scores. However, in the previous study, a small practice effect was also found, which did not occur in the present study.

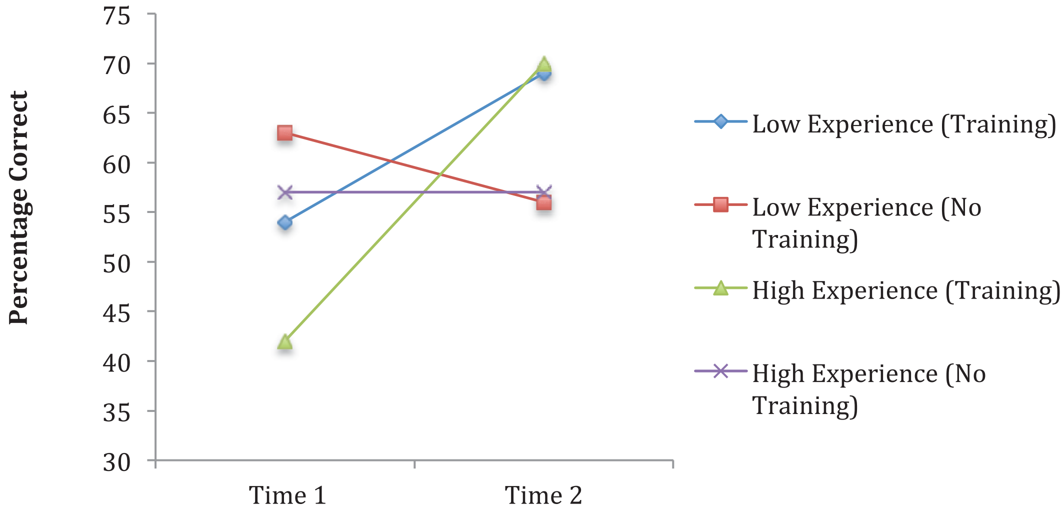



*Figure 22.* Change in scores for identifying a single Sad expression from Time 1 to Time 2 for all groups.

There was a significant interaction between Time and Training, $F(1,21) = 6.87$, $p = .02$, $\eta_p^2 = .25$, showing a large effect of training. These results suggest that training is somewhat effective, but is most effective for those with a High level of Experience.

There was no main effect of Experience Level, Training, or interaction between Experience Level and Training, all *Fs* < 1.

***Disgust***

The results for recognising single Disgust expressions are displayed graphically in Figure 23 showing that Training increased the scores for both the High Experience, Time 1 (*M* = .55) to Time 2 (*M* = .70) and Low Experience groups, Time 1 (*M* = .70) to Time 2 (*M* = .81). For the groups who did not receive training, scores also increased over time; High Experience, Time 1 (*M* = .60) to Time 2 (*M* = .81) and Low Experience groups, Time 1 (*M* = .64) to Time 2 (*M* = .72). In the previous study, Training appeared to benefit only the High Experience group but in the present study both groups derived benefit from training. Also, previously those in the High Experience group who did not receive training showed a decline in scores over time, whereas in the current study, the scores of both the Low Experience and High Experience (untrained) groups increased, suggesting a practice effect.

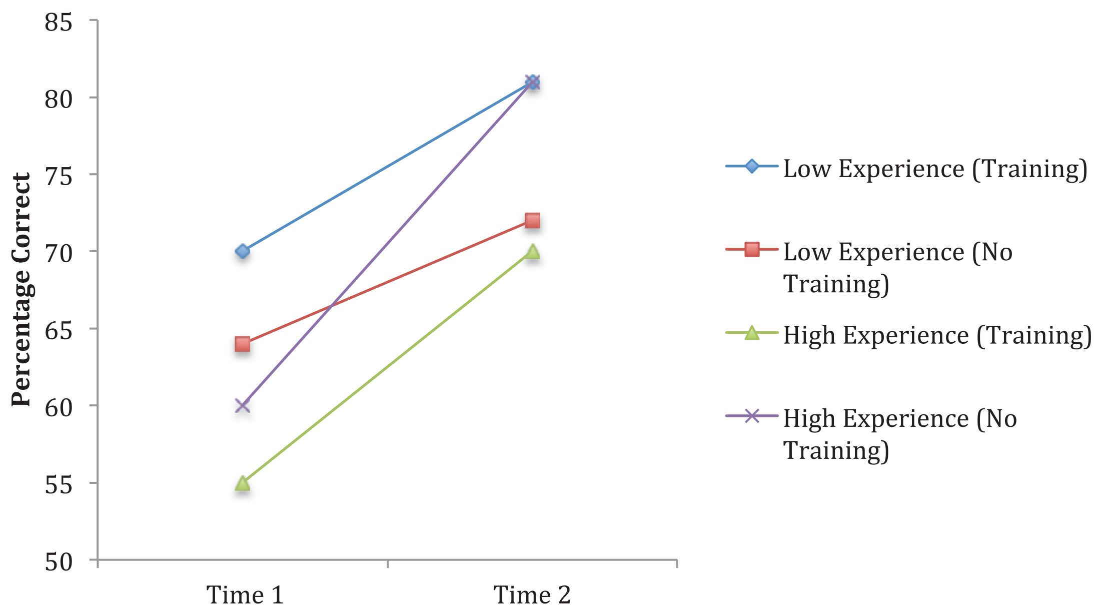



*Figure 23.* Change in scores for identifying a single Disgust expression from Time 1 to Time 2 for all groups.

The main effect of Time on accurate recognition of a Disgust emotion produced a large effect and was significant, $F(1,21) = 16.00$, $p = .001$, $\eta_p^2 = .43$, supporting the suggestion a practice effect occurred.

There was no interaction between Time and Training, or Time, Experience Level, and Training, both *Fs* < 1. There was no main effect of Training, *F* < 1.

## Masked Emotion Expressions

### *Happy*

Once again, as with the single Happy expressions, the scores for correctly identifying a masked Happy expression were not normally distributed as can be seen in the histograms in Figure 24. Despite this, ANOVAs for masked Happy expressions were still conducted, but caution is required when interpreting results.

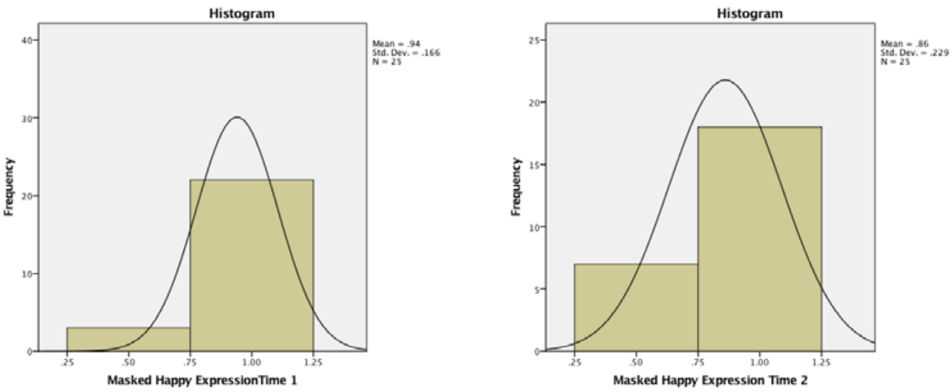



*Figure 24.* Distribution of scores for correct identification of a Masked Happy expression at Time 1 and Time 2 for all participants. Note that the frequency ranges on the ordinates of the two plots are different.

The mean scores for each group are shown in Table 12 and graphically presented in Figure 25, showing that except for the Low Experience (no training) participants, scores declined over time irrespective of whether training occurred or not.

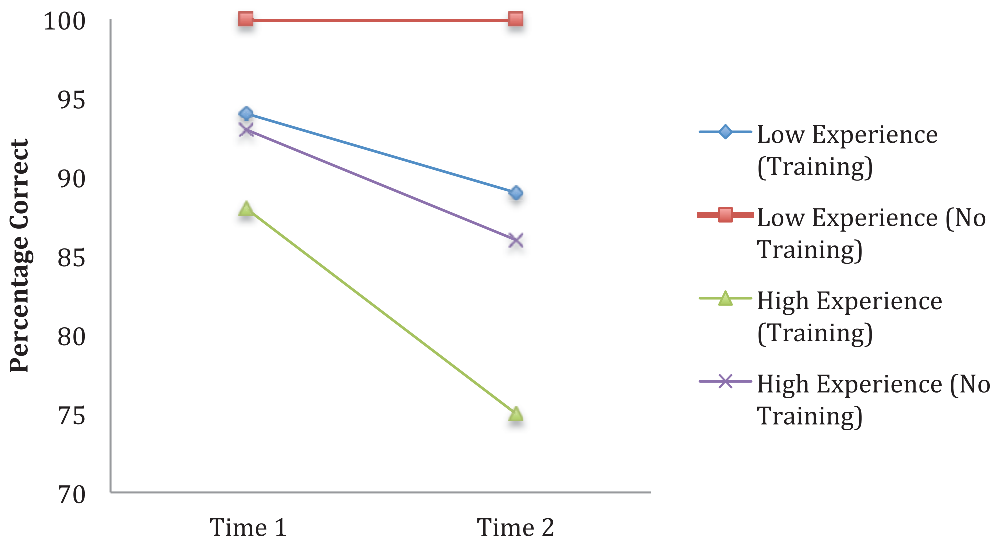



*Figure 25.* Change in scores for identifying a Masked Happy expression from Time 1 to Time 2 for all groups.

There was no significant main effect of Time on accurate recognition of a masked Happy expression. However, Figure 25 shows a definite trend with scores for three of the groups decreasing at Time 2. There was no interaction between Time and Experience Level, or between Time and Training, or Time, Training, and Experience Level, all *F*s < 1.

Overall, for this particular emotion, it would seem that training and practice are both ineffective for improving recognition of masked happy expressions, particularly for those more experienced participants. However, there are reasons for treating these results with suspicion, a matter taken up in the Discussion section.

***Sad***

Figure 26 reveals that Training increased emotion recognition for masked Sadness, for the High, but not the Low Experience group. This is unlike the effect of Training on recognition of a single Sad expression, which was effective for both low and high experienced participants.

The Low Experience (Training) group's mean score decreased from Time 1 (*M* = .48) to Time 2 (*M* = .45). The High Experience (Training) group's mean score increased by

42 percentage points from Time 1 ($M$ = .29) to Time 2 ($M$ = .71). The Low Experience (No Training) group's mean score increased by 10 percentage points, and the High Experience (No Training) group's mean score also increased by 16 percentage points from Time 1 to Time 2, suggesting a practice effect. This is also dissimilar to the results for single sad expressions where no practice effect was observed.

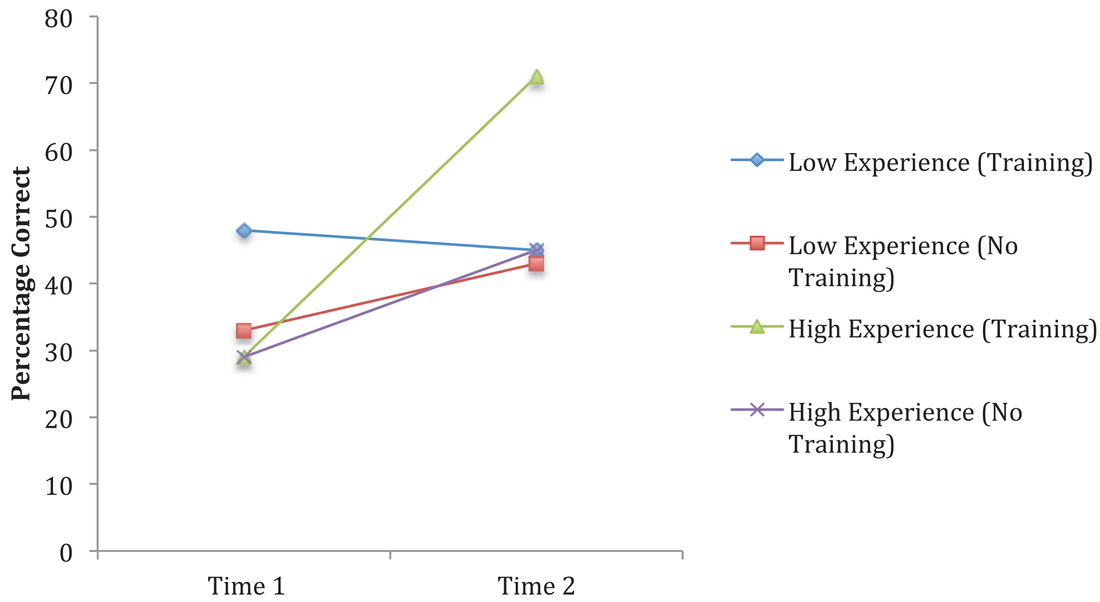



*Figure 26.* Change in scores for identifying a masked Sad expression from Time 1 to Time 2 for all groups.

There was a main effect of Time on accurate recognition of a masked Sad expression, $F(1,21) = 6.97$, $p = .02$, $\eta_p^2 = .25$, with a large effect. However, this main effect of Time was qualified by a significant interaction between Time and Experience Level, $F(1,21) = 4.54$, $p = .05$, $\eta_p^2 = .18$, showing a medium effect. This arose from the increase in scores over time for the High Experience groups, particularly for those who received training, in contrast to the Low Experience groups, where only the untrained group's performance improved over time. This would suggest that those with more experience benefit from both training and practice. There was no interaction between Time and Training.

There was no main effect of Experience Level, and no interaction between Experience Level and Training, both $Fs < 1$.

### *Disgust*

The results for masked Disgust are displayed graphically in Figure 27 showing that for all four groups, neither Time nor Training had any great impact on scores, with these remaining reasonably static from Time 1 to Time 2.

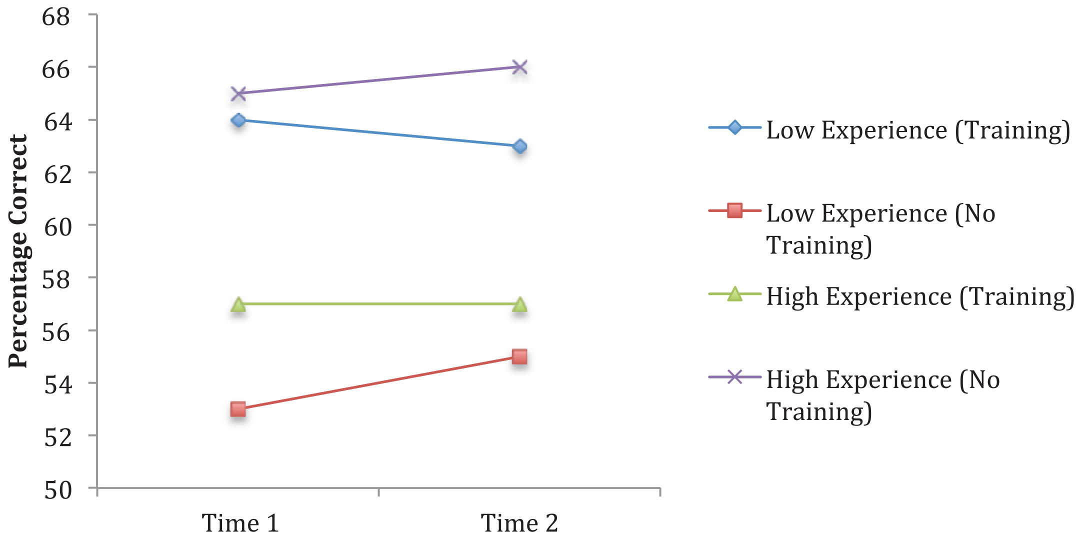



*Figure 27.* Change in scores for identifying a masked Disgust expression from Time 1 to Time 2 for all groups.

As would be expected from observing Figure 27, there was no main effect of Time, no interaction between Time and Experience Level, Time and Training, Experience Level and Training, and no three-way interaction between Time, Experience Level, and Training, all *Fs* < 1.

There was also no main effect of Experience Level, or of Training, both *Fs* < 1. In summary, there was very little effect of Time, Training, or Experience Level, on scores for detecting masked Disgust expressions.

## Leaked Emotion Expressions

### *Happy*

The mean scores for accurately detecting a leaked Happy expression for each group are shown in Table 12 and visually displayed in Figure 28. Both High Experience groups' scores increased from Time 1 (Training, *M* = .07; No-Training *M* = .06) to

Time 2 (Training, *M* = .22; No-Training *M* = .13), whereas both Low Experience groups' scores decreased over time, Time 1 (Training, *M* = .16; No-Training *M* = .18) to Time 2 (Training, *M* = .10; No-Training *M* = .08).

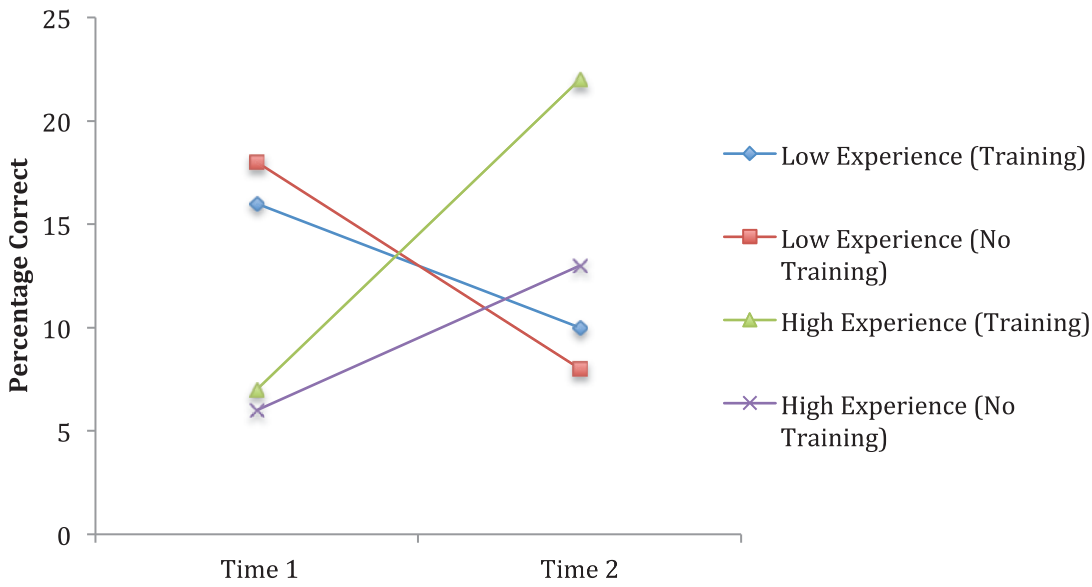



Figure 28. Change in scores for identifying a leaked Happy expression from Time 1 to Time 2 for all groups.

There was no main effect of Time on accurate recognition of a leaked Happy expression, $F < 1$, but there was a significant interaction between Time and Experience Level, $F(1,21) = 18.42$, $p < .001$, $\eta_p^2 = .47$, producing a large effect. Both High Experience groups improved over time but as Figure 28 indicates the trained High Experience group improved more than the untrained group. There was no interaction between Time, Training, and Experience Level, $F < 1$.

There were no main effects of Training or Experience Level, and no interaction between Training and Experience Level, all *F*s < 1.

***Sad***

Figure 29 shows that training increased emotion recognition ability for leaked Sadness, but only for those with high experience, from Time 1 (*M* = .08) to Time 2 (*M* = .42). This is similar to the results for masked sadness. For the Low Experience group, Training did not affect their scores from Time 1 (*M* = .19) to Time 2 (*M* = .19). The Low Experience Level group that did not receive Training decreased in scores from Time 1 (*M* = .20) to Time 2 (*M* = .10), whereas the High Experience (no-training) group's scores changed by only one percentage point from Time 1 to Time 2.

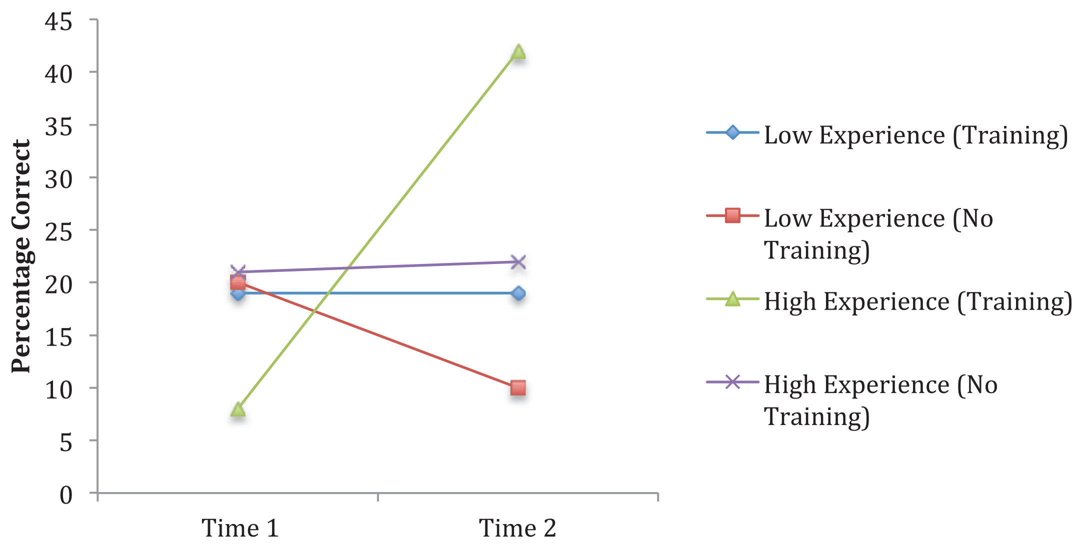



*Figure 29.* Change in scores for identifying a leaked Sad expression from Time 1 to Time 2 for all groups.

There was a significant medium-large interaction between Time and Experience Level, $F(1,21) = 4.67$, $p = .04$, $\eta_p^2 = .18$, and Time and Training $F(1,21) = 4.66$, $p = .04$, $\eta_p^2 = .18$. As can be seen in Figure 29, these results are largely due to the High Experience group's scores increasing by 34 percentage points.

There was no main effect of Training, and no interaction between Training and Experience Level, both *F*s < 1.

***Disgust***

Figure 30 displays the change in mean scores for detecting leaked Disgust for all groups from Time 1 to Time 2. All of the scores increased except for the Low Experience Training group.

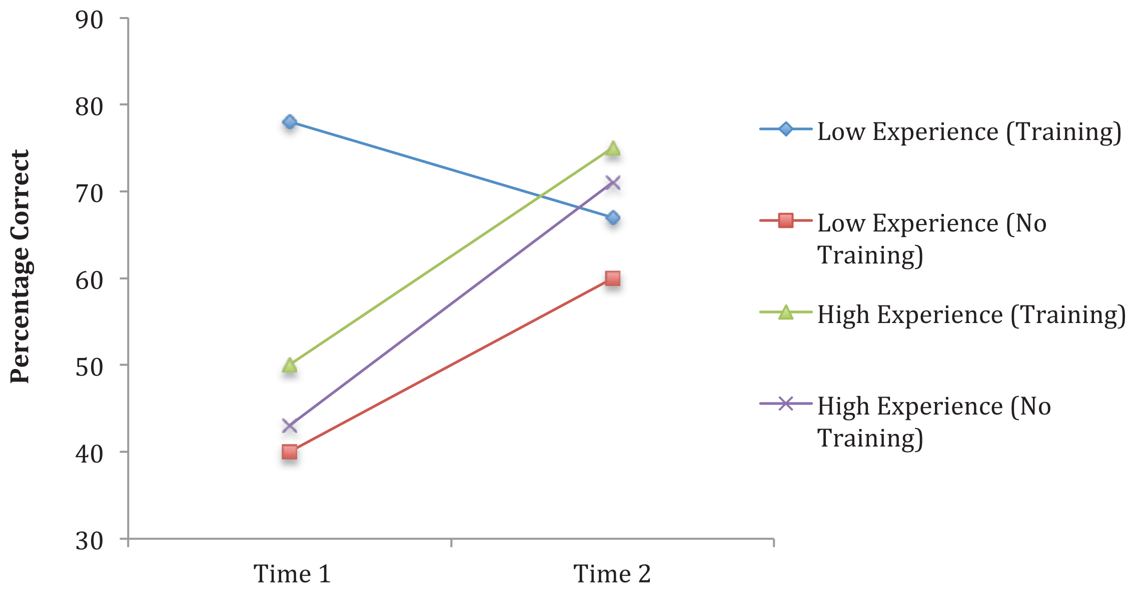



*Figure 30.* Change in scores for identifying a leaked Disgust expression from Time 1 to Time 2 for all groups. Note, the scores for the untrained High and Low Experience groups are equal.

The ANOVA analysis revealed a marginal main effect of Time, $F(1,21) = 4.06$, $p = .06$, $\eta_p^2 = .16$, with a medium effect size. There were no other main effects or interactions, all $Fs < 1$.

These results suggest that a practice effect occurred, and this was more notable for the more experienced participants. It is also noted that the scores for recognising leaked Disgust were considerably higher than those for recognising leaked Happiness or Sadness, a finding which was also observed in Study 1. Once again this would suggest that even without training, leaked Disgust is easier to detect than leakage of the other two emotions.

## ANALYSIS OF DECEPTION DETECTION

The analysis was performed as described for Study 1, Stage 2. A Group (Training vs No-Training) x Experience Level (Low Experience vs High Experience) x Time (Time 1, Time 2) ANOVA was carried out for each variable obtained from the Signal Detection analysis. As the primary interest was deception detection ability and bias, only these two variables were analysed for this study. The descriptive results are presented in Table 13.

## RESULTS FOR DECEPTION DETECTION AND BIAS

Table 13

*Mean Group Values for Time 1 and Time 2 for P(A), and c, for Low Experience and High Experience across Experimental Groups*

| | Training | | | | No Training | | | |
|---|---|---|---|---|---|---|---|---|
| | Low experience *M* (*SD*) | | High experience *M* (*SD*) | | Low experience *M* (*SD*) | | High experience *M* (*SD*) | |
| | T1 | T2 | T1 | T2 | T1 | T2 | T1 | T2 |
| *P(A)* | .59 (.11) | .56 (.09) | .54 (.05) | .63 (.11) | .68 (.08) | .65 (.07) | .58 (.11) | .60 (.08) |
| *C* | .18 (.23) | .36 (.16) | .11 (.57) | .24 (.44) | .12 (.38) | .05 (.40) | .38 (.31) | .20 (.33) |

*Note.* Low experience (training) *n*=9; High experience (training) *n*=4; Low experience (no-training) *n*=5; High experience (no-training) *n*=7.

### Sensitivity - *P(A)*

As shown in Figure 31, the scores for both High Experience groups have increased from Time 1 to Time 2, with the greatest increase occurring for the group that received training, Time 1 (*M* = .54,) to Time 2 (*M* = .63), reflecting an increase of nine percentage points, compared with a two percentage point increase for the High Experience (untrained) group. However, for both the Low Experience groups, scores decreased by three percentage points from Time 1 (*M* = .68*, M* = .59*)* to Time 2 (*M* = .65, *M* = .56*)*, irrespective of whether Training occurred or not. What appears to have an impact on scores at Time 2 is Experience Level rather than Training, albeit a smallish impact. That is, the High Experience participants improved in sensitivity from Time 1 to Time 2, with those receiving training showing the greatest increase, and Low Experience participants' sensitivity declined from Time 1 to Time 2, irrespective of Training.

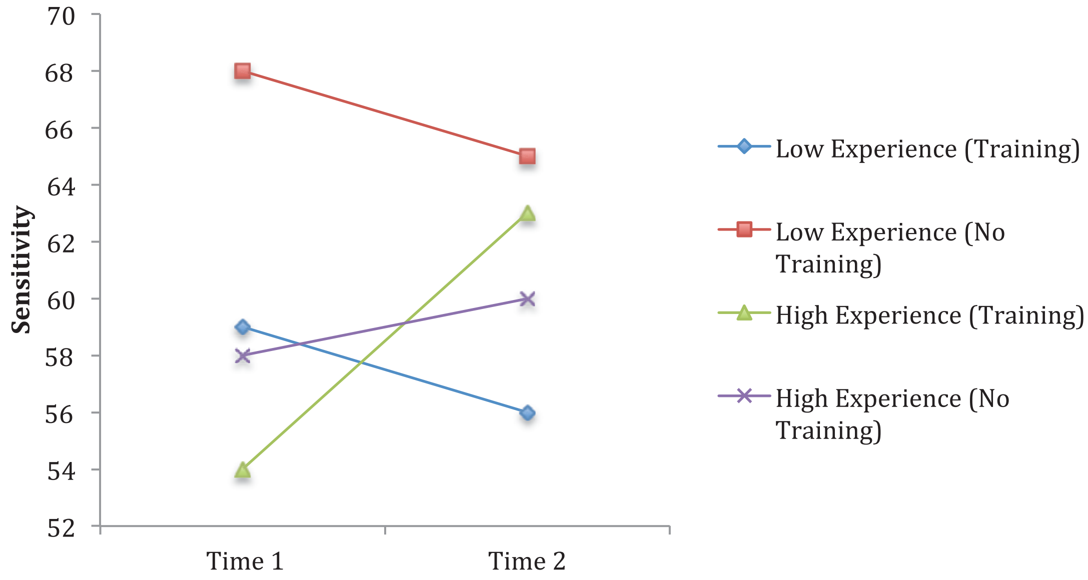



*Figure 31.* Indices of sensitivity as measured by *P(A)* for all groups for Time 1 and Time 2.

There was no main effect of Time, and no interaction between Time and Training, both *F*s < 1. There was a significant interaction between Time and Experience Level, $F$ (1,21) = 6.07, $p < .05$, $\eta_p^2 = .22$, a large effect. This occurred as a result of both high experience groups increasing in sensitivity over time, whilst both low experience groups decreased in their ability to detect deception.

Training had no impact on the less experienced participants but for those with more experience, their scores improved more over time with training than the High Experienced untrained group, who only demonstrated a modest increase with practise. There was no main effect of Experience Level, $F < 1$. As in the first study, the High Experience group appeared to benefit the most from training.

## Response Bias (*c*)

Figure 32 displays the differences in response bias between the groups. Both low and high experience groups that received training demonstrated a tendency towards being more conservative following training, that is, they were more likely to say genuine at Time 2 than they were at Time 1. Whereas the two groups that did not receive training became more liberal at Time 2, that is, they were more likely to say fake at Time 2 than they were at Time 1.

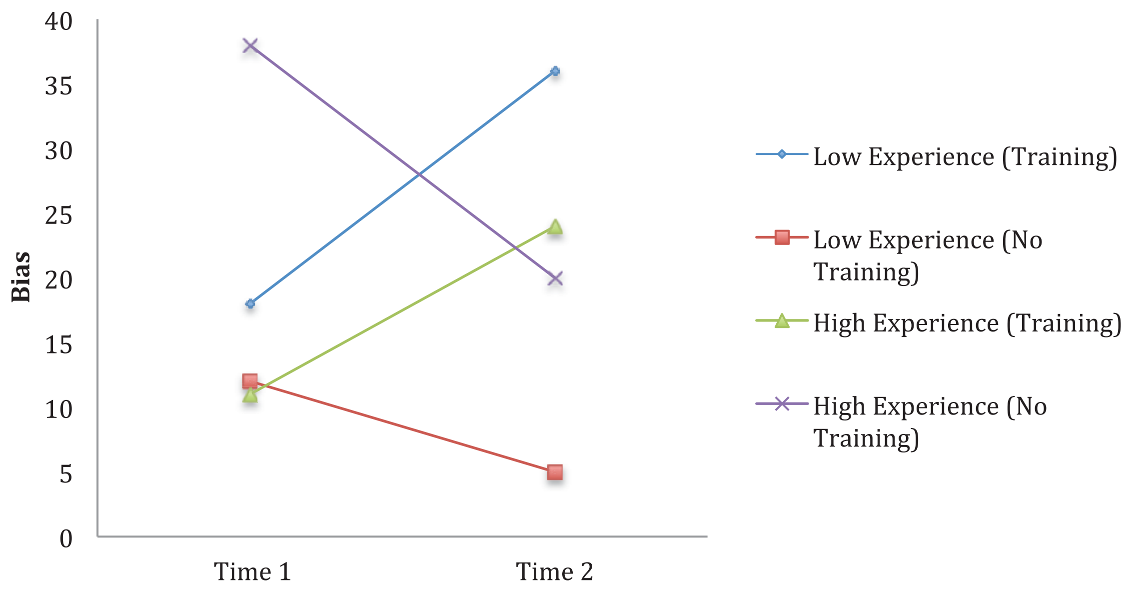



*Figure 32.* Change in response bias from Time 1 to Time 2 for all groups.

There was no main effect of Time, and no interaction between Time and Experience Level, and no three-way interaction between Time, Training, and Experience Level, all *Fs* < 1. There was a near significant interaction between Time and Training, $F(1,21) = 3.21$, $p = .09$, $\eta_p^2 = .13$, a medium-large effect size, suggesting that Training has an impact on bias with both training groups becoming more conservative after training.

There was no main effect of Training, or Experience Level, both *Fs* < 1.

Mean scores for all groups for *c* were > 0, indicating an overall conservative response bias, meaning participants had a slight bias in favour of responding *genuine*. After training, both Low and High Experience groups demonstrated a more conservative bias, which suggests the training increases the likelihood of responding *genuine,* whereas those who did not receive training tended to become more liberal with their responses. These results are similar to those in Study 1, where all participants also displayed a conservative bias. However, in Study 1, training had no impact on bias.

# CHAPTER TEN: DISCUSSION – STUDY TWO

## SYNOPSIS OF STUDY 2'S AIMS

As for Study 1, the goal of this study was to examine therapists' abilities in recognising emotional facial expressions, whether training enhances this ability, and whether level of experience would impact on therapist performance. This study also explored therapists' ability to detect deception from facial expressions, and whether this, too, could be enhanced by training. Unlike in Study 1, age was not correlated with experience. Therefore, the hypothesis that younger, but less experienced clinicians would outperform their older, more experienced, counterparts was not pursued. As Hypothesis 2 was no longer confounded by the effects of age, it was hypothesized that experienced clinicians would demonstrate greater accuracy at emotion recognition than less experienced clinicians.

Again, an aim of the study was to explore whether therapists were able to detect deceptive expressions at levels above chance, and whether training in recognising subtle expressions of emotion would enhance deception detection. It was hypothesised that, as found in Study 1, therapists would be able to detect deception at above chance levels. It was also of interest to find out if training in key features of subtle emotion expressions would improve deception detection.

Although Study 2 is an attempt to partially replicate Study 1, there are some key differences which meant replication was limited. In Study 2, the emotion recognition training programme used in Study 1 had been updated and the earlier version was no longer available. This posed a major stumbling block to making across-study comparisons. Also, there are several differences in the demographic characteristics of participants. Although there was not a significant difference in the mean age of participants in both studies, $t(66) = 1.35$, $p > .05$, Study 1 ($M = 42.6$ years; $SD = 11.4$), Study 2 ($M = 38.6$ years; $SD = 9.73$), there was a significant difference in years of experience, $t(66) = 2.02$, $p = .05$, between the participants in Study 1 ($M = 11.80$ years; $SD = 10.66$) and Study 2 ($M = 7.02$ years; $SD = 8.59$). Also, in Study 1, the therapists with less experience were significantly younger than their more experienced counterparts. This was not the case in Study 2 where there was no significant difference in age between the low and high experience groups. Another key difference between the groups in Study 1 and Study 2 was that the participants in Study 1 mainly worked in mental health settings or in private practice, with only two participants (5%) working for the Department of Corrections. In Study 2, eleven (48%) participants worked for the Corrections Department whilst the remainder worked in private practice or for the District Health Board.

To increase the power of the overall study, consideration was given to pooling the data collected in both studies and performing ANOVAs. However, due to the differences in methodology and participant characteristics as outlined above, this was considered to be inappropriate. Furthermore, combining the data collected from two individual studies after analyses have been performed may also be considered “data peeking” (Hartgerink, Van Aert, Nuijten, Wicherts, & Van Assen, 2016). The temptation to do this can occur where significance testing is conducted, revealing encouraging but non-significant results, leading researchers to collect more data. This practice is discouraged (see Simmons, Nelson, & Simonsohn, 2011) as augmenting a dataset can increase the chance of a Type 1 error.

In the following discussion, comparisons are made between Studies 1 and 2. However, in light of the fundamental differences between the studies, caution is required in interpreting these.

## FINDINGS FROM THE EMOTION RECOGNITION TASK

### Single Emotion Expressions

Training in recognising single expressions of happiness[5] had no significant effect on scores but there was a small-medium effect (non-significant) of an interaction between Time and Experience level. This was demonstrated by the scores of those with most experience increasing with training, whereas the less experienced therapists' scores decreased with training. For the two groups that were not trained, a similar pattern was evident, with the more experienced therapists' scores remaining unchanged, whilst the low experience therapists' scores showed a similar pattern to the low experience (trained) group with a decline in accuracy over time noted.

Overall, high accuracy levels for detecting happiness were also noted, with scores ranging from 90 to 96 percent accuracy, which can impose ceiling effects limiting the effects of training. This is unsurprising as happiness is reported to be the most universally recognised expression (Ekman & Friesen, 1971). With regard to Hypothesis 1 that training enhances emotion recognition, there was no effect of training for either group; therefore, this hypothesis was not supported.

These results when compared with those from Study 1, show a similar pattern for High Experience participants, whereby Training increased scores from Time 1 to Time 2, and for those who did not receive training, their scores remained static over time in both studies. However, with regard to less experienced participants, a different pattern emerged. Previously, the Low Experience group's scores remained unchanged from Time 1 to Time 2 after training, whereas in the current study, their scores declined at Time 2. The untrained Low Experience group's scores also declined at Time 2, whereas in the previous study they had increased slightly. However, there was a difference in year's of experience between the Low Experience groups in each study. Firstly, the Low Experience group in Study 1 was

[5] Caution is needed when interpreting the happiness effect due to the data violating the normal distribution assumption.

more experienced, with a significant difference in the mean year's of experience t(34) = 2.37, $p < .05$, between the participants in Study 1 ($M$ = 3.16 years; $SD$ = 2.42) and Study 2 ($M$ = 1.82 years; $SD$ = 0.87). There was no significant difference ($t(34)$ = .39, $p > .05$) in age between the two groups. Therefore, the difference between the two groups from Study 1 and Study 2 may be due to those in Study 2 being considerably less experienced when compared with those in Study 1.

As for Study 1, in relation to recognition of single sad expressions, there was a significant and large effect of training for both the highly experienced and less experienced groups with increases in scores from Time 1 to Time 2 of 28% and 15%, respectively, supporting Hypothesis 1. There was a non-significant small-medium interaction between Training and Experience level, which is demonstrated by the high experience group increasing their scores more following training. Unlike in Study 1, no practice effect was seen. As for single Happy expressions, the high experience (untrained) group's scores remained the same over time, whereas the low experience (untrained) group's scores declined over time. As mentioned previously, the SETT training draws attention to reliable signs of sadness, which appear to be quite distinct and once learned, can be readily recognized. It is interesting to note that in the current study, the high experience group's scores increased by 28 percentage points, compared with 18 points previously in Study 1, whereas the low experience group score's increase was similar to that of previously (15% vs 17%). However, this comparison may be confounded by the more experienced participants in Study 1 being significantly older than those in Study 2 t(30) = 2.27, $p < .05$, Study 1 ($M$ = 50.57 years; $SD$ = 8.77) and Study 2 ($M$ = 42.82 years; $SD$ = 10.06). Therefore, although more experienced participants benefit more from training when compared with those with less experience, younger more experienced therapists seem to perform the best. There was also a significant difference in years of experience between the two groups, t(30) = 2.72, $p < .05$, Study 1 ($M$ = 20.86 years; $SD$ = 8.01) and Study 2 ($M$ = 13.64 years; $SD$ = 9.50) making direct comparisons across studies unreliable.

Although the high experience group demonstrated the most benefit from training, their score at Time 2 was almost the same as the Low Experience (Training) group (70% vs. 69%). This is similar to the results obtained in Study 1, where training resulted in participants reaching a similar level of competence after training, irrespective of their skill level prior to training. It is possible that there is a common level of competence that can be achieved through training, rather than training developing participants to a higher level of expertise. However, in both Study 1 and Study 2, each participant only completed 90 mins of training. It would be interesting to explore whether an extended period of training, or repetition of training at regular intervals with 'real life' practice in between, would increase scores further.

Hypothesis 1 was supported with training being effective for improving recognition of single Sad expressions. Hypothesis 2 was not supported, with no difference in skill level noted between Low and High Experience groups.

With regard to recognition of single Disgust expressions, all groups showed a significant improvement in scores at Time 2, irrespective of whether or not they received training. There was a modest interaction between Time and Training, and also between Training and Experience level, which was demonstrated by both High Experience groups score's increasing more over time compared with those with less experience, but interestingly, the High Experience (no training) group increased the most. However, their score at Time 2 was the same as those with Low Experience who received training. It would seem, therefore, that with regard to recognising single expressions of Disgust, given the large and significant effect of Time, practice is most influential in increasing skill at recognition of this emotion.

When compared with Study 1, these results show a different pattern. Previously, training only improved accuracy for high experienced individuals, and adversely affected the performance of those with low experience. No practice effect was observed in Study 1. It was posited that the reason for the decline on performance in Study 1, was the younger age of the less experienced therapists, given the literature has reported that disgust recognition improves with increasing age (e.g., Calder et al., 2003; Isaacowitz et al., 2007; Suzuki et al., 2007). However, in Study 2, there was no age difference between the Low and High Experience groups, which may account for the different results recorded by the Low Experience participants.

Hypothesis 1 was not supported for training improving recognition of single Disgust expressions. Hypothesis 2 was not supported.

## Masked Emotion Expressions

Once again, as found with recognition for single happy expressions, a ceiling effect occurred for recognition of masked happy expressions, and this was particularly noticeable for one group (Low Experience, no training) which scored 100% at Time 1 and Time 2. Unlike in Study 1, no practice effect was noted in this study. The results for recognition of this emotion need to be regarded with caution. As mentioned previously, a large limitation is that there were only three masked happiness expressions, due to the difficulty obtaining video clips of this expression. For this expression type, Hypothesis 1 was not supported as training did not enhance recognition of masked happiness. With regard to Hypothesis 2, there was no effect of experience level on accurate recognition of masked happiness.

For recognition of masked sad expressions, training had a greater impact on the accuracy of more experienced therapists compared with the less experienced ones. This contrasts with the results found in Study 1, where those with less experience improved more with training. After training, the High Experience group scored 71%, compared with the next highest groups which both scored 45%, showing a greater improvement in recognition of masked sadness for those with more experience when compared with the other three groups.

As training was only effective for those with a high level of experience, Hypothesis 1 is only partially supported. Hypothesis 2 was also only partially supported as the effect of experience level was only evident after training.

For masked disgust expressions, training was not effective for either of the groups. There was also no practice effect, and no difference in skill due to level of experience. Therefore, Hypotheses 1 and 2 are not supported for this emotion type.

This finding is unsurprising as the literature (e.g., Calder et al., 2003; West et al., 2012) consistently reports a disgust recognition advantage amongst older adults. However, in this study, there were no age differences between the groups and therefore differences in skill level would be unlikely.

## Leaked Emotion Expressions

The results found for accurate recognition of leaked happiness were similar to those observed in Study 1. The interaction between Time and Experience was significant, and there was also a medium effect of an interaction between Time and Training, although this was not significant. This was demonstrated by both High Experience groups improving over time, with the training group improving the most. For the less experienced, there was a small decline in scores over time irrespective of whether training occurred or not.

Hypothesis 1 received partial support in that training was effective, but only for those with more experience. Hypothesis 2 was also partially supported as the more experienced participants were more accurate at recognising this emotion, but only after training.

For leaked sad expressions, the results were similar to leaked happiness, in that training was only effective for more experienced participants whose scores increased over time by 34 percentage points. The other three groups' scores remained fairly static from Time 1 to Time 2. The more experienced group who received training demonstrated greater skill at recognising leaked sadness. These results contrast with Study 1, where no effect of training was found for recognition of leaked sad expressions.

In relation to recognition of leaked disgust, it appears that practice has more effect than training, with an almost significant, medium-large effect of Time. This is similar to the results found for single disgust expressions, although for single expressions, all four groups improved with practice, whereas for leaked disgust, the Low Experience group which received training declined in accuracy over time. There was also some interaction between Time and Experience, with the scores for both groups with more experience increasing over time, whereas only the untrained low experience group improved with time. For recognition of leaked disgust, Hypothesis 1 was not supported as only time (practice), not training, improved scores. There was also no main effect of experience, therefore no support for Hypothesis 2.

As for leaked sadness, these results contrast with Study 1 which found no effect of training or experience. However, similar to Study 1, the scores for detecting leaked disgust were also greater than recognition of leaked happiness or sadness, demonstrating the recognisability of this particular emotion. For example, the scores for recognition of leaked happiness ranged from 6% to 18%, for leaked sadness 8% to 41%, but recognition for leaked disgust ranged from 40% to 78%. This suggests that even without training, leaked disgust is readily detectable, and that practice alone enhances recognition.

ten Brinke and Porter's (2012) study of genuine versus deceptive pleaders reported that deceptive pleaders were more likely to express disgust in the lower facial regions (40.0% vs. 16.3%). It is possible, therefore, that more sensitive recognition of this emotion may be helpful when working with populations who may be deceptive, such as within a corrections setting. However, as noted by ten Brinke and Porter, the expression does not reveal its source, and therefore further investigation would be required to determine the cause of the expression.

## SUMMARY OF EMOTION FINDINGS

In Study 2, training was most effective for enhancing recognition of sad emotions, with a significant and large effect for participants of both experience levels for recognising a single sad emotion. However, with regard to masked and leaked sadness, training was only effective for participants with a high level of experience. For happiness, training was only effective for highly experienced participants, and only for recognition of leaked happiness, not for single or masked happy expressions. With regard to disgust recognition, although there was no effect of training, a practice effect was observed for recognition of single and leaked disgust, with this being more notable for those with a high level of experience.

In summary, as was the case in Study 1, these results show that training is more effective for those with a higher level of experience, except for training in recognition of single sad expressions which was effective for both experience levels.

As proposed for Study 1, the reason for the greater effect being dependent on experience appears to be that the lower scores exhibited by the more experienced group prior to training provided greater scope for improvement, and at Time 2, scores for those with more experience were not significantly higher than the less experienced groups. Except for masked and leaked sadness, where the High Experience training group's scores after training were 16 and 22 percentage points higher than the next group, respectively.

In Study 1, it was suggested that the difference in age between the Low Experience and High Experience groups may have accounted for the different scores. However, in Study 2, even where there was no age difference between low and high experience groups, the more experienced clinicians again showed the most benefit from training compared with their less experienced counterparts. Furthermore, in some instances, both practice and training led to declines in performance for those with less experience. For this reason, there does not appear to be any benefit in offering facial emotion recognition training to less experienced clinicians. However, for those with more experience, training enhanced accuracy in recognising facial expressions.

## FINDINGS FROM THE DECEPTION DETECTION TASK

### Sensitivity *P(A)*

As for Study 1, participants' accuracy at correctly discerning deceptive facial expressions from genuine expressions was measured using the non-parametric measure, *P(A).* Both High Experience groups' mean accuracy scores increased over time, with the trained group showing the greatest increase. Whereas both low experience groups declined over time, irrespective of training, with the low experience group which received training performing the worst overall. The ANOVA revealed the interaction between time and experience level was significant with a large effect, as reflected by the results as described above. However, unlike in Study 1 where the group with more experience achieved the highest scores at Time 2, in the current study, they were not, overall as a group, better than the less experienced therapists at detecting deception.

Overall, the results for accuracy at deception detection were very similar to those found in Study 1. The range of mean scores in Study 1 was 60% to 69%, whereas in Study 2 it was 59% to 68%. There were also large variations in individual accuracy, ranging from 36% to 77%, compared with Study 1 where individual scores ranged from 44% to 82%. Once again, although there is individual variation, this suggests that overall as a group, therapists perform at above chance levels for detecting emotional deception and this finding is consistent with the literature (e.g., Ekman et al., 1999; Wright Whelan et al., 2015).

### Bias (*c*)

As for Study 1, all participants were more inclined to respond *genuine* rather than *fake,* representing a conservative bias. In Study 1, there was no effect of training; however, in the current study, there was a medium-large effect of training (non-significant), whereby training increased a conservative bias for both high and low experience therapists. Interestingly, for the two groups that did not receive training, their response bias became more liberal over time.

In summary, the current study, unlike Study 1, found experience level had no impact on deception detection ability, which is consistent with results reported in the literature (e.g., Ekman & O'Sullivan, 1991; Vrij & Mann, 2001a; Wright Whelan et al., 2015) where no correlation was found between accuracy for veracity and professional experience. With regard to bias, the results of Study 2 were similar to those from Study 1 in that all participants displayed a conservative bias. However, in Study 2, the training increased the tendency towards conservatism, irrespective of experience level whereas in Study 1, training had no impact.

Although it has been highlighted that there are some fundamental differences between the studies necessitating caution when making comparisons, perhaps the most consistent finding across Studies 1 and 2 is that the more experienced clinicians benefit the most from training in recognising facial emotions.

Some explanations for the differences between the low and high experience groups in Study 1 could possibly be accounted for by the differences in age between the groups, but age cannot explain the differences found in Study 2. A number of trainee characteristics related to training outcomes have been identified in the literature (e.g., Tziner, Fisher, Senior, & Weisberg, 2007), some of which may account for the differences between the low experience and high experience groups in the current study. Self-efficacy, an individual's belief in success, can influence training outcomes and was found by Tziner et al. to be significantly related to training grade. It is possible that the more experienced clinicians have experienced more success over the course of their career, and therefore expect to improve their skills with training. It was also suggested by Tziner et al. that self-efficacy enhances motivation to learn.

Motivation to learn and to persist when training material is difficult was hypothesized by Noe (1986) to influence training outcomes, and was reported by Tziner et al. (2007) to have a significant positive effect on training grade. Perhaps the more experienced clinicians were more motivated to acquire new skills than their less experienced counterparts, and therefore invested greater effort. Many of the less experienced therapists had completed only a few years of practice since completing clinical training and were possibly less open to learning new skills due to continuing to develop and consolidate nascent clinical skills.

Test anxiety is known to be a major disruptive factor in test performance with studies relating test anxiety to significant performance decrements (Culler & Holahan, 1980). Clinical training involves repeated assessment which may increase test anxiety in students due high expectations and the consequences of not performing well. Less experienced clinicians may still experience a certain level of test anxiety due to the proximity to their clinical training, which may have adversely affected their test performance. Instrumentality of training also influences training outcomes. An individual's perception that training will lead to valued outcomes, such as new work related skills, can influence achievement. It is possible that the more experienced clinicians were more able to identify the relevance of learning facial emotion recognition skills, compared with their less experienced counterparts, which enhanced the effect of training.

## SUMMARY OF DECEPTION DETECTION FINDINGS

Although in Study 2 training had no effect on increasing accuracy for deception detection, there was a large interaction of Time and Experience Level, and a small-medium interaction of Training and Experience Level, demonstrating once again, a difference between the low and high experience groups, with accuracy increasing over time for those with more experience and reducing over time with less experience. A similar pattern was seen in Study 1 although there was a stronger advantage of experience increasing deception detection accuracy.

In Study 1, there was a positive correlation between years of experience and deception detection, ($r$ = .39, $N$ = 43, p < .01) at Time 2, a finding which was contrary to Ekman and O'Sullivan (1991) who reported a negative correlation between experience level and deception detection. However, the results obtained in Study 2 revealed no correlation between years of experience and deception detection accuracy, ($r$ = .01, $N$ = 25, $p > .05$). Even though more experienced participants obtained greater benefit from training, overall accuracy levels were similar for both low and high experience therapists. Further research with an increased number of participants is needed to resolve these different outcomes for the important potential association between deception detection and level of clinical experience.

## EVALUATION OF NEW SETT VERSION

As previously stated, the purpose of this second study was to carry out a preliminary evaluation of the effectiveness of the updated version of the SETT training tool. Only very limited comparisons across the two studies can be made due to the major differences between the two participant groups with regard to age and experience level.

Both SETT versions showed that training led to improved recognition of most emotion types, and mainly for those with more experience. The second version of SETT was more effective for training recognition of all three types of sad emotion expressions. With regard to detecting deception, the results were reasonably similar across both studies, in that training interacts with experience level, increasing accuracy amongst high experience participants whereas low experience participants decreased in accuracy over time, irrespective of training.

That the two versions of SETT resulted in reasonable agreement on the outcomes mentioned above, even though the samples were small and quite different, suggests that SETT, Version 2 may be robust in dealing with substantial variations in sample types and sample sizes.

# CHAPTER ELEVEN: LIMITATIONS, STRENGTHS, and SUGGESTIONS for FUTURE RESEARCH

## LIMITATIONS

On completion of these two studies, various limitations became apparent that might have influenced the results.

### Stage 1. Creation of Emotional Stimuli

This study set out to measure therapists' ability to recognise facial expressions of emotion. Expressions of emotion that occur in the social world are generally embedded in a particular context, and are accompanied by verbal and other non-verbal information such as body language and posture. For these reasons, developing a set of stimuli with ecological validity that accurately assesses the phenomenon of facial emotion expressions has many challenges. Although the use of dynamic rather than static images in the current investigation is one step towards increasing validity, there were a number of flaws in the stimuli used. Firstly, expressions were made by individuals who were alone in a psychology laboratory, thus stripped of a social context. And secondly, the vantage point was full-frontal, where often social communication occurs from a variety of positions.

The IAPS images used to evoke emotional responses for the creation of emotion expressions have been normed on a large sample; as well, the slides were selected from discrete emotion categories provided by Mikels et al. (2005). However, there was no qualitative check with participants to determine what emotion they actually felt whilst responding to the images. Also, the request to express a particular emotion may have reduced the spontaneity of the expression, even when congruent with the underlying emotion. As a result, these expressions would not have met Matsumoto et al.'s (2012) definition of a naturally occurring expression which is one *not* produced by request. To produce a more natural, spontaneous expression, an instruction to "respond as you naturally would to the images" along with eliciting from the participant what emotion they were feeling could have produced more authentic expressions. A stimulus check could have been done following the presentation of the slideshow by having participants view the slideshow for a second time, and choosing an emotion term and intensity rating that matched their experience. A similar procedure was performed by Porter et al. (2011) where they asked participants to rate their own emotional reactions (valence and arousal) to the images.

There was also no examination of participants' self-consciousness about being videoed, which could have affected their emotion expressions. A self-consciousness rating similar to that used by Malatesta et al. (1987) could have revealed if the participants were experiencing any discomfort. But, it is also possible that the use of a questionnaire about self-consciousness might increase camera shyness. Despite this possible limitation, the process used mitigated the risk of self-consciousness influencing emotion expressions. All participants were alone when video-recorded which may have helped to put them at ease. Furthermore, during an informal debrief after recording had taken place, participants were asked if they experienced any discomfort about being filmed. Almost all reported that the slideshow was so engrossing that they rapidly forgot about the camera filming them. The discrete use of the laptop's inbuilt camera and a familiarity of this cohort of participants with commonly used technology utilising facial cameras, such as Facetime and Skype, are all likely to have diminished any self-consciousness aroused through filming.

Financial limitations prevented obtaining a larger set of video clips from which a more balanced set for each emotion could be drawn. As far as possible, an attempt was made to balance each emotion type within each veracity category, with priority given to evenly balancing the veracity categories. That way, an observer would obtain only a chance total accuracy score if they were to judge all expressions as deceptive, or all as genuine. As mentioned previously, happy expressions are the most successful masking expression, resulting in a larger number of single happy expressions to select from. Whereas there were very few happy expressions where leakage of the underlying emotion was evident. As expected, in line with previous research on expression of insincere emotions, a large number (15 out of 25) of the fake expressions contained some leakage. As it appeared to be difficult for expressers to fake an expression without leakage occurring, to produce more single faked emotion expressions, neutral slides could have been utilised. Only two of the genuine expressions showed leakage of a second emotion. Both of these were a combination of disgust and sadness, and may be an artefact of the expresser experiencing both of these emotions simultaneously; a qualitative check would have determined if this was the case. This blending of two expressions occurred despite using only images from discrete emotion categories as defined by Mikels et al. (2005). It is possible this resulted from there being some residual emotion still felt from a previous slide. Although it was not a primary aim of the present study to confirm Darwin's inhibition hypothesis, considerable leakage of deceivers' felt emotions occurred during the emotion induction process, lending support to this hypothesis.

Another factor that may have influenced the present results was that the mean age of the people expressing emotions in the video-clips was only 22.3 years. This may have given the younger, less experienced participants an advantage, as studies have found younger adults outperform older adults when target expressers are also young, but no age-related differences are noted when recognising emotions in older adults (Richter, Dietzel, & Kunzmann, 2011). Also, Murphy et al. (2010) reported that older adults performed better than younger adults when targets were a mix of both young and older faces. If a mixture of younger and older faces had been used, this may have resulted in better performance by the more experienced therapists.

The use of low-stakes deception is a serious constraint. The literature has described how cues to deceptive behaviour increase when the deceiver is motivated to succeed. In the current study, there was no incentive for expressers to be convincingly deceptive which may have influenced their facial behaviour, and, for ethical reasons, it is difficult to generate high-stakes deception in a laboratory setting. The expressers in the video stimuli formed their facial expressions for the experiment; as a result, although they were also exposed to emotionally evocative stimuli (images and music), this is unlikely to have replicated what would occur during a real-life, highly emotive situation. As emotionally intense lies increase clues for detection, participant emotions could have been intensified by employing a user-controlled slideshow. This would have allowed participants to spend as long as they chose to when looking at the images which might have intensified the emotional experience, resulting in more leakage and/or inconsistent expressions. Also, viewing a range of images selected from different emotion categories within only two mins could have resulted in a cross-over of emotion experiences. An alternative would have been to video-record participants watching a short film containing emotionally arousing material which could allow for more natural development of emotions by providing a richer context. Having participants report their emotional experience would provide information about the emotion(s) subjectively experienced. Another option could have been to use more authentic facial stimuli capturing spontaneous and posed expressions. For example, Matsumoto et al. (2012) studied video-tapes of athletes expressing emotions at the end of medal matches and Machado et al. (1999) used segments of video-recordings of client therapy sessions.

Despite this limitation arising from using a low-stakes induction process, many of the expressions obtained during the video-recording process contained evidence of emotional leakage during masking of expressions as defined by the FACS coder, and consistent with what was being viewed on the slideshow. Furthermore, the deception detection scores above chance were similar to those recorded in other studies using high-stakes deception.

**Stage 2: The Emotion Recognition and Deception Detection Task (DDERT)**

Another limitation is the format in which the data were collected. Each participant completed the DDERT task independently, at different times of day, and in different locations. Participants may have experienced fatigue when conducting the DDERT task at Time 2, particularly those in the training group, as this followed 90 min of on-screen training which may have resulted in tiredness adversely affecting their performance. Whereas for the non-training group, they were instructed to continue "business as usual" which may have been less demanding than concentrating on training. Also, some of the participants completed the training at home during the evening after work. This is not an optimum time to engage in training and may have reduced the efficacy due to fatigue.

Participants in the no-training group were requested to occupy themselves with "work as usual" for the length of the training period with the exception of avoiding any face-to-face contact with clients during that time. As the data collection was conducted remotely, the researcher relied upon the participants to adhere to the instructions given.

**No baseline**

In normal social interactions, participants are exposed to the others' full repertoire of behaviour during which deceptive and honest communication occurs. In the current study, the observers were all unfamiliar with the expressers. As it has been reported that veracity judgements are increased by familiarity with a communicator's truthful behaviour (Brandt & Miller, 2009; Ekman & Friesen, 1974), this may have impacted on the results. However, in some real-life situations, snap judgements need to be made without any prior communication; for example, a duty officer examining a client during a crisis interview.

**Complexity of task**

While previous research has largely focused on either emotion recognition or deception detection, the current study investigated both of these skills simultaneously. It could be that the increased complexity of the task resulted in lower scores than might have been obtained if the two constructs of interest had been investigated separately. If the order of the two tasks had been randomised, a check could have been made to see if either coming first influenced the second.

Finally, the study only examined accuracy for three emotions (Happy, Sad, Disgust). This narrow focus limited the evaluation of the effectiveness of the intervention which provided training in seven universal emotion expressions. Additionally, observations of expressions were limited to 5 s, with no other normally observed cues such as vocal tone or body language. In a natural setting, a more extended period of time could be expected coupled with various contextual clues. However, with regard to detecting veracity of expressions, previous studies have not found that interviewing increases accuracy, with passive observers being more accurate in detecting deception than active interviewers. Also, interviewing increases a truth bias and is detrimental to detecting deceit (Vrij & Mann, 2001b).

**SETT**

The first study used the original version of the Subtle Emotion Training Tool (SETT). Whilst this presented information about the seven universal expressions: Happy, Sad, Anger, Fear, Surprise, Contempt, and Disgust, these were all displayed by the same young Caucasian woman. The rationale presented by the training programme was that using one face would eliminate the distractions that may occur when looking at different facial features of a number of faces. In April 2014, an updated version of SETT was released which has superior features when compared with the original SETT such as more expressions, posed by a range of individuals from both sexes and various ethnicities. The commentary information about features of each expression is also more comprehensive. Study 2 was conducted using this latest release of the SETT.

Both Study 1 and Study 2 only examined the impact of training on recognition of Happiness, Sadness, and Disgust. Therefore, the SETT was only partially tested. It would be interesting to explore whether the present findings would be replicated if other emotions such as anger, contempt, surprise, and fear were included. This would allow for more general conclusions to be drawn in relation to Hypotheses 1 and 2. Although this is desirable, if all emotions had been investigated, the number of stimuli would have been far too many. As it was, the DDERT with 50 clips took approximately 15 mins to complete. Inclusion of other emotion types would have been impractical as the task would have been too lengthy, possibly leading to fatigue and loss of motivation amongst participants.

**Sample size**

Sample size was another limitation of both Study 1 and Study 2. Previous studies have used varying sample sizes, such as *N*=310 (Porter et al., 2002) for investigating deception detection, and *N*=227 in Calder et al.'s (2003) second experiment examining age differences in emotion recognition. The sample size for the present studies was largely governed by the availability of participants. Power calculations determined that a total of 64 people in each group was required to reach a power of .80 for a medium effect size. Thus, both Study 1 and Study 2 were underpowered, increasing the risk of Type II errors occurring. However, the present studies focused on the effect size of results, looking beyond statistical significance. Due to the small sample sizes, detailed analyses were not possible to investigate other factors such as handedness and gender.

**Limited number of deception-interested psychologists**

Finally, there are issues with regard to the domain of practice of the clinician participants. In Study 1, many of those recruited were employed in either private practice or working for district health boards, where there may be less emphasis on being alert to deception. Of the 43 participants, only two worked for the Department of Corrections, where there is likely to be a stronger interest in detecting client deception, and also possibly more opportunities to develop skill in detecting deception. As a result, clinicians who may have a stronger interest in this area were not well represented in Study 1. Although in Study 2, eleven corrections psychologists participated, this is still a relatively small number when compared with Ekman's et al.'s (1999) study whose deception interested psychologist sample was 107. As Ekman et al. found, psychologists with a special interest in deception outperformed other groups of psychologists. Future work could include a larger number of psychologists working in forensic or corrections settings.

As has been outlined in the sections above, there is a large number of methodological factors that influence the outcome measures used when investigating both ERA and deception detection. Firstly, in relation to ERA, there is no consistency across studies when defining age groups to explore the impact of age on ERA. Secondly, the stimuli used have ranged from posed or genuine emotion expressions displayed in static, photographs, in either colour or black-and-white, to video clips of varying length, or computer generated morphed images. For investigating deception detection, many studies have used videos of people giving honest or deceptive accounts. Not all of these have used a measuring system to determine whether differences exist between the two levels of veracity. Some have used low-stakes methods, whilst others have obtained stimuli involving high-stakes pleaders. Finally, there is varying methodology in the analysis of accuracy, with some studies reporting percentage correct, whilst others have used SDT statistics. Consequently, it is perhaps unsurprising to find such variable outcomes across studies.

## STRENGTHS

Despite the limitations outlined above, this study is the first known to examine clinical psychologists' emotion recognition abilities and whether a brief intervention can enhance these skills. It is also one of very few studies that have evaluated deception detection amongst this group in recent years. A strength of the study was the use of the FACS coding system to objectively categorise the facial expressions, rather than relying on subjective judgements. FACS is able to discern between genuine and masked expressions by identifying the absence or presence of reliable facial muscle movements relative to the emotion being expressed (faked or genuine). FACS is widely acknowledged to be one of the most comprehensive and systematic methods for scoring facial behaviour (Matsumoto, 2004).

The majority of studies investigating emotion recognition accuracy have used still photographs of high-intensity, full-face, single emotion prototypical expressions, often in black and white. However, most every day spontaneous expressions tend to be more subtle, and often include blends of emotions. The use of dynamic, rather than static expressions added a more realistic element in the present investigation, particularly as the expressions used included subtle emotion blends rather than full-face, high intensity expressions. As found by Bould et al. (2008), recognition of subtle expressions is enhanced by movement, hence the use of dynamic rather than static expressions. Also, the study used full-colour video-recordings as expressions shown in black-and-white can alter the perceived intensity of emotion and be seen as more expressive compared with colour images (Barr & Kleck, 1995).

Although the universality of expressions has been strongly argued, facial expressions are still subject to culturally-bound 'display rules' (Ekman et al., 1987) and 'in-group' advantages (Elfenbein & Ambady, 2002). The use of a broad cultural group in the creation of the facial expression video clips was representative of New Zealand's multi-cultural society and of the population that psychologists are likely to encounter in therapy practice. Although culturally representative, the participants that created the video stimuli were all university students of a narrow age range (mean age = 22.3 years). Differences have been found in research whereby older adults demonstrated an advantage over younger adults in recognising expressions of young and old expressers, compared with only young expressers. Thus, the use of a more diverse age group, which is more representative of clients presenting for therapy, would have been preferable. Nevertheless, there was a good balance between the sexes (9 male and 10 female). Of the 53 clips selected for the DDERT, 33 showed female expressions. It was not possible to equally balance the sexes as priority was given to balancing veracity first, followed by emotion categories.

**Signal detection**

Using SDT as a measure of accuracy has a theoretical advantage over the use of percentage correct scores, an approach adopted by many studies. SDT provides a sensitivity measure which is independent of response bias, and, as has been demonstrated above, the participants in the current study all exhibited an inclination to respond *genuine* rather than *fake* to the expressions they viewed. However, despite using a different approach to that of other studies, it is still meaningful to make comparisons with other results. As a part of their meta-analysis, Bond Jr and DePaulo (2006) compared the accuracy measures of percentage correct, log odds ratio, and *d′*, reporting that the three accuracy measures were very highly correlated.

## SUGGESTIONS FOR FUTURE RESEARCH

Further research could investigate whether deception is more detectable in individual emotions. For example, in Porter et al.'s (2011) study, in addition to exploring inconsistent emotional expressions that occur during deception, they examined observers' accuracy in judging the veracity of each separate emotion expression. They found that observers were only able to detect deception in happiness at levels above chance compared with other emotion types.

As emotion recognition has been theorised to be one of the core components of the construct of emotional intelligence (EI; Mayer et al., 2008), an important avenue for future research could be to examine the convergence of ERA with measures of other abilities linked to EI such as the Empathy Quotient (EQ; Allison, Baron-Cohen, Wheelwright, Stone, & Muncer, 2011), or the expression and regulation of emotions (Elfenbein et al., 2002). Also of importance is not just whether people can recognise facial expressions, but also how they react to and use that information (Ekman, 2007).

It is possible that the role of a therapist requires one to be a high self-monitor, that is, to monitor and control one's expressive behaviour and self-presentation whilst being sensitive to others' expressions and presentation in social situations. Snyder (1974) found that individuals high in self-monitoring more accurately recognised facial expressions compared to those with low self-monitoring capabilities. The therapist role also provides a lot of face-to-face contact with clients as well as colleagues. It would be interesting to investigate further the relationship between self-monitoring and emotion recognition abilities. A small relationship between self-monitoring and deception detection has also been reported (Aamondt & Custer, 2006).

## CONCLUSION

The resurgence of interest in emotion research has led to understanding the importance of emotion in psychotherapy. Exploring emotion is seen as foundational to the therapeutic process as people constantly try to make sense of their emotions. Accordingly, accurately recognising affective expressions is key to assisting clients to understand and use their emotions.

### Value of Emotion Recognition Training

Given the potential significance to the therapeutic relationship of accurately recognising emotions in others, it is surprising that direct empirical analysis of this issue has been so neglected. It has been amply demonstrated that denied or suppressed emotions can reveal themselves in the face involuntarily. Self-reports of emotions are not always accurate, either intentionally or not, whereas the face can reveal the underlying emotion. Current clinical psychology training programmes emphasise the importance of building a therapeutic alliance and responding empathically to clients' emotions. Despite this, there is no specific element of attending to facial cues to enhance communication, with the focus being largely on verbal accounts. As emotion recognition is dependent on both verbal and non-verbal channels, it makes sense, then, to include training in non-verbal cues.

There is a lack of literature on training students of applied psychology in recognition of emotions from facial expressions. Although various authors, (e.g., Foa & Kozak, 1986; Pos et al., 2003) describe the importance of accurate use and reflection of emotions in therapy, there has been little research assessing clinicians' ability to accurately recognise facial emotion expressions, or on whether inclusion of such skills would be beneficial to applied psychology training programmes. However, given the findings from both studies presented here, it would appear that with regard to Study 1, less experienced and younger clinicians are able to recognise emotion expressions more accurately than older, more experienced clinicians, without training. In Study 1, scores of more experienced clinicians were worse before training, and following training, were similar to those of less experienced clinicians. However, in some cases, the less experienced participants' scores declined with training. Therefore, there does not appear to be any benefit in including facial emotion recognition training in clinical training programmes, for both young and inexperienced psychologists. However, the results from both Study 1 and Study 2 demonstrate that training is effective for older, and more experienced psychologists, so could therefore be utilised as part of ongoing professional development training.

**Value of Deception Training**

In general, therapists value honesty and attempt to work collaboratively with clients in order to establish a therapeutic alliance. Nonetheless, undetected deception can potentially damage this relationship and there are risks for therapists who may be duped by deception, for example, anger or embarrassment (Curtis & Hart, 2015). Often clients deceive for reasons such as protecting others or impression management and the purpose of investigating whether training can enhance deception detection is not to train psychologists to be professional lie detectors, but rather to enhance their observation of the client and be alert to when deception is occurring in order to assist the client to evaluate the function of the deception. Previous research (e.g., Shaw et al., 2013) suggests that training programmes focusing on emotional cues to deceit can enhance detection of high-stakes and emotionally intense deception from a level of chance to 80%. Although their training was more comprehensive than that used in the current study, the brief intervention employed here did enhance deception detection, particularly for older, more experienced therapists who were already performing at above chance level.

In conclusion, considerable research has focused on facial emotion recognition in a broad range of populations, yet very little has examined therapists' skill in this area, despite the likely importance of this in the therapeutic alliance. Furthermore, many studies have used single, usually static, facial expressions, whereas in therapy, emotion perception skill relies on interpretation of dynamic, subtle, and blended emotion expressions. This study attempted to fill this gap in the literature by examining the skill of therapists using more ecologically valid stimuli, and also by investigating whether training can enhance these skills. Also of interest was the impact of experience level on emotion recognition skill, and therapists' deception detection skills, both of which are topics that have been neglected in the research literature.

The results provide mixed support for the first hypothesis, that training would enhance emotion recognition skills. In both Study 1 and Study 2 this effect was mediated by experience level. Secondly, in Study 1, it was found that less experienced, younger, therapists showed superior emotion recognition skills for happiness and disgust, but not for sadness, only partially supporting the second hypothesis. This hypothesis was not explored in Study 2 as there were no age differences between the low and high experience groups. Instead, it was hypothesised that more experienced therapists would demonstrate superior accuracy at recognising emotion expressions. It was found that for most expressions, the more experienced therapists showed greater accuracy for emotion recognition compared with the less experienced therapists. This was particularly noticeable at Time 2, with practice or training enhancing accuracy. The third hypothesis, that therapists would perform at above chance levels for detecting deception, was also supported in both Study 1 and Study 2. It was also found that training generally enhanced this skill in therapists with a greater level of experience relative to those who were younger and less experienced.

# REFERENCES


Aamondt, M. G., & Custer, H. (2006). Who can best catch a liar? A meta-analysis of individual differences in detecting deception. *Forensic Examiner, 15*, 6-12.

Adolphs, R., Damasio, H., Tranel, D., & Damasio, A. R. (1996). Cortical systems for the recognition of emotion in facial expressions. *Journal of Neuroscience, 16*, 7678-7687.

Allison, C., Baron-Cohen, S., Wheelwright, S. J., Stone, M. H., & Muncer, S. J. (2011). Psychometric analysis of the Empathy Quotient (EQ). *Personality and Individual Differences, 51*, 829-835. doi: 10.1016/j.paid.2011.07.005

Ambadar, Z., Schooler, J. W., & Conn, J. F. (2005). Deciphering the enigmatic face: The importance of facial dynamics in interpreting subtle facial expressions. *Psychological Science, 16*, 403-410.

Ansfield, M. E., DePaulo, B. M., & Bell, K. L. (1995). Familiarity effects in nonverbal understanding: Recognizing our own facial expressions and our friends'. *Journal of Nonverbal Behavior, 19*, 135-149. doi: 10.1007/BF02175501

Archinard, M., Haynal-Reymond, V., & Heller, M. (2000). Doctor's and patients' facial expressions and suicide reattempt risk assessment. *Journal of Psychiatric Research, 34*, 261-262. doi: 10.1016/s0022-3956(00)00011-x

Ashwin, C., Chapman, E., Colle, L., & Baron-Cohen, S. (2006). Impaired recognition of negative basic emotions in autism: a test of the amygdala theory. *Social neuroscience, 1*, 349-363.

Barod, J. C., Koff, E., Perlman Lorch, M., & Nicholas, M. (1986). The expression and perception of facial emotion in brain-damaged patients. *Neuropsychologia, 24*, 169-180. doi: 10.1016/0028-3932(86)90050-3

Barr, C. L., & Kleck, R. E. (1995). Self-other perception of the intensity of facial expressions of emotion: Do we know what we show? *Journal of Personality and Social Psychology, 68*, 608-618.

Benes, K. M. (1995). Review of the Facial Action Coding System. In J. C. Conoley & J. C. Impara (Eds.), *The twelfth mental measurements yearbook*. Nebraska: Buros Institute.

Bergstrom-Isacsson, M., Lagerkvist, B., Holck, U., & Gold, C. (2013). How facial expressions in a Rett syndrome population are recognised and interpreted by those around them as conveying emotions. *Research in Developmental Disabilities, 34*, 788-794. doi: 10.1016/j.ridd.2012.10.011

Blair, R. J. R. (2005). Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness and Cognition, 14*, 698-718. doi: 10.1016/j.concog.2005.06.004

Boecker, L., Likowski, K. U., Pauli, P., & Weyers, P. (2015). The face of schadenfreude: Differentiation of joy and schadenfreude by electromyography. *Cognition and Emotion, 29*, 1117-1125. doi: 10.1080/02699931.2014.966063

Boiger, M., & Mesquita, B. (2012). The construction of emotion in interactions, relationships, and cultures. *Emotion Review, 4*, 221-229. doi: 10.1177/1754073912439765

Bonanno, G. A., & Keltner, D. (1997). Facial expressions of emotion and the course of conjugal bereavement. *Journal of Abnormal Psychology, 106*, 126-137. doi: 10.1037/0021-843x.106.1.126

Bonanno, G. A., Noll, J. G., Trickett, P. K., Keltner, D., Putnam, F. W., LeJeune, J., & Anderson, C. (2002). When the face reveals what words do not: Facial expressions of emotion, smiling, and the willingness to disclose childhood sexual abuse. *Journal of Personality and Social Psychology, 83*, 94-110. doi: 10.1037//0022-3514.83.1.94

Bond Jr, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*, 214-234. doi: 10.1207/s15327957pspr1003_2

Bond Jr, C. F., & Uysal, A. (2007). On lie detection "wizards". *Law and Human Behavior, 31*, 109-115. doi: 10.1007/s10979-006-9016-1

Bould, E., & Morris, N. (2008). Role of motion signals in recognizing subtle facial expressions of emotion. *British Journal of Psychology, 99*, 167-189. doi: 10.1348/000712607X206702

Bould, E., Morris, N., & Wink, B. (2008). Recognising subtle emotional expressions: The role of facial movements. *Cognition and Emotion, 22*, 1569-1587. doi: 10.1080/02699930801921156

Brandt, D. R., & Miller, G. R. (2009). Effects of self-monitoring and familiarity on deception detection. *Communication Quarterly, 28*, 3-10.

Buck, R. (1999). The biological affects: A typology. *Psychological Review, 106*, 301-336.

Cain, D. J. (2007). What every therapist should know, be and do. Contributions from humanistic psychotherapies. *Journal of Contemporary Psychotherapies, 37*, 3-10.

Calder, A. J., Keane, J., Manly, T., Sprengelmeyer, R., Scott, S., Nimmo-Smith, I., & Young, A. W. (2003). Facial expression recognition across the adult life span. *Neuropsychologia, 41*, 195-202. doi: 10.1016/s0028-3932(02)00149-5

Camras, L. A., Oster, H., Campos, J. J., Miyake, K., & Bradshaw, D. (1992). Japanese and American infants' responses to arm restraint. *Developmental Psychology, 28*, 578-583.

Chapman, H. A., Kim, D. A., Susskind, J. M., & Anderson, A. K. (2009). In bad taste: Evidence for the oral origins of moral disgust. *Science, 323*, 1222-1226. doi: 10.1126/science.1165565

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences.* (3rd ed.). New York: Routledge.

Cohn, J. F., & Ekman, P. (2005). Measuring facial action. In J. A. Harrigan, R. Rosenthal, & K. R. Scherer (Eds.), *The new handbook of methods in nonverbal behavior research [electronic resource]*. New York: Oxford University Press.

Cohn, J. F., Zlochower, A. J., Lien, J., & Kanade, T. (1999). Automated face analysis by feature point tracking has high concurrent validity with manual FACS coding. *Psychophysiology, 36*, 35-43. doi: 10.1017/S0048577299971184

Culler, R. E., & Holahan, C. J. (1980). Test anxiety and academic performance: The effects of study-related behaviors. *Journal of Educational Psychology, 72*, 16-20. doi: 10.1037/0022-0663.72.1.16

Curtis, D. A., & Hart, C. L. (2015). Pinocchio's nose in therapy: Therapists' beliefs and attitudes toward client deception. *International Journal for the Advancement of Counselling, 37*, 279-292. doi: 10.1007/s10447-015-9243-6

Dalili, M. N., Penton-Voak, I. S., Harmer, C. J., & Munafo, M. R. (2015). Meta-analysis of emotion recognition deficits in major depressive disorder. *Psychological Medicine, 45*, 1135-1144. doi: 10.1017/s0033291714002591

Darwin, C. (1872/1955). *The expression of the emotions in man and animals*. New York: Philosophical Library. (Original work published 1872).

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*, 113-126. doi: 10.1037/0022-3514.44.1.113

Demenescu, L. R., Mathiak, K. A., & Mathiak, K. (2014). Age-and gender-related variations of emotion recognition in pseudowords and faces. *Experimental Aging Research, 40*, 187-207. doi: 10.1080/0361073X.2014.882210

DePaulo, B. M., Kirkendol, S. E., Kashy, D. A., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70*, 979-995.

DePaulo, B. M., Malone, B. E., Lindsay, J. J., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*, 74-118.

Deschamps, P. K. H., Schutte, I., Kenemans, J. L., Matthys, W., & Schutter, D. J. L. G. (2012). Electromyographic responses to emotional facial expressions in 6-7year olds: A feasibility study. *International Journal of Psychophysiology, 85*, 195-199. doi: 10.1016/j.ijpsycho.2012.05.004

Dimberg, U. (1997). Facial reactions: Rapidly evoked emotional responses. *Journal of Psychophysiology, 11*, 115-123.

Dimberg, U., Andréasson, P., & Thunberg, M. (2011). Emotional empathy and facial reactions to facial expressions. *Journal of Psychophysiology, 25*, 26-31. doi: 10.1027/0269-8803/a000029

Dimberg, U., & Thunberg, M. (1998). Rapid facial reactions to emotional facial expressions. *Scandinavian Journal of Psychology, 39*, 39-45.

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

Dimberg, U., Thunberg, M., & Grunedal, S. (2002). Facial reactions to emotional stimuli: Automatically controlled emotional responses. *Cognition and Emotion, 16*, 449-471. doi: 10.1080/02699930143000356

Donlon, T. F. (1995). Review of the Facial Action Coding System. In J. C. Conoley & J. C. Impara (Eds.), *The twelfth mental measurements yearbook.* Nebraska: Buros Institute.

Duchenne, G. B. (1862/1990). *The mechanism of human facial expression (R. A. Cuthbertson, Ed. and Trans)*. New York: Cambridge University Press (Original work published 1862).

Eerola, T., & Vuoskoski, J. K. (2011). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music, 39*, 18-49. doi: 10.1177/0305735610362821

Ekman, P. (1993). Facial expression and emotion. *American Psychologist, 48*, 384-392.

Ekman, P. (2003). Darwin, Deception, and Facial Expression. *Annals of the New York Academy of Sciences, 1000*, 205-221. doi: 10.1196/annals.1280.010

Ekman, P. (2007). *Emotions revealed: recognizing faces and feelings to improve communication and emotional life* (2nd ed.). New York: St. Martin's Press.

Ekman, P. (2009). *Telling lies: Clues to deceit in the marketplace, politics, and marriage*. New York: W. W. Norton & Co.

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review, 3*, 364-370. doi: 10.1177/1754073911410740

Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The Duchenne smile: Emotional expression and brain physiology II. *Journal of Personality and Social Psychology, 58*, 342-353.

Ekman, P., Freisen, W. V., & Ancoli, S. (1980). Facial signs of emotional experience. *Journal of Personality and Social Psychology, 39*, 1125-1134. doi: 10.1037/h0077722

Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry, 32*, 88-106.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology, 17*, 124-129. doi: 10.1037/h0030377

Ekman, P., & Friesen, W. V. (1974). Detecting deception from the body or face. *Journal of Personality and Social Psychology, 29*, 288-298. doi: 10.1037/h0036006

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face*. Englewood Cliffs, NJ: Prentice-Hall.

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., & Friesen, W. V. (1982). Felt, false, and miserable smiles. *Journal of Nonverbal Behavior, 6*, 238-252. doi: 10.1007/bf00987191

Ekman, P., Friesen, W. V., & Hager, J. C. (1978/2002). *Facial action coding system: The manual on CD ROM*

Ekman, P., Friesen, W. V., & O'Sullivan, M. (1988). Smiles when lying. *Journal of Personality and Social Psychology, 54*, 414-420.

Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., . . . Tzavaras, A. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology, 53*, 712-717.

Ekman, P., Friesen, W. V., & Simons, R. C. (1985). Is the startle reaction an emotion? *Journal of Personality and Social Psychology, 49*, 1416-1426. doi: 10.1037/0022-3514.49.5.1416

Ekman, P., Hager, J. C., & Friesen, W. V. (1981). The symmetry of emotional and deliberate facial actions. *Psychophysiology, 18*, 101-106.

Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist, 46*, 913-920.

Ekman, P., & O'Sullivan, M. (2006). From flawed self-assessment to blatant whoppers: The utility of voluntary and involuntary behavior in detecting deception. *Behavioral Sciences and the Law, 24*, 673-686. doi: 10.1002/bsl.729

Ekman, P., O'Sullivan, M., & Frank, M. G. (1999). A few can catch a liar. *Psychological Science, 10*, 263-266.

Ekman, P., & Rosenberg, E. L. (Eds.). (2005). *What the face reveals*. New York: Oxford University Press Inc.

Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science, 164*, 86-88.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin, 128*, 203-235. doi: 10.1037//0033-2909.128.2.203

Elfenbein, H. A., Foo, M. D., White, J., Tan, H. H., & Aik, V. C. (2007). Reading your counterpart: The benefit of emotion recognition accuracy for effectiveness in negotiation. *Journal of Nonverbal Behavior, 31*, 205-223. doi: 10.1007/s10919-007-0033-7

Elfenbein, H. A., Marsh, A. A., & Ambady, N. (2002). Emotional intelligence and the recognition of emotion from facial expressions. In L. F. Barrett & P. Salovey (Eds.), *The wisdom in feeliing: Psychological processes in emotional intelligence*. New York: The Guilford Press.

Endres, J., & Laidlaw, A. (2009). Micro-expression recognition training in medical students: A pilot study. *BMC Medical Education, 9*. doi: 10.1186/1472-6920-9-47

Etcoff, N. L., Ekman, P., Magee, J. J., & Frank, M. G. (2000). Lie detection and language comprehension: People who can't understand words are better picking up lies about emotions. *Nature, 405*, 139.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear. Exposure to corrective information. *Psychological Bulletin, 99*, 20-35. doi: 10.1037/0033-2909.99.1.20

Frank, M. G., & Stennett, J. (2001). The forced-choice paradigm and the perception of facial expressions of emotion. *Journal of Personality and Social Psychology, 80*, 75-85. doi: 10.1037//0022-3514.80.1.75

Gosselin, P., Perron, M., & Beaupre, M. (2010). The voluntary control of facial action units in adults. *Emotion, 10*, 266-271. doi: 10.1037/a0017748

Gottman, J., Levenson, R., & Woodin, E. (2001). Facial expressions during marital conflict. *Journal of Family Communication, 1*, 37-57. doi: 10.1207/S15327698JFC0101_06

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Grimshaw, G. M., Bulman-Fleming, M. B., & Ngo, C. (2004). A signal-detection analysis of sex differences in the perception of emotional faces. *Brain and Cognition, 54*, 248-250. doi: 10.1016/j.bandc.2004.02.029

Grohol, J. M. (2008). 10 common reasons to lie to your therapist. Retrieved 28 September 2015 from http://psychcentral.com/blog/archives/2008/02/06/10-common-reasons-to-lie-to-your-therapist/

Guilford, J. P. (1929). An experiment in learning to read facial expression. *Journal of Abnormal and Social Psychology, 24*, 191-202. doi: 10.1037/h0069973

Hager, J. C., & Ekman, P. (2005). The asymetry of facial actions is inconsistent with models of hemispheric specialization. In P. Ekman & E. L. Rosenberg (Eds.), *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)* (2nd ed., pp. 40-62). New York: Oxford University Press.

Haidt, J., McCauley, C., & Rozin, P. (1994). Individual differences in sensitivity to disgust: A scale sampling seven domains of disgust elicitors. *Personality and Individual Differences, 16*, 701-713. doi: 10.1016/0191-8869(94)90212-7

Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological Bulletin, 85*, 845-857. doi: 10.1037/0033-2909.85.4.845

Hampson, E., van Anders, S. M., & Mullin, L. I. (2006). A female advantage in the recognition of emotional facial expressions: test of an evolutionary hypothesis. *Evolution and Human Behavior, 27*, 401-416. doi: 10.1016/j.evolhumbehav.2006.05.002

Hartgerink, C. H. J., Van Aert, R. C. M., Nuijten, M. B., Wicherts, J. M., & Van Assen, M. A. L. M. (2016). Distributions of p-values smaller than .05 in psychology: What is going on? *PeerJ, 2016*. doi: 10.7717/peerj.1935

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry, 51*, 59-67.

Henley, S. M. D., Novak, M. J. U., Frost, C., King, J., Tabrizi, S. J., & Warren, J. D. (2012). Emotion recognition in Huntington's disease: A systematic review. *Neuroscience and Biobehavioral Reviews, 36*, 237-253. doi: 10.1016/j.neubiorev.2011.06.002

Henry, J. D., Ruffman, T., McDonald, S., O'Leary, M. A. P., Phillips, L. H., Brodaty, H., & Rendell, P. G. (2008). Recognition of disgust is selectively preserved in Alzheimer's disease. *Neuropsychologia, 46*, 1363-1370. doi: 10.1016/j.neuropsychologia.2007.12.012

Herrera, E., Cuetos, F., & Rodríguez-Ferreiro, J. (2011). Emotion recognition impairment in Parkinson's disease patients without dementia. *Journal of the Neurological Sciences, 310*, 237-240. doi: 10.1016/j.jns.2011.06.034

Hess, U., Blairy, S., & Kleck, R. E. (1997). The intensity of emotional facial expressions and decoding accuracy. *Journal of Nonverbal Behavior, 21*, 241-257.

Hess, U., & Thibault, P. (2009). Darwin and emotion expression. *American Psychologist, 64*, 120-128. doi: 10.1037/a0013386

Hoffmann, H., Kessler, H., Eppel, T., Rukavina, S., & Traue, H. C. (2010). Expression intensity, gender and facial emotion recognition: Women recognize only subtle facial emotions better than men. *Acta Psychologica, 135*, 278-283. doi: 10.1016/j.actpsy.2010.07.012

Hudson, S. M., Marshall, W. L., Wales, D., McDonald, E., Bakker, L. W., & McLean, A. (1988). Emotional recognition skills of sex offenders. *Annals of Sex Research, 6*, 199-211. doi: 10.1007/BF00849561

Hutchison, A. N., & Gerstein, L. H. (2012). What's in a face? Counseling trainees' ability to read emotions. *Training and Education in Professional Psychology, 6*, 100-112.

Ickes, W., Gesn, P. R., & Graham, T. (2000). Gender differences in empathic accuracy: Differential ability or differential motivation? *Personal Relationships, 7*, 95-110.

Isaacowitz, D. M., Löckenhoff, C. E., Lane, R. D., Wright, R., Sechrest, L., Riedel, R., & Costa, P. T. (2007). Age differences in recognition of emotion in lexical stimuli and facial expressions. *Psychology and Aging, 22*, 147-159. doi: 10.1037/0882-7974.22.1.147

Izard, C. E. (1992). Basic emotions, relations among emotions, and emotion-cognition relations. *Psychological Review, 99*, 561-565.

Izard, C. E. (1994). Innate and universal facial expressions: Evidence from developmental and cross-cultural research. *Psychological Bulletin, 115*, 288-299.

James, W. (1894). Discussion: The physical basis of emotion. *Psychological Review, 1*, 516-529.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences, 31*, 559-575+612-621. doi: 10.1017/s0140525x08005293

Kanade, T., Cohn, J. F., & Tian, Y. (2000). *Comprehensive database for facial expression analysis.* Paper presented at the 4th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2000, Grenoble.

Kim, Y. T., Kwon, D. H., & Chang, Y. (2011). Impairments of facial emotion recognition and theory of mind in methamphetamine abusers. *Psychiatry Research, 186*, 80-84. doi: 10.1016/j.psychres.2010.06.027

Kirita, T., & Endo, M. (1995). Happy face advantage in recognizing facial expressions. *Acta Psychologica, 89*, 149-163. doi: 10.1016/0001-6918(94)00021-8

Klein, M. H., Mathieu, P. L., Gendlin, E. T., & Kiesler, D. J. (1969). *The Experiencing Scale: A research training manual*. Madison: University of Wisconsin Extension Bureau of Audiovisual Instruction.

Kreutz, G., Ott, U., Teichmann, D., Osawa, P., & Vaitl, D. (2008). Using music to induce emotions: Influences of musical preference and absorption. *Psychology of Music, 36*, 101-126. doi: 10.1177/0305735607082623

Krumhuber, E. G., Kappas, A., & Manstead, A. S. R. (2013). Effects of dynamic aspects of facial expressions: A review. *Emotion Review, 5*, 41-46. doi: 10.1177/1754073912451349

Kunz, M., Peter, J., Huster, S., & Lautenbacher, S. (2013). Pain and disgust: The facial signaling of two aversive bodily experiences. *PLoS ONE, 8*, art. no. e83277. doi: 10.1371/journal.pone.0083277

Landis, C. (1924). Studies of emotional reactions. II. General behavior and facial expression. *Journal of Comparative Psychology, 4*, 447-510. doi: 10.1037/h0073039

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). International Affective Pictures System (IAPS): Technical manual and affective ratings. NIMH Center for the Study of Emotion and Attention.

Lavrakas, P. J., & Maier, R. A. (1979). Differences in human ability to judge veracity from the audio medium. *Journal of Research in Personality, 13*, 139-153. doi: 10.1016/0092-6566(79)90026-6

Lewinski, P., den Uyl, T. M., & Butler, C. (2014). Automated facial coding: Validation of basic emotions and FACS AUs in FaceReader. *Journal of Neuroscience, Psychology, and Economics, 7*, 227-236.

Lints-Martindale, A. C., Hadjistavropoulos, T., Barber, B., & Gibson, S. J. (2007). A psychophysical investigation of the facial action coding system as an index of pain variability among older adults with and without Alzheimer's disease. *Pain Medicine, 8*, 678-689. doi: 10.1111/j.1526-4637.2007.00358.x

Lutz, C., & White, G. (1986). The anthropology of emotions. *Annual Review of Anthropology, 15*, 405-436.

Machado, P. P. P., Beutler, L. E., & Greenberg, L. S. (1999). Emotion recognition in psychotherapy: Impact of therapist level of experience and emotional awareness. *Journal of Clinical Psychology, 55*, 39-57.

Malatesta, C. Z., Izard, C. E., Culver, C., & Nicolich, M. (1987). Emotion communication skills in young, middle-aged, and older women. *Psychology and Aging, 2*, 193-203.

Malcolm, S. R., & Keenan, J. P. (2005). Hemispheric asymmetry and deception detection. *Laterality, 10*, 103-110. doi: 10.1080/13576500442000274

Marsh, P. J., Green, M. J., Russell, T. A., McGuire, J., Harris, A., & Coltheart, M. (2010). Remediation of facial emotion recognition in schizophrenia: Functional predictors, generalizability, and durability. *American Journal of Psychiatric Rehabilitation, 13*, 143-170. doi: 10.1080/15487761003757066

Marsh, P. J., Luckett, G., Russell, T., Coltheart, M., & Green, M. J. (2012). Effects of facial emotion recognition remediation on visual scanning of novel face stimuli. *Schizophrenia Research, 141*, 234-240. doi: 10.1016/j.schres.2012.08.006

Matsumoto, D. (2004). Paul Ekman and the legacy of universals. *Journal of Research in Personality, 38*, 45-51. doi: 10.1016/j.jrp.2003.09.005

Matsumoto, D., Consolacion, T., Yamada, H., & Suzuki, R. (2002). American-Japanese cultural differences in judgements of emotional expressions of different intensities. *Cognition and Emotion, 16*, 721-747. doi: 10.1080/02699930143000608

Matsumoto, D., & Hwang, H. C. (2014). Judgments of subtle facial expressions of emotion. *Emotion, 14*, 349-357. doi: 10.1037/a0035237

Matsumoto, D., & Hwang, H. S. (2011). Evidence for training the ability to read microexpressions of emotion. *Motivation and Emotion, 35*, 181-191. doi: 10.1007/s11031-011-9212-2

Matsumoto, D., Olide, A., Schug, J., Willingham, B., & Callan, M. (2009). Cross-cultural judgments of spontaneous facial expressions of emotion. *Journal of Nonverbal Behavior, 33*, 213-238. doi: 10.1007/s10919-009-0071-4

Matsumoto, D., & Willingham, B. (2009). Spontaneous facial expressions of emotion of congenitally and noncongenitally blind individuals. *Journal of Personality and Social Psychology, 96*, 1-10. doi: 10.1037/a0014037

Matzke, B., Herpertz, S. C., Berger, C., Fleischer, M., & Domes, G. (2014). Facial reactions during emotion recognition in borderline personality disorder: A facial electromyography study. *Psychopathology, 47*, 101-110. doi: 10.1159/000351122

Maurage, P., Grynberg, D., Noël, X., Joassin, F., Hanak, C., Verbanck, P., . . . Philippot, P. (2011). The "Reading the Mind in the Eyes" test as a new way to explore complex emotions decoding in alcohol dependence. *Psychiatry Research, 190*, 375-378. doi: 10.1016/j.psychres.2011.06.015

Mayer, J. D., DiPaolo, M., & Salovey, P. (1990). Perceiving affective content in ambiguous visual stimuli: A component of emotional intelligence. *Journal of Personality Assessment, 54*, 772-781.

Mayer, J. D., Roberts, R. D., & Barsade, S. G. (2008). Human abilities: Emotional intelligence. *Annual Review of Psychology, 59*, 507-536. doi: 10.1146/annurev.psych.59.103006.093646

McNicol, D. M. (1972). *A primer of signal detection theory*. Mahwah, NJ: Lawrence Erlbaum Associates.

Mehu, M., Mortillaro, M., Banziger, T., & Scherer, K. R. (2012). Reliable facial muscle activation enhances recognizability and credibility of emotional expression. *Emotion, 12*, 701-715.

Mikels, J. A., Fredrickson, B. L., Larkin, G. R., Lindberg, C. M., Maglio, S. J., & Reuter-Lorenz, P. A. (2005). Emotional category data on images from the international affective picture system. *Behavior Research Methods, 37*, 626-630.

Murphy, N. A., Lehrfeld, J. M., & Isaacowitz, D. M. (2010). Recognition of posed and spontaneous dynamic smiles in young and older adults. *Psychology and Aging, 25*, 811-821. doi: 10.1037/a0019888

Najt, P., Bayer, U., & Hausmann, M. (2013). Models of hemispheric specialization in facial emotion perception-A reevaluation. *Emotion, 13*, 159-167. doi: 10.1037/a0029723

Niedenthal, P. M., Brauer, M., Halberstadt, J. B., & Innes-Ker, Å. H. (2001). When did her smile drop? Facial mimicry and the influences of emotional state on the detection of change in emotional expression. *Cognition and Emotion, 15*, 853-864. doi: 10.1080/02699930143000194

Noe, R. A. (1986). Trainees' attributes and attitudes: Neglected influences on training effectiveness. *Academy of management review, 11*, 736-749.

O'Sullivan, M., & Ekman, P. (2004). The wizards of deception detection. In P. A. Granhag & L. A. Stromwall (Eds.), *The detection of deception in forensic contexts* (pp. 269-286). Cambridge: Cambridge University Press.

Oatley, K. (2004). *Emotions: a brief history*. Malden, MA: Blackwell Publishing.

Oatley, K., Keltner, D., & Jenkins, J. M. (2006). *Understanding emotions*. Malden, MA: Blackwell Publishing.

Ortony, A., & Turner, T. J. (1990). What's basic about basic emotions? *Psychological Review, 97*, 315-331.

Pete Blair, J., Levine, T. R., & Vasquez, B. E. (2015). Producing deception detection expertise. *Policing, 38*, 71-85. doi: 10.1108/pijpsm-09-2014-0092

Peterson, R. L., & Pennington, B. F. (2012). Developmental dyslexia. *The Lancet, 379*(9830), 1997-2007. doi: 10.1016/S0140-6736(12)60198-6

Phillips, L. H., Henry, J. D., Hosie, J. A., & Milne, A. B. (2008). Effective regulation of the experience and expression of negative affect in old age. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences, 63*, P138-P145.

Porter, S., & Brinke, L. T. (2008). Reading between the lies: Identifying concealed and falsified emotions in universal facial expressions. *Psychological Science, 19*, 508-514.

Porter, S., Campbell, M. A., Stapleton, J., & Birt, A. R. (2002). The influence of judge, target, and stimulus characteristics on the accuracy of detecting deceit. *Canadian Journal of Behavioural Science, 34*, 172-185.

Porter, S., Juodis, M., ten brinke, L. M., Klein, R., & Wilson, K. (2010). Evaluation of the effectiveness of a brief deception detection training program. *Journal of Forensic Psychiatry and Psychology, 21*, 66-76.

Porter, S., ten Brinke, L., & Wallace, B. (2011). Secrets and lies: Involuntary leakage in deceptive facial expressions as a function of emotional intensity. *Journal of Nonverbal Behavior, 36*, 23-37. doi: 10.1046/j.0956-7976.

Porter, S., Woodworth, M., & Birt, A. R. (2000). Truth, lies, and videotape: An investigation of the ability of federal parole officers to detect deception. *Law and Human Behavior, 24*, 643-658.

Pos, A. E., Greenberg, L. S., Goldman, R. N., & Korman, L. M. (2003). Emotional processing during experiential treatment of depression. *Journal of Consulting and Clinical Psychology, 71*, 1007-1016. doi: 10.1037/0022-006X.71.6.1007

Rahman, Q., Wilson, G. D., & Abrahams, S. (2004). Sex, sexual orientation, and identification of positive and negative facial affect. *Brain and Cognition, 54*, 179-185. doi: 10.1016/j.bandc.2004.01.002

Richter, D., Dietzel, C., & Kunzmann, U. (2011). Age differences in emotion recognition: The task matters. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences, 66 B*, 48-55. doi: 10.1093/geronb/gbq068

Rinn, W. E. (1984). The neuropsychology of facial expression: A review of the neurological and psychological mechanisms for producing facial expressions. *Psychological Bulletin, 95*, 52-77. doi: 10.1037/0033-2909.95.1.52

Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting Psychology, 21*, 95-103.

Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. *Neuroscience and Biobehavioral Reviews, 32*, 863-881. doi: 10.1016/j.neubiorev.2008.01.001

Ruffman, T., Murray, J., Halberstadt, J., & Vater, T. (2012). Age-related differences in deception. *Psychology and Aging, 27*, 543-549. doi: 10.1037/a0023380

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin, 115*, 102-141.

Russell, J. A., Bachorowski, J. A., & Fernandez-Dols, J. M. (2003). Facial and vocal expressions of emotion. *Annual Review of Psychology, 54*, 329-349. doi: 10.1146/annurev.psych.54.101601.145102

Russell, T. A., Chu, E., & Phillips, M. L. (2006). A pilot study to investigate the effectiveness of emotion recognition remediation in schizophrenia using the micro-expression training tool. *British Journal of Clinical Psychology, 45*, 579-583.

Russell, T. A., Green, M. J., Simpson, I., & Coltheart, M. (2008). Remediation of facial emotion perception in schizophrenia: Concomitant changes in visual attention. *Schizophrenia Research, 103*, 248-256. doi: 10.1016/j.schres.2008.04.033

Sasson, N. J., Pinkham, A. E., Richard, J., Hughett, P., Gur, R. E., & Gur, R. C. (2010). Controlling for response biases clarifies sex and age differences in facial affect recognition. *Journal of Nonverbal Behavior, 34*, 207-221. doi: 10.1007/s10919-010-0092-z

Sayette, M. A., Cohn, J. F., Wertz, J. M., Perrott, M. A., & Parrott, D. J. (2001). A psychometric evaluation of the facial action coding system for assessing spontaneous expression. *Journal of Nonverbal Behavior, 25*, 167-185. doi: 10.1023/a:1010671109788

Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review, 69*, 379-399. doi: 10.1037/h0046234

Scherer, K. R. (2009). Emotions are emergent processes: They require a dynamic computational architecture. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*, 3459-3474. doi: 10.1098/rstb.2009.0141

Seligman, M. E. P. (2011). *Flourish : a visionary new understanding of happiness and well-being.* New York: Free Press.

Sentencing Act. (2002). Sentencing Act. Retrieved from http://www.legislation.govt.nz/act/public/2002/0009/latest/DLM135342.html

Shaw, J., Porter, S., & ten Brinke, L. (2013). Catching liars: Training mental health and legal professionals to detect high-stakes lies. *Journal of Forensic Psychiatry and Psychology, 24*, 145-159. doi: 10.1007/s10979-006-9016-1;

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*, 1359-1366. doi: 10.1177/0956797611417632

Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology, 30*, 526-537. doi: 10.1037/h0037039

Somerville, L. H., & Whalen, P. J. (2006). Prior experience as a stimulus category confound: an example using facial expressions of emotion. *Social cognitive and affective neuroscience, 1*, 271-274. doi: 10.1093/scan/nsl040

Srinivasan, R., Golomb, J. D., & Martinez, A. M. (2016). A neural basis of facial action recognition in humans. *Journal of Neuroscience, 36*, 4434-4442. doi: 10.1523/JNEUROSCI.1704-15.2016

Sullivan, S., Ruffman, T., & Hutton, S. B. (2007). Age differences in emotion recognition skills and the visual scanning of emotion faces. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences, 62*, P53-P60.

Suzuki, A., Hoshino, T., Shigemasu, K., & Kawamura, M. (2007). Decline or improvement? Age-related differences in facial expression recognition. *Biological Psychology, 74*, 75-84. doi: 10.1016/j.biopsycho.2006.07.003

ten Brinke, L., & Porter, S. (2012). Cry me a river: Identifying the behavioral consequences of extremely high-stakes interpersonal deception. *Law and Human Behavior, 36*, 469-477. doi: 10.1037/h0093929

ten Brinke, L., Porter, S., & Baker, A. (2012). Darwin the detective: Observable facial muscle contractions reveal emotional high-stakes lies. *Evolution and Human Behavior, 33*, 411-416. doi: 10.1007/s10919-011-0120-7;

Thompson, A. E., & Voyer, D. (2014). Sex differences in the ability to recognise non-verbal displays of emotion: A meta-analysis. *Cognition and Emotion, 28*, 1164-1195. doi: 10.1080/02699931.2013.875889

Tomkins, S. (1962). *Affect, imagery, consciousness, Vol. 1: The positive affects*. New York: Springer.

Tomkins, S., & McCarter, R. (1964). What and where are the primary affects? Some evidence for a theory. *Perceptual and Motor Skills, 18*, 119-158.

Tracy, J. L., Randles, D., & Steckler, C. M. (2015). The nonverbal communication of emotions. *Current Opinion in Behavioral Sciences, 3*, 25-30. doi: 10.1016/j.cobeha.2015.01.001

Tziner, A., Fisher, M., Senior, T., & Weisberg, J. (2007). Effects of trainee characteristics on training effectiveness. *International Journal of Selection and Assessment, 15*, 167-174. doi: 10.1111/j.1468-2389.2007.00378.x

Vrij, A., Granhag, P. A., Mann, S., & Leal, S. (2011). Outsmarting the liars: Toward a cognitive lie detection approach. *Current Directions in Psychological Science, 20*, 28-32.

Vrij, A., & Mann, S. (2001a). Telling and detecting lies in a high-stake situation: The case of a convicted murderer. *Applied Cognitive Psychology, 15*, 187-203. doi: 10.1002/1099-0720(200103/04)15:2<187::AID-ACP696>3.0.CO;2-A

Vrij, A., & Mann, S. (2001b). Who killed my relative? Police officers' ability to detect real-life high-stake lies. *Psychology, Crime and Law, 7*, 119-132.

Warren, G., Schertler, E., & Bull, P. (2009). Detecting deception from emotional and unemotional cues. *Journal of Nonverbal Behavior, 33*, 59-69. doi: 10.1007/s10919-008-0057-7

Waxer, P. H. (1983). Emotional deceit: False words versus false action. *Motivation and Emotion, 7*, 365-376. doi: 10.1007/bf00991646

West, J. T., Horning, S. M., Klebe, K. J., Foster, S. M., Cornwell, R. E., Perrett, D., . . . Davis, H. P. (2012). Age effects on emotion recognition in facial displays: From 20 to 89 years of age. *Experimental Aging Research, 38*, 146-168. doi: 10.1080/0361073X.2012.659997

Wong, B., Cronin-Golomb, A., & Neargarder, S. (2005). Patterns of visual scanning as predictors of emotion identification in normal aging. *Neuropsychology, 19*, 739-749. doi: 10.1037/0894-4105.19.6.739

Wood, P. M., & Kroese, B. S. (2007). Enhancing the emotion recognition skills of individuals with learning disabilities: A review of the literature. *Journal of Applied Research in Intellectual Disabilities, 20*, 576-579. doi: 10.1111/j.1468-3148.2006.00355.x

Wright Whelan, C., Wagstaff, G., & Wheatcroft, J. M. (2015). High stakes lies: police and non-police accuracy in detecting deception. *Psychology, Crime and Law, 21*, 127-138. doi: 10.1080/1068316x.2014.935777

Zuckerman, M., Lipets, M. S., Koivumaki, J. H., & Rosenthal, R. (1975). Encoding and decoding nonverbal cues of emotion. *Journal of Personality and Social Psychology, 32*, 1068-1076. doi: 10.1037/0022-3514.32.6.1068

# APPENDICES

## APPENDIX A

## Poster Advertisement

SCHOOL OF PSYCHOLOGY
Te Kura Hinengaro Tangata
Private Bag 11 222
Palmerston North 4442
New Zealand
T 64 6 356 9099 extn 2040
F 64 6 350 5673

Massey University
COLLEGE OF HUMANITIES AND SOCIAL SCIENCES
Te Kura Pūkenga Tangata

# DECEPTION DETECTION

## and Emotion Recognition Abilities of Therapists

Can you

## Lie to Me?

**We need people for the study who believe they are successful deceivers**

Participation in this project will take **half an hour** of your time and the session will be video-recorded.

As a thank you for participating, you will be compensated $15 for your time. ☺

### What does the study involve?

**You will be shown a range of emotion-evoking slides whilst being video-recorded. For part of the study, you will be asked to act in a deceptive manner.**

**If you would like an information sheet, or have any questions, please contact:**

**Alexa Curtis** email: emotion.deception@gmail.com text: 021 401 915
**Ass. Prof. John Podd** School of Psychology, Massey University, Palmerston Nth
Email: J.V.Podd@massey.ac.nz ph: 06 3569099 ext. 2067

"This project has been evaluated by peer review and judged to be low risk. Consequently, it has not been reviewed by one of the University's Human Ethics Committees. The researcher(s) named above are responsible for the ethical conduct of this research.

If you have any concerns about the conduct of this research that you wish to raise with someone other than the researcher(s), please contact Professor John O'Neill, Director, Research Ethics, telephone 06 350 5249, email humanethics@massey.ac.nz".

# APPENDIX B

## Information Sheet

Massey University
COLLEGE OF HUMANITIES AND SOCIAL SCIENCES
Te Kura Pūkenga Tangata

SCHOOL OF PSYCHOLOGY
Te Kura Hinengaro Tangata
Private Bag 11 222
Palmerston North 4442
New Zealand
T 64 6 356 9099 extn 2040
F 64 6 350 5673

# *Deception Detection and Emotion Recognition Abilities of Therapists*

## INFORMATION SHEET

My name is Alexa Curtis and I am a student at Massey University. I am conducting this research project as a requirement for the completion of the Doctoral Programme in Clinical Psychology. My supervisor for this project is Associate Professor John Podd.

**Project Description and Invitation**
This project aims to explore the emotion recognition abilities of Clinical Psychologists and whether they are able to detect deception. The study will also investigate whether these skills can be enhanced by training.

The first part of the study involves creating video images of people responding deceptively to emotionally evocative images. These videos will be shown to the psychologists in the second part of the study.

You are invited to participate in this research.

**Who can participate?**
- Participants will be recruited through advertisements on campus.
- People who are interested can contact the researcher via email or text message.
- Men and women aged between 18 and 45 years will be eligible to participate.
- People who have had facial Botox treatment are ineligible to participate.

**Project Procedures**
- Participants will be shown a range of emotionally evocative slides whilst being video-recorded. Segments of the video will be shown to participants in part 2 of the study. The video images for each participant will be edited and approximately 20 seconds of video per person will be used for part 2 of the study.
- Participants will be recorded individually in a session lasting between 15-30 minutes.

**Participant's Rights**

- You are under no obligation to accept this invitation to complete this task. If you choose to participate, you are free to exit the session at any time. You may ask any questions about the study at any time during participation. The video images will be securely stored and will only be viewed by the researchers and the psychologists participating in part 2 of the study.
- You agree to be video-recorded on the understanding that your name will not be used by the researcher. You can ask for the recorder to be turned off at any time during the session.
- You will be given access to a summary of the project findings when it is concluded.
- You will be reimbursed $15 for your time.
- Please note: some of the images that will be shown may be disturbing to some people. You have the right to withdraw from the study at any time.

**Project Contacts**

**Alexa Curtis** email: emotion.deception@gmail.com text: 021 401 915
**Ass. Prof. John Podd** School of Psychology, Massey University, Palmerston Nth
Email: J.V.Podd@massey.ac.nz ph: 06 3569099 ext. 2067

If you have any further questions please contact the researcher or chief supervisor. If you would like to receive a summary of the research findings following completion, please send an email to the researcher at emotion.deception@gmail.com, and one will be provided to you as soon as it is available.

"This project has been evaluated by peer review and judged to be low risk. Consequently, it has not been reviewed by one of the University's Human Ethics Committees. The researcher(s) named above are responsible for the ethical conduct of this research.

If you have any concerns about the conduct of this research that you wish to raise with someone other than the researcher(s), please contact Professor John O'Neill, Director, Research Ethics, telephone 06 350 5249, email humanethics@massey.ac.nz".

# APPENDIX C

## Participant Consent Form

MASSEY UNIVERSITY
COLLEGE OF HUMANITIES
AND SOCIAL SCIENCES
TE KURA PŪKENGA TANGATA

# *Deception Detection and Emotion Recognition Abilities of Therapists*

## PARTICIPANT CONSENT FORM - INDIVIDUAL

I have read the Information Sheet and have had the details of the study explained to me. My questions have been answered to my satisfaction, and I understand that I may ask further questions at any time.

I agree/do not agree to the interview being image recorded.

I wish/do not wish to have data placed in an official archive.

I agree to participate in this study under the conditions set out in the Information Sheet.

**Signature:** .................................................... **Date:** ....................

**Full Name - printed** ....................................................................

Te Kunenga
ki Pūrehuroa

School of Psychology - Te Kura Hinengaro Tangata
Private Bag 11222, Palmerston North 4442, New Zealand T +64 6 356 9099 extn 2040 F +64 6 350 5673
http://psychology.massey.ac.nz

# APPENDIX D

## Participant Questionnaire

MASSEY UNIVERSITY
COLLEGE OF HUMANITIES
AND SOCIAL SCIENCES
TE KURA PŪKENGA TANGATA

**Deception Detection and Emotion Recognition Abilities of Therapists**

## Participant questionnaire

**If you would like a summary of the completed study, please provide your email address.**

Email address : ______________________________

**Please complete the following to assist with analysis of results**

Sex: M / F Date of Birth: ____________________

**Occupation** ______________________________

**(If student, please state what year of study, e.g. 2nd year)**

**Thank you for participating in this research** ☺

Te Kunenga ki Pūrehuroa

School of Psychology - Te Kura Hinengaro Tangata
Private Bag 11222, Palmerston North 4442, New Zealand T +64 6 356 9099 extn 2040 F +64 6 350 5673
http://psychology.massey.ac.nz

# APPENDIX E

## Instructions for Powerpoint Slideshow

**INSTRUCTIONS FOR POWERPOINT SLIDESHOW**

Thank you for coming today, I appreciate your participation in this study.

We are interested in the deception detection and emotion recognition abilities of therapists.

In order to investigate this, we will be creating videos of both genuine and deceptive facial expressions of emotion.

You will be shown a slideshow of various pictures intended to evoke a range of emotions. I have some prototypes of the universal emotions that you will be asked to express. (Show prototypes to participant.)

While watching the slideshow, an instruction will appear for you to express an emotion. At times the instruction will match the emotion you may be experiencing, at other times the instruction may be for an emotion that is quite different to what you are feeling. Please just do your best to express the particular emotion requested. **You only have to express the emotion for the slide following the instruction, then return to a neutral expression until a further instruction appears.**

I am sure you have had plenty of practise in your life at being deceptive. For example, think about when you've been given a gift that you don't like, how you have to smile and pretend that "it's just what you wanted".

Remember, it is important that in between expressing the requested emotions, that you relax and allow your face to return to a neutral expression.

After watching the slideshow, you will be shown a series of jokes. For each joke I would like you to give it a rating from 0 to 10 where 0 = not funny at all, and 10 = extremely funny.  I will record your ratings of each joke.

# APPENDIX F

## Images used from the International Affective Picture System

**Images used from the International Affective Picture System**

| Picture number | Description | Valence Mean (*SD*) | Arousal Mean (*SD*) | Category (Mikels et al., 2005) |
|---|---|---|---|---|
| 7224 | File cabinets | 4.45 (1.36) | 2.81 (1.94) | |
| 7041 | Baskets | 4.99 (1.12) | 2.60 (1.78) | |
| 7217 | Clothes rack | 4.82 (0.99) | 2.43 (1.64) | |
| 3150 | Mutilation | 2.26 (1.57) | 6.55 (2.20) | Disgust |
| 3000 | Mutilation | 1.45 (1.20) | 7.26 (2.10) | Disgust |
| 5250 | Nature | 6.08 (2.01) | 3.64 (2.27) | |
| 9320 | Vomit | 2.65 (1.92) | 4.93 (2.70) | Disgust |
| 9300 | Dirty | 2.26 (1.76) | 6.00 (2.41) | Disgust |
| 5982 | Sky | 7.61 (1.48) | 4.51 (2.85) | |
| 9042 | StickThruLip | 3.15 (1.89) | 5.78 (2.48) | Disgust |
| 2352.2 | BloodyKiss | 2.09 (1.50) | 6.25 (2.10) | Disgust |
| 7705 | Cabinet | 4.77 (1.02) | 2.65 (1.88) | |
| 2141 | GrievingFem | 2.44 (1.64) | 5.00 (2.03) | Sadness |
| 2205 | Hospital | 1.95 (1.58) | 4.53 (2.23) | Sadness |
| 5711 | Field | 6.62 (1.65) | 3.03 (1.96) | |
| 3350 | Infant | 1.88 (1.67) | 5.72 (2.23) | Sadness |
| 3301 | InjuredChild | 1.80 (1.28) | 5.21 (2.26) | Sadness |
| 5010 | Flower | 7.14 (1.50) | 3.00 (2.25) | |
| 2900 | Crying boy | 2.45 (1.42) | 5.09 (2.15) | Sadness |
| 9252 | Deadbody | 1.98 (1.59) | 6.64 (2.33) | Sadness |
| 7052 | Clothespins | 5.33 (1.32) | 3.01 (2.02) | |
| 1811 | Monkeys | 7.62 (1.59) | 5.12 (2.25) | |
| 1999 | Mickey | 7.43 (1.47) | 4.77 (2.40) | AmCE* |
| 7490 | Window | 5.52 (1.41) | 2.42 (2.23) | |
| 1463 | Kittens | 7.45 (1.46) | 4.79 (2.19) | AmC |

| | | | | |
|---|---|---|---|---|
| 1710 | Puppies | 8.34 (1.12) | 5.41 (2.34) | U |
| 7025 | Stool | 4.63 (1.17) | 2.71 (2.20) | |
| 8031 | Skier | 6.76 (1.39) | 5.58 (2.24) | E |
| 5621 | Skydivers | 7.57 (1.42) | 6.99 (1.95) | AwE |
| 1850 | Camels | 6.15 (1.52) | 4.06 (2.14) | |
| 1920 | Porpoise | 7.90 (1.48) | 4.27 (2.53) | |
| 5910 | Fireworks | 7.80 (1.23) | 5.59 (2.55) | |
| 2070 | Baby | 8.17 (1.46) | 4.51 (2.74) | |
| 8180 | Cliffdivers | 7.12 (1.88) | 6.59 (2.12) | |
| 2791 | Balloons | 6.64 (1.70) | 3.83 (2.09) | |
| 2091 | Girls | 7.68 (1.43) | 4.51 (2.28) | |

*Am = Amusement; C = Contentment; E = Excitement; Aw = Awe; U = Undifferentiated

## APPENDIX G

## Statements by New Zealand Rugby Players

## Statements by New Zealand Rugby Players

**Jono Gibbs**, Chiefs

"Nobody in Rugby should be called a genius. A genius is a guy like Norman Einstein."

**Rodney So'ialo**, Hurricanes, on University

"I'm going to graduate on time, no matter how long it takes."

**Colin Cooper**, Hurricanes head coach

"You guys line up alphabetically by height." And, "You guys pair up in groups of three, then line up in a circle."

**Chris Masoe** (Hurricanes) on whether he had visited the Pyramids during his visit to Egypt .

"I can't really remember the names of the clubs that we went to."

**Colin Cooper** on Paul Tito

"He's a guy who gets up at six o'clock in the morning regardless of what time it is."

**Kevin Senio** ( Auckland ), on Night Rugby vs Day Games

"It's basically the same, just darker."

**David Nucifora** ( Auckland ) talking about Troy Flavell

"I told him, 'Son, what is it with you... Is it ignorance or apathy?' He said, 'David, I don't know and I don't care.'

**David Holwell** (Hurricanes) when asked about the upcoming season:

"I want to reach for 150 or 200 points this season, whichever comes first."

**Ma'a Nonu**

"Colin has done a bit of mental arithmetic with a calculator."

**Phil Waugh**

"We actually got the winning try three minutes from the end but then they scored."

**Jerry Collins**

"I've never had major knee surgery on any other part of my body."

**Tony Brown**

"That kick was absolutely unique, except for the one before it which was identical."

**Tana Umaga**

"I owe a lot to my parents, especially my mother and father."

**Doc Mayhew**

"Sure there have been injuries and deaths in rugby, but none of them serious."

**Anton Oliver**

"If history repeats itself, I should think we can expect the same thing again."

**Ewan McKenzie**

"I never comment on referees and I'm not going to break the habit of a lifetime for that prat."

**Murray Mexted**

(1) "Andy Ellis the 21 year old, who turned 22 a few weeks ago"

(2) "He scored that try after only 22 seconds - totally against the run of play."

(3) "I would not say he (Rico Gear) is the best left winger in the Super14, but there are none better."

(4) "Well, either side could win it, or it could be a draw."

(5) "Strangely, in slow motion replay, the ball seemed to hang in the air for even longer."

**Murray Deaker**:

"Have you ever thought of writing your autobiography?"

**Tana Umaga**:

"On what?"

## APPENDIX H

## Invitation to Psychologists to Participate in the Study

Dear Clinical Psychologist

I am extending you an invitation to participate in an exciting research study, part of which will enable you to undertake some professional development training.

My name is Alexa Curtis, I am undertaking a doctorate in clinical psychology at Massey University. I am supervised by Associate Professor John Podd. Dr Shane Harvey is responsible for the clinical aspects of the study.

For my doctoral thesis I am interested in a crucial component of Emotional Intelligence – the recognition of emotion from facial expressions. I will be investigating this skill in clinical psychologists and whether training enhances this ability. I am also interested in whether training increases the ability to detect deception in facial expressions.

**What is involved?**

A maximum of two hours of your time is required during which time you would complete a questionnaire on emotional awareness and a deception detection and emotion recognition task (pre- and post-training).

You would also be provided with training in reading subtle expressions of emotion (Subtle Expression Training Tool) developed by Paul Ekman through the Ekman Institute. This would take 1.5 hours and is provided at no cost to you.

I would visit you at your worksite, or a location convenient to you, at a time during late August or September that suits you. If you wish, I could also provide you with your pre- and post-training scores. You would have continued access to your on-line training scores.

Following the study, you would also have unlimited access to the on-line training.

This study is limited to 40 participants. Participants will be randomly assigned to either training or no-training. However, participants in the no-training group will be given the opportunity to complete the training. Further information is outlined in the attached information sheet.

If you are interested in participating, please email me at:

**emotion.deception@gmail.com**

Thank you for taking the time to consider this invitation.

Regards

Alexa Curtis
*Doctoral Candidate*
*Massey University*
*Palmerston North*

Te Kunenga ki Pūrehuroa

**School of Psychology - Te Kura Hinengaro Tangata**
Private Bag 11222, Palmerston North 4442, New Zealand **T** +64 6 356 9099 extn 2040 **F** +64 6 350 5673
**http://psychology.massey.ac.nz**

## APPENDIX I

## Information Sheet

Massey University
COLLEGE OF HUMANITIES AND SOCIAL SCIENCES
Te Kura Pūkenga Tangata

SCHOOL OF PSYCHOLOGY
Te Kura Hinengaro Tangata
Private Bag 11 222
Palmerston North 4442
New Zealand
T 64 6 356 9099 extn 2040
F 64 6 350 5673

# *Deception Detection and Emotion Recognition Abilities of Therapists*

**INFORMATION SHEET**

## Who is doing this research?

My name is Alexa Curtis and I am a student at Massey University. I am conducting this research project as a requirement for the completion of a Doctorate in Clinical Psychology. My chief supervisor is Associate Professor John Podd. Clinical Psychologist, Dr Shane Harvey, is responsible for the clinical aspects of the study.

## Project Description and Invitation

This project explores the emotion recognition and deception detection abilities of Clinical Psychologists. The study will also investigate whether these skills can be enhanced by training.

You are invited to participate in this research.

## Who can participate?

Clinical Psychologists who are currently in practice and have not previously completed any training in facial emotion recognition.

## What is involved?

Participants will be randomly assigned to either a training or no-training condition. All participants will view video clips of facial expressions and make judgements of veracity, and select an emotion term to match the expression. The training group will complete an on-line training programme in subtle expressions (Subtle Expression Training Tool) through the Ekman Institute (http:/www.paulekman.com), after which all participants will again rate the genuine/deceptive facial expression video clips.

Participants assigned to the non-training group will be able to return to their regular work schedule for 1.5 hours between the two emotion recognition tasks. However, no face-to-face client work can be conducted during this time. Participants in the no-training group will be provided with access to the on-line training at the completion of the second emotion recognition task, and can undertake this training at their own convenience.

Participants will also complete a 55-item questionnaire on emotional practice. This will be sent to you before our meeting and can be completed prior to the training session. The questionnaire will be re-administered one month following training. You will be asked whether you consent to this re-administration, and if so, asked to provide contact details. The information collected from this questionnaire, along with the pre-training data, will be used to inform an Honours Project conducted by Amanda Johnsen, on emotional competencies of therapists. This project is supervised by Dr Shane Harvey and Dr Don Baken.

## Participant's Rights

If you decide to participate, you have the right to:

- decline to answer any particular question;
- withdraw from the study at any time;
- ask any questions about the study at any time during participation;
- provide information on the understanding that your name will not be used unless you give permission to the researcher;
- be given access to a summary of the project findings when it is concluded.

## Project Contacts

| *MAIN PROJECT* | |
|---|---|
| **Researcher** | **Chief Supervisor** |
| **Alexa Curtis** | **Associate Professor John Podd** |
| School of Psychology<br>Massey University<br>Turitea Campus<br>Palmerston North | School of Psychology<br>Massey University<br>Turitea Campus<br>Palmerston North |
| Ph: 021 401 915<br>emotion.deception@gmail.com | Ph: 06 3569099 ext. 2067<br>J.V.Podd@massey.ac.nz |
| *HONOURS PROJECT* | |
| **Researcher** | **Chief Supervisor** |
| **Amanda Johnsen** | **Dr. Shane Harvey** |
| School of Psychology<br>Massey University<br>Turitea Campus<br>Palmerston North | School of Psychology<br>Massey University<br>Turitea Campus<br>Palmerston North |
| amanda.johnsen.1@uni.massey.ac.nz | Ph. 06 3569099 ext. 81742<br>S.T.Harvey@massey.ac.nz |

If you have any further questions please contact the researcher(s) or chief supervisor(s). If you would like to receive a summary of the research findings following completion, please send an email to the researcher at emotion.deception@gmail.com, and one will be provided to you as soon as it is available.

"This project has been evaluated by peer review and judged to be low risk. Consequently, it has not been reviewed by one of the University's Human Ethics Committees. The researcher(s) named above are responsible for the ethical conduct of this research.

If you have any concerns about the conduct of this research that you wish to raise with someone other than the researcher(s), please contact Professor John O'Neill, Director, Research Ethics, telephone 06 350 5249, email humanethics@massey.ac.nz".

## APPENDIX J

## Participant Consent Form

Massey University
COLLEGE OF HUMANITIES AND SOCIAL SCIENCES
Te Kura Pūkenga Tangata

SCHOOL OF PSYCHOLOGY
Te Kura Hinengaro Tangata
Private Bag 11 222
Palmerston North 4442
New Zealand
T 64 6 356 9099 extn 2040
F 64 6 350 5673

# *Deception Detection and Emotion Recognition Abilities of Therapists*

## PARTICIPANT CONSENT FORM - INDIVIDUAL

I have read the Information Sheet and have had the details of the study explained to me. My questions have been answered to my satisfaction, and I understand that I may ask further questions at any time.

I wish/do not wish to have data placed in an official archive.

I agree to participate in this study under the conditions set out in the Information Sheet.

**Signature:** .................................................................... **Date:** ....................

**Full Name - printed** ....................................................................

**APPENDIX K(A)**

**Instructions for Training Group Participants**

*Deception Detection and Emotion Recognition Abilities of Therapists*

**Instructions**

**Please read carefully through the entire set of instructions before commencing.**

**Once you have read the instructions, please email me the date/time you have set aside to complete the data collection. I will then phone you prior to the study to answer any questions you may have.**

**There are four parts to the data collection:**

| | |
|---|---|
| **Part 1.** | **Complete the Questionnaire and Consent Form** |
| **Part 2.** | **Deception Detection and Emotion Recognition Task – Trial 1** |
| **Part 3.** | **On-line training for participants assigned to the training group**<br><br>**Administrative tasks for participants assigned to the no-training group** |
| **Part 4.** | **Deception Detection and Emotion Recognition Task – Trial 2** |

**<u>It is important you do not preview the CD-ROM prior to Part 2.</u>**

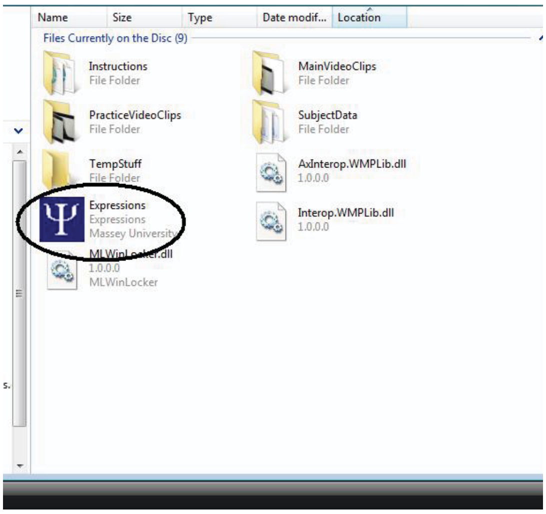
**Part 1. Questionnaire and Consent Form**

These may be completed prior to the data collection session.

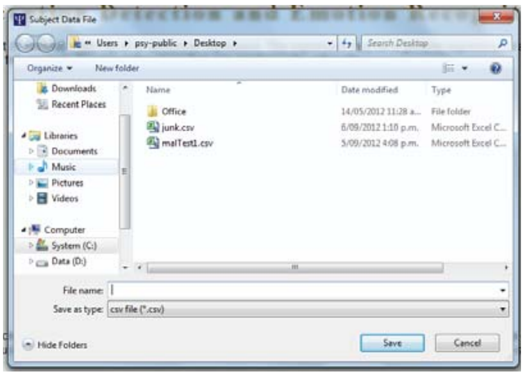
**Part 2 – Part 4 must be completed within the two-hour timeframe and without any interruption.**

**Part 2. CD-ROM - Deception Detection and Emotion Recognition Task – Trial 1**

Step 1. Insert the CD-ROM into the CD drive on your computer. If your computer is set to auto-run, click “open folder to view files” (otherwise navigate to the DVD/CD drive). Double-click on the “Expressions” file folder, a screen similar to the following screen will appear:

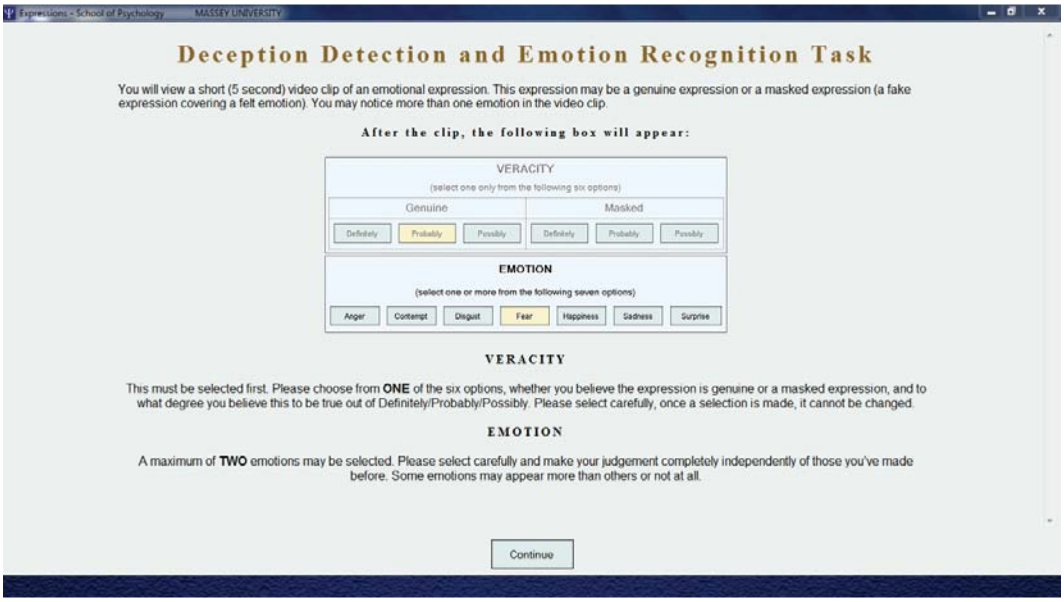



Double-click on the Expressions Psychology Icon and the next screen will appear.



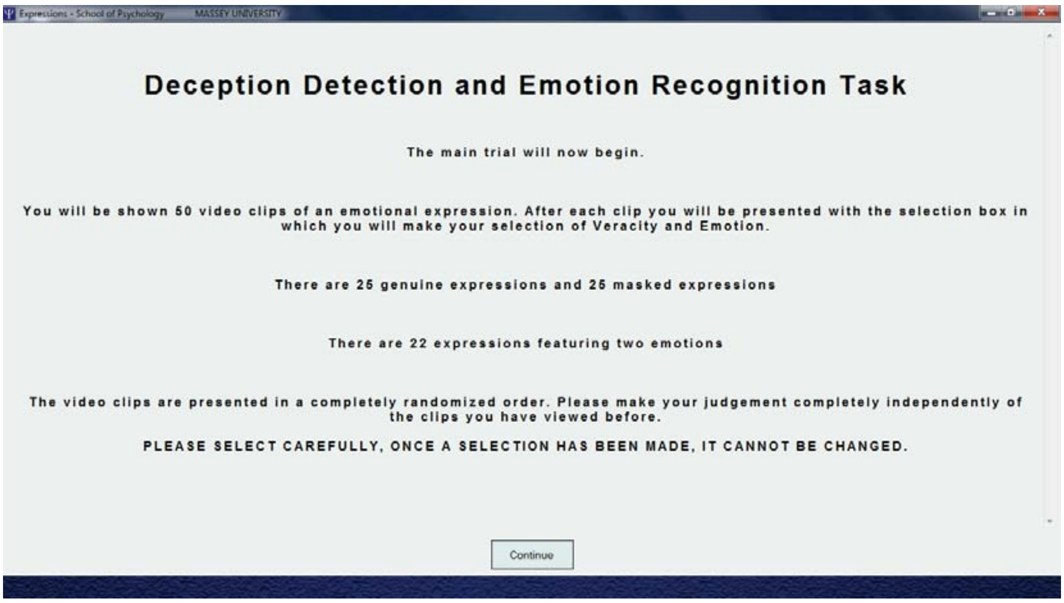

Step 2. The programme will ask you to enter a file name. Type in P4a and click ‘save’. This will save the file to your desktop.

The following screen will appear:




Please read ALL the instructions on this screen before clicking 'continue'. You will then complete a practice trial containing three video clips. Following each video clip, you will be asked to select whether you believe the expression is genuine or masked, and what emotion(s) you believe is/are expressed. A maximum of two emotions may be selected.

After the practice trial, the following screen will appear:



Step 3. You will now view 50 video clips, similar to those you viewed in the practice trial. For each clip, you will be asked to select the veracity of the expression and emotion(s) you believe is/are being expressed.

There are 25 genuine expressions and 25 masked expressions

There are 22 expressions featuring two emotions

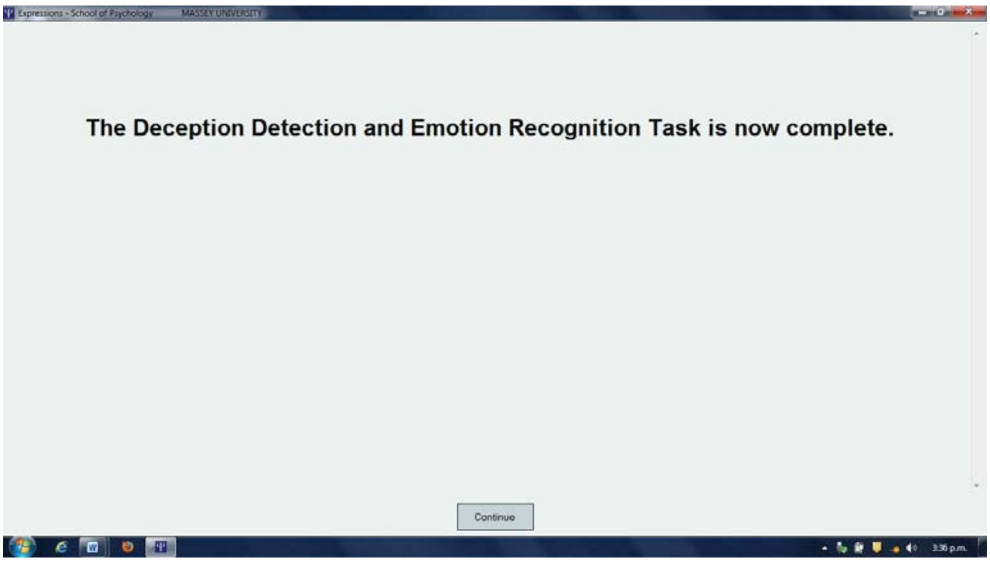
The video clips are presented in a completely randomized order. Please make your judgement completely independently of the clips you have viewed before.

Some emotions appear more than others or not at all.

Please select carefully, once a selection has been made, it cannot be changed.

Step 4. Click 'continue' to start the main trial.

When you have viewed all 50 video clips, the following screen will appear:




Step 5. Click 'continue' and the programme will close.

You have been randomly assigned to the training group.

**Part 3. Instructions for training group.**

Please email me at emotion.deception@gmail.com and advise me that you have completed Part 2. CD-ROM - Deception Detection and Emotion Recognition Task – Trial 1. I will then email you instructions to access the on-line Subtle Expressions Training Tool (SETT).

**SETT TRAINING**

Once you have logged on and accessed the training, read the 'Introduction' and 'About' sections.

Step 1. Please complete the 'Learn' session for all seven emotions
– Fear, Surprise, Happy, Angry, Contempt, Sad, and Disgust.

For each expression, you will see several examples of the expression together with a description of the expression.

Step 2. After completing the 'Learn' session, complete the 'Practice' session. At the end of the Practice session, you will receive a score. Make a note of this score.

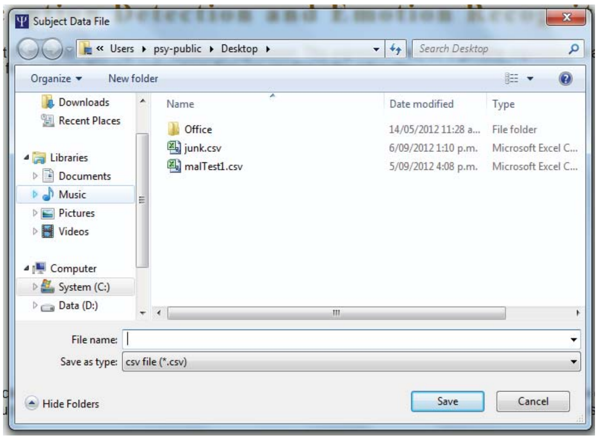
Repeat Step 1 and Step 2 above, making a note of your score for the second 'Practice' session.

Review any of the seven emotional expressions that you wish to.

**This takes approximately 90 minutes to go through the 'Learn' and 'Practice' sessions. When you have finished, please email me your two scores.**

**Part 4. Deception Detection and Emotion Recognition Task – Trial 2**

Repeat all the steps above for Part 2, except for in Step 2, when requested to enter a file name, this time type in P4b. This will save a second file to your desktop.




When the trial is complete, locate the two files saved to your desktop. These will be named P4a.csv and P4b.csv and look similar to the one shown below


Insert these two data files into an email and return them via email to emotion.deception@gmail.com

Return the CD-ROM in the post-paid envelope, together with the signed consent form, questionnaire and these instructions.

**This now completes the data collection.**

**Thank you for participating in this study, your time and input is greatly appreciated. I hope you found it an interesting and enjoyable experience.**

**Remember you now have ongoing access to the SETT training.**

## APPENDIX K(B)

## Instructions for No Training Group Participants

*Deception Detection and Emotion Recognition Abilities of Therapists*

**Instructions**

**Please read carefully through the entire set of instructions before commencing.**

**Once you have read the instructions, please email me the date/time you have set aside to complete the data collection. I will then phone you prior to the study to answer any questions you may have.**

**There are four parts to the data collection:**

| | |
|---|---|
| **Part 1.** | **Complete the Questionnaire and Consent Form** |
| **Part 2.** | **Deception Detection and Emotion Recognition Task – Trial 1** |
| **Part 3.** | **On-line training for participants assigned to the training group**<br><br>**Administrative tasks for participants assigned to the no-training group** |
| **Part 4.** | **Deception Detection and Emotion Recognition Task – Trial 2** |

**It is important you do not preview the CD-ROM prior to Part 2.**

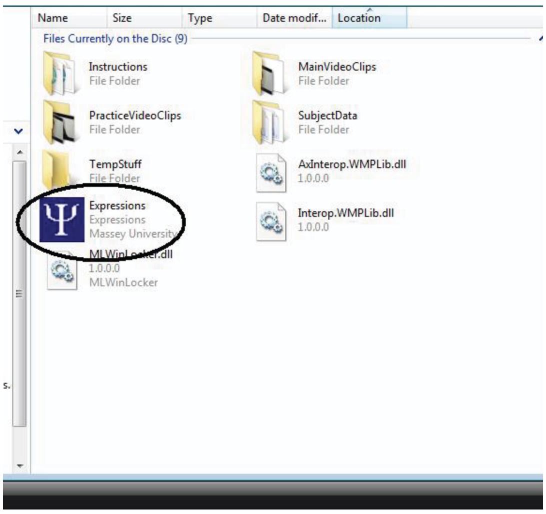
**Part 1. Questionnaire and Consent Form**

These may be completed prior to the data collection session.

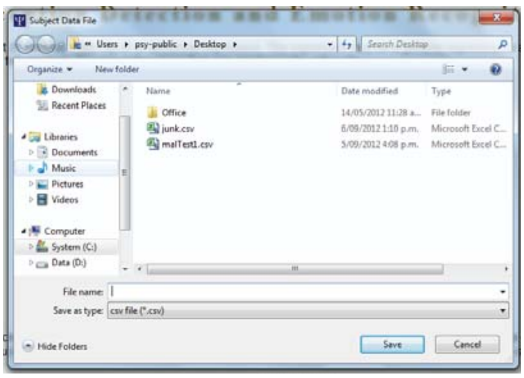
**Part 2 – Part 4 must be completed within the two-hour timeframe and without any interruption.**

**Part 2. CD-ROM - Deception Detection and Emotion Recognition Task – Trial 1**

Step 1. Insert the CD-ROM into the CD drive on your computer. If your computer is set to auto-run, click “open folder to view files” (otherwise navigate to the DVD/CD drive). Double-click on the “Expressions” file folder, a screen similar to the following screen will appear:

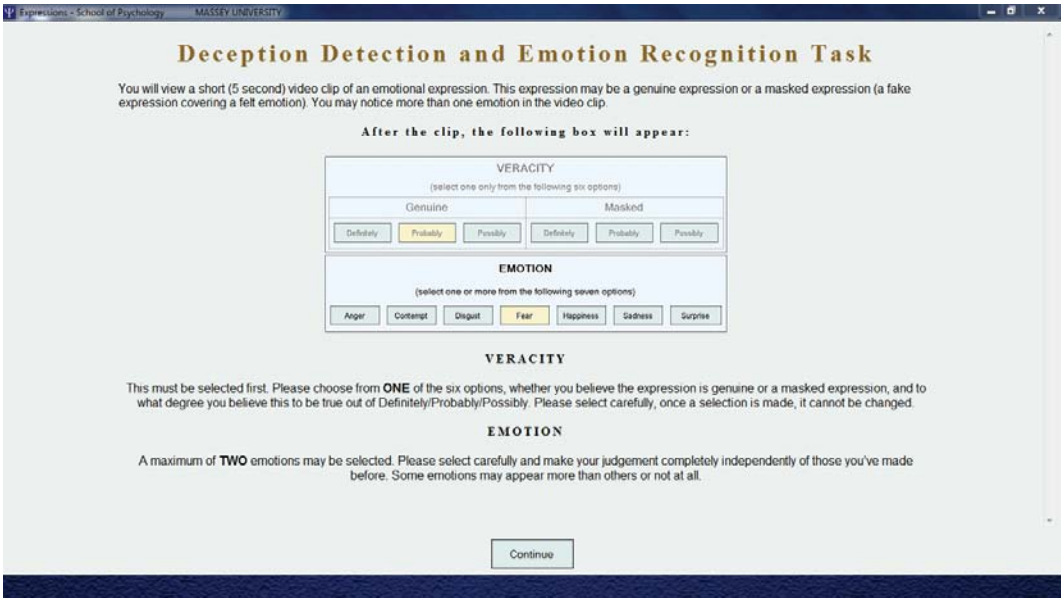



Double-click on the Expressions Psychology Icon and the next screen will appear.



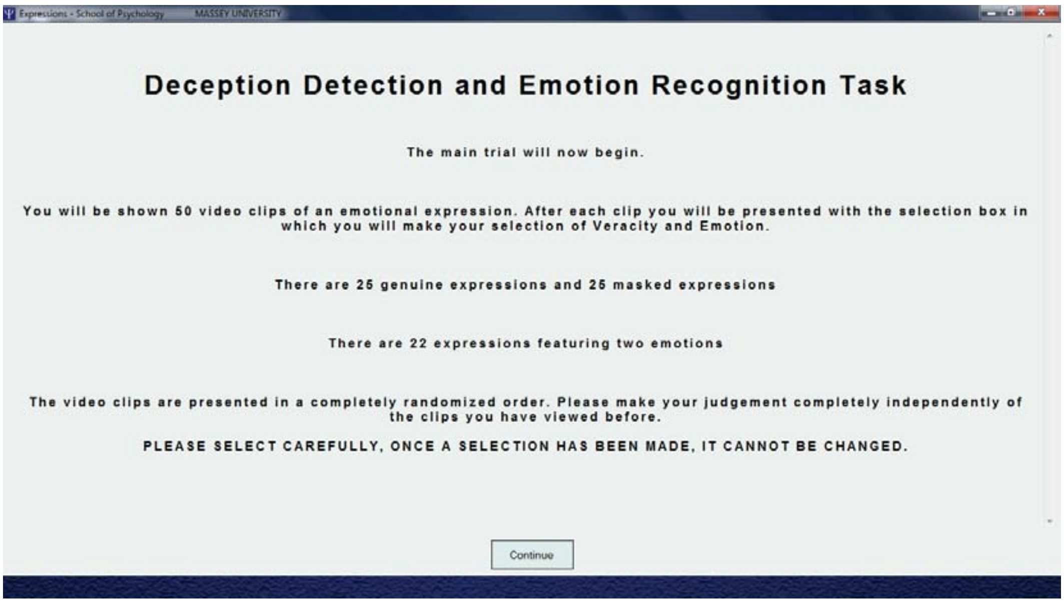

Step 2. The programme will ask you to enter a file name. Type in P1a and click ‘save’. This will save the file to your desktop.

The following screen will appear:




Please read ALL the instructions on this screen before clicking ‘continue’. You will then complete a practice trial containing three video clips. Following each video clip, you will be asked to select whether you believe the expression is genuine or masked, and what emotion(s) you believe is/are expressed. A maximum of two emotions may be selected.

After the practice trial, the following screen will appear:



Step 3. You will now view 50 video clips, similar to those you viewed in the practice trial. For each clip, you will be asked to select the veracity of the expression and emotion(s) you believe is/are being expressed.

There are 25 genuine expressions and 25 masked expressions

There are 22 expressions featuring two emotions

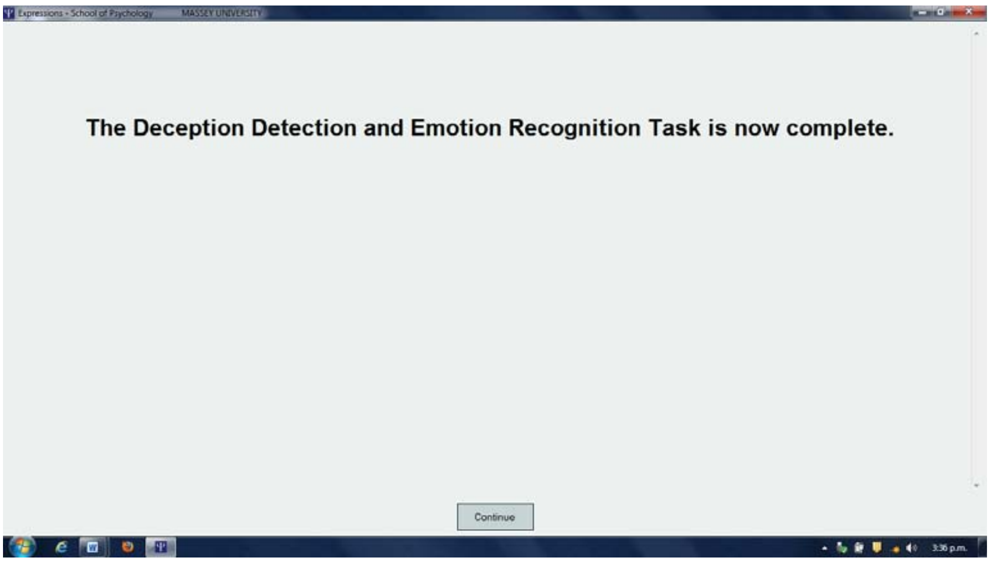
The video clips are presented in a completely randomized order. Please make your judgement completely independently of the clips you have viewed before.

Some emotions appear more than others or not at all.

Please select carefully, once a selection has been made, it cannot be changed.

Step 4. Click ‘continue’ to start the main trial.

When you have viewed all 50 video clips, the following screen will appear:




Step 5. Click ‘continue’ and the programme will close.

You have been randomly assigned to the no-training group.

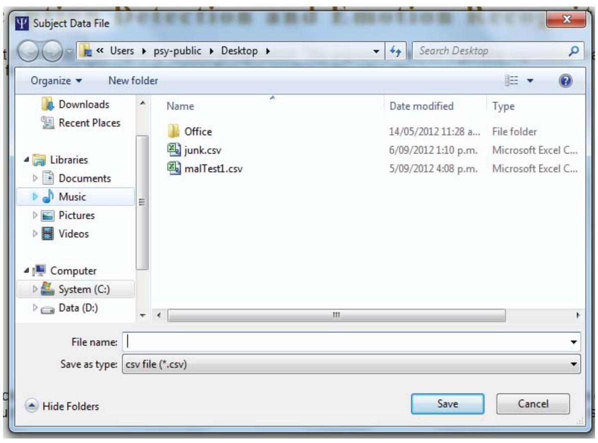
Part 3. Instructions for no-training group.

For the next 1.5 hours you can continue business as usual, as long as there is no face-to-face contact with clients. Please complete the following log detailing the type of work carried out during the 1.5 hour interval. Eg. Casenotes, wrote reports, etc.

| Type of work |
|---|
| |
| |
| |
| |
| |
| |

**Part 4. CD-ROM - Deception Detection and Emotion Recognition Task – Trial 2**

Repeat all the steps above for Part 2, except for in Step 2, when requested to enter a file name, this time type in P1b.  This will save a second file to your desktop.



When the trial is complete, locate the two files saved to your desktop. These will be named P1a.csv and P1b.csv and look similar to the one shown below


Insert these two data files into an email and return them via email to emotion.deception@gmail.com

Return the CD-ROM in the post-paid envelope, together with the signed consent form, questionnaire and these instructions.

Once I have received the two data files, I will email you your logon and password to access the Subtle Emotion Training Tool (SETT) from the Ekman Institute.

**This now completes the data collection.**

**Thank you for participating in this study, your time and input is greatly appreciated. I hope you found it an interesting and enjoyable experience.**

# APPENDIX L

## Ethics Approval

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA

30 September 2011

Alexa Curtis
33 Riveredge Terrace
RD 20
**LEVIN 5570**

Dear Alexa

**Re: Clinical Psychologists' Ability to Interpret Emotion and Detect Deception from Facial Expressions**

Thank you for your Low Risk Notification which was received on 29 September 2011.

Your project has been recorded on the Low Risk Database which is reported in the Annual Report of the Massey University Human Ethics Committees.

The low risk notification for this project is valid for a maximum of three years.

Please notify me if situations subsequently occur which cause you to reconsider your initial ethical analysis that it is safe to proceed without approval by one of the University's Human Ethics Committees.

Please note that travel undertaken by students must be approved by the supervisor and the relevant Pro Vice-Chancellor and be in accordance with the Policy and Procedures for Course-Related Student Travel Overseas. In addition, the supervisor must advise the University's Insurance Officer.

**A reminder to include the following statement on all public documents:**

> *"This project has been evaluated by peer review and judged to be low risk. Consequently, it has not been reviewed by one of the University's Human Ethics Committees. The researcher(s) named above are responsible for the ethical conduct of this research.*
>
> *If you have any concerns about the conduct of this research that you wish to raise with someone other than the researcher(s), please contact Professor John O'Neill, Director (Research Ethics), telephone 06 350 5249, e-mail humanethics@massey.ac.nz".*

Please note that if a sponsoring organisation, funding authority or a journal in which you wish to publish requires evidence of committee approval (with an approval number), you will have to provide a full application to one of the University's Human Ethics Committees. You should also note that such an approval can only be provided prior to the commencement of the research.

Yours sincerely

J. O'Neill

John G O'Neill (Professor)
**Chair, Human Ethics Chairs' Committee and**
**Director (Research Ethics)**

cc Assoc Prof John Podd
School of Psychology
**PN320**

Assoc Prof Mandy Morgan, HoS
School of Psychology
**PN320**

Massey University Human Ethics Committee
Accredited by the Health Research Council

**Research Ethics Office,** Massey University, Private Bag 11222, Palmerston North 4442, New Zealand
**T** +64 6 350 5573 +64 6 350 5575 **F** +64 6 350 5622
**E** humanethics@massey.ac.nz animalethics@massey.ac.nz gtc@massey.ac.nz
www.massey.ac.nz

**MASSEY UNIVERSITY**
**ALBANY**

2 July 2015

Alexa Curtis
33 Riveredge Terrace
RD20
**Levin**

Dear Alexa

**Re: Clinical Psychologists' ability to interpret emotion and detect deception from facial expressions**

Thank you for your Low Risk Notification which was received on 2 July 2015.

Your project has been recorded on the Low Risk Database which is reported in the Annual Report of the Massey University Human Ethics Committees.

You are reminded that staff researchers and supervisors are fully responsible for ensuring that the information in the low risk notification has met the requirements and guidelines for submission of a low risk notification.

The low risk notification for this project is valid for a maximum of three years.

Please notify me if situations subsequently occur which cause you to reconsider your initial ethical analysis that it is safe to proceed without approval by one of the University's Human Ethics Committees.

Please note that travel undertaken by students must be approved by the supervisor and the relevant Pro Vice-Chancellor and be in accordance with the Policy and Procedures for Course-Related Student Travel Overseas. In addition, the supervisor must advise the University's Insurance Officer.

**A reminder to include the following statement on all public documents:**

> *"This project has been evaluated by peer review and judged to be low risk. Consequently, it has not been reviewed by one of the University's Human Ethics Committees. The researcher(s) named above are responsible for the ethical conduct of this research.*
>
> *If you have any concerns about the conduct of this research that you wish to raise with someone other than the researcher(s), please contact Dr Brian Finch, Director (Research Ethics), telephone 06 356 9099, extn 86015, e-mail humanethics@massey.ac.nz".*

Please note that if a sponsoring organisation, funding authority or a journal in which you wish to publish requires evidence of committee approval (with an approval number), you will have to provide a full application to one of the University's Human Ethics Committees. You should also note that such an approval can only be provided prior to the commencement of the research.

Yours sincerely

Brian T Finch (Dr)
**Chair, Human Ethics Chairs' Committee and**
**Director (Research Ethics)**

cc Associate Professor John Podd
School of Psychology
**Palmerston North**

Professor James Liu
Head of School of Psychology
**Albany campus**

**Massey University Human Ethics Committee**
**Accredited by the Health Research Council**

## APPENDIX M

## ANOVA tables – Study 1

**ANOVA table for Single Happy Expressions**

**Test of Within-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Time | 0.02 | 1 | 0.02 | 3.53 | 0.07 | 0.08 | 0.45 |
| Time*Training | 0.00 | 1 | 0.00 | 0.85 | 0.36 | 0.02 | 0.15 |
| Time* Experience Level | 0.00 | 1 | 0.00 | 0.27 | 0.61 | 0.01 | 0.08 |
| Time*Training*Experience Level | 0.02 | 1 | 0.02 | 3.60 | 0.07 | 0.09 | 0.46 |
| Error (Time) | 0.19 | 39 | 0.01 | | | | |

**Test of Between-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Training | 0.01 | 1 | 0.01 | 0.40 | 0.53 | 0.01 | 0.10 |
| Experience Level | 0.05 | 1 | 0.05 | 2.67 | 0.11 | 0.06 | 0.36 |
| Training*Experience Level | 0.02 | 1 | 0.02 | 1.41 | 0.24 | 0.04 | 0.21 |
| Error | 0.68 | 39 | 0.02 | | | | |

**ANOVA table for Single Sad Expressions**

**Test of Within-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Time | 0.42 | 1 | 0.42 | 27.84 | 0.00 | 0.42 | 1.00 |
| Time*Training | 0.03 | 1 | 0.03 | 1.93 | 0.17 | 0.05 | 0.27 |
| Time* Experience Level | 0.00 | 1 | 0.00 | 0.04 | 0.85 | 0.00 | 0.05 |
| Time*Training*Experience Level | 0.00 | 1 | 0.00 | 0.05 | 0.83 | 0.00 | 0.06 |
| Error (Time) | 0.60 | 39 | 0.02 | | | | |

**Test of Between-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Training | 0.41 | 1 | 0.41 | 5.95 | 0.02 | 0.13 | 0.66 |
| Experience Level | 0.05 | 1 | 0.05 | 0.73 | 0.40 | 0.02 | 0.13 |
| Training*Experience Level | 0.18 | 1 | 0.18 | 2.66 | 0.11 | 0.06 | 0.36 |
| Error | 2.68 | 39 | 0.07 | | | | |

**ANOVA table for Single Disgust Expressions**

**Test of Within-Subjects Effects**

| **Source of Variation** | **Sum of Squares** | ***df*** | **Mean Squares** | ***F*** | **Sig.** | **Partial Eta Squared** | **Power** |
|---|---|---|---|---|---|---|---|
| Time | 0.00 | 1 | 0.00 | 0.08 | 0.78 | 0.00 | 0.06 |
| Time*Training | 0.01 | 1 | 0.01 | 0.58 | 0.45 | 0.02 | 0.12 |
| Time* Experience Level | 0.00 | 1 | 0.00 | 0.01 | 0.93 | 0.00 | 0.05 |
| Time*Training*Experience Level | 0.06 | 1 | 0.06 | 3.25 | .08 | 0.08 | 0.42 |
| Error (Time) | 0.73 | 39 | 0.02 | | | | |

**Test of Between-Subjects Effects**

| **Source of Variation** | **Sum of Squares** | ***df*** | **Mean Squares** | ***F*** | **Sig.** | **Partial Eta Squared** | **Power** |
|---|---|---|---|---|---|---|---|
| Training | 0.01 | 1 | 0.01 | 0.35 | 0.56 | 0.01 | 0.09 |
| Experience Level | 0.06 | 1 | 0.06 | 1.36 | 0.25 | 0.03 | 0.21 |
| Training*Experience Level | 0.02 | 1 | 0.02 | 0.53 | 0.47 | 0.01 | 0.11 |
| Error | 1.59 | 39 | 0.04 | | | | |

**ANOVA tables for Masked Happy Expressions**

**Test of Within-Subjects Effects**

| **Source of Variation** | **Sum of Squares** | ***df*** | **Mean Squares** | ***F*** | **Sig.** | **Partial Eta Squared** | **Power** |
|---|---|---|---|---|---|---|---|
| Time | 0.10 | 1 | 0.10 | 2.19 | 0.15 | 0.05 | 0.30 |
| Time*Training | 0.05 | 1 | 0.05 | 1.08 | 0.31 | 0.03 | 0.17 |
| Time* Experience Level | 3.76 | 1 | 3.76 | 0.00 | 0.98 | 0.00 | 0.05 |
| Time*Training*Experience Level | 0.01 | 1 | 0.01 | 0.17 | 0.68 | 0.00 | 0.07 |
| Error (Time) | 1.84 | 39 | 0.05 | | | | |

**Test of Between-Subjects Effects**

| **Source of Variation** | **Sum of Squares** | ***df*** | **Mean Squares** | ***F*** | **Sig.** | **Partial Eta Squared** | **Power** |
|---|---|---|---|---|---|---|---|
| Training | 0.00 | 1 | 0.00 | 0.16 | 0.90 | 0.00 | 0.05 |
| Experience Level | 0.00 | 1 | 0.00 | 0.16 | 0.90 | 0.00 | 0.05 |
| Training*Experience Level | 0.22 | 1 | 0.22 | 3.75 | 0.06 | 0.09 | 0.47 |
| Error | 2.33 | 39 | 0.06 | | | | |

## ANOVA tables for Masked Sad Expressions

**Test of Within-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Time | 0.18 | 1 | 0.18 | 4.77 | 0.04 | 0.11 | 0.57 |
| Time*Training | 0.01 | 1 | 0.01 | 0.15 | 0.70 | 0.00 | 0.07 |
| Time* Experience Level | 0.00 | 1 | 0.00 | 0.02 | 0.88 | 0.00 | 0.05 |
| Time*Training*Experience Level | 0.06 | 1 | 0.06 | 1.43 | 0.24 | 0.04 | 0.21 |
| Error (Time) | 1.50 | 39 | 0.04 | | | | |

**Test of Between-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Training | 0.28 | 1 | 0.28 | 3.08 | 0.09 | 0.07 | 0.40 |
| Experience Level | 0.04 | 1 | 0.04 | 0.43 | 0.51 | 0.01 | 0.10 |
| Training*Experience Level | 0.18 | 1 | 0.18 | 1.99 | 0.16 | 0.05 | 0.28 |
| Error | 3.58 | 39 | 0.09 | | | | |

## ANOVA tables for Masked Disgust Expressions

**Test of Within-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Time | 0.00 | 1 | 0.00 | 0.10 | 0.75 | 0.00 | 0.06 |
| Time*Training | 0.01 | 1 | 0.01 | 0.21 | 0.65 | 0.01 | 0.07 |
| Time* Experience Level | 0.00 | 1 | 0.00 | 0.08 | 0.79 | 0.00 | 0.06 |
| Time*Training*Experience Level | 0.02 | 1 | 0.02 | 1.10 | 0.30 | 0.03 | 0.18 |
| Error (Time) | 0.87 | 39 | 0.02 | | | | |

**Test of Between-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Training | 0.11 | 1 | 0.11 | 2.76 | 0.10 | 0.07 | 0.37 |
| Experience Level | 0.50 | 1 | 0.50 | 12.42 | 0.00 | 0.24 | 0.93 |
| Training*Experience Level | 0.44 | 1 | 0.44 | 10.93 | 0.00 | 0.22 | 0.90 |
| Error | 1.58 | 39 | 0.04 | | | | |

**ANOVA tables for Leaked Happy Expressions**

**Test of Within-Subjects Effects**

| **Source of Variation** | **Sum of Squares** | ***df*** | **Mean Squares** | ***F*** | **Sig.** | **Partial Eta Squared** | **Power** |
|---|---|---|---|---|---|---|---|
| Time | 0.00 | 1 | 0.00 | 0.89 | 0.35 | 0.02 | 0.15 |
| Time*Training | 9.55 | 1 | 9.55 | 0.02 | 0.89 | 0.00 | 0.05 |
| Time* Experience Level | 0.00 | 1 | 0.00 | 1.09 | 0.30 | 0.03 | 0.18 |
| Time*Training*Experience Level | 0.00 | 1 | 0.00 | 0.83 | 0.37 | 0.02 | 0.14 |
| Error (Time) | 0.19 | 39 | 0.01 | | | | |

**Test of Between-Subjects Effects**

| **Source of Variation** | **Sum of Squares** | ***df*** | **Mean Squares** | ***F*** | **Sig.** | **Partial Eta Squared** | **Power** |
|---|---|---|---|---|---|---|---|
| Training | 0.00 | 1 | 0.00 | 0.05 | 0.82 | 0.00 | 0.06 |
| Experience Level | 0.00 | 1 | 0.00 | 0.06 | 0.81 | 0.00 | 0.06 |
| Training*Experience Level | 5.62 | 1 | 5.62 | 0.01 | 0.94 | 0.00 | 0.05 |
| Error | 0.37 | 39 | 0.01 | | | | |

**ANOVA tables for Leaked Sad Expressions**

**Test of Within-Subjects Effects**

| **Source of Variation** | **Sum of Squares** | ***df*** | **Mean Squares** | ***F*** | **Sig.** | **Partial Eta Squared** | **Power** |
|---|---|---|---|---|---|---|---|
| Time | 9.86 | 1 | 9.86 | 0.01 | 0.94 | 0.00 | 0.05 |
| Time*Training | 0.01 | 1 | 0.01 | 0.62 | 0.44 | 0.02 | 0.12 |
| Time* Experience Level | 0.01 | 1 | 0.01 | 0.57 | 0.45 | 0.01 | 0.11 |
| Time*Training*Experience Level | 0.02 | 1 | 0.02 | 0.92 | 0.34 | 0.02 | 0.16 |
| Error (Time) | 0.62 | 39 | 0.02 | | | | |

**Test of Between-Subjects Effects**

| **Source of Variation** | **Sum of Squares** | ***df*** | **Mean Squares** | ***F*** | **Sig.** | **Partial Eta Squared** | **Power** |
|---|---|---|---|---|---|---|---|
| Training | 0.04 | 1 | 0.04 | 2.78 | 0.10 | 0.07 | 0.37 |
| Experience Level | 0.00 | 1 | 0.00 | 0.22 | 0.32 | 0.01 | 0.07 |
| Training*Experience Level | 0.01 | 1 | 0.01 | 0.43 | 0.52 | 0.01 | 0.10 |
| Error | 0.62 | 39 | 0.02 | | | | |

**ANOVA tables for Leaked Disgust Expressions**

**Test of Within-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Time | 0.07 | 1 | 0.07 | 0.77 | 0.39 | 0.02 | 0.14 |
| Time*Training | 0.00 | 1 | 0.00 | 0.02 | 0.89 | 0.00 | 0.05 |
| Time* Experience Level | 0.00 | 1 | 0.00 | 0.02 | 0.89 | 0.00 | 0.05 |
| Time*Training*Experience Level | 0.00 | 1 | 0.00 | 0.01 | 0.91 | 0.00 | 0.05 |
| Error (Time) | 3.55 | 39 | 0.09 | | | | |

**Test of Between-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Training | 0.07 | 1 | 0.07 | 0.89 | 0.54 | 0.01 | 0.09 |
| Experience Level | 0.00 | 1 | 0.00 | 0.01 | 0.94 | 0.00 | 0.05 |
| Training*Experience Level | 0.00 | 1 | 0.00 | 0.01 | 0.92 | 0.00 | 0.05 |
| Error | 7.05 | 39 | 0.18 | | | | |

**ANOVA tables for Sensitivity *P(A)***

**Test of Within-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Time | 0.03 | 1 | 0.00 | 0.30 | 0.59 | 0.01 | 0.08 |
| Time*Training | 0.00 | 1 | 0.00 | 0.06 | 0.81 | 0.00 | 0.06 |
| Time* Experience Level | 0.01 | 1 | 0.01 | 2.01 | 0.17 | 0.05 | 0.28 |
| Time*Training*Experience Level | 4.69 | 1 | 4.69 | 0.01 | 0.92 | 0.00 | 0.05 |
| Error (Time) | 0.17 | 39 | 0.00 | | | | |

**Test of Between-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Training | 0.04 | 1 | 0.04 | 3.73 | 0.06 | 0.09 | 0.47 |
| Experience Level | 0.04 | 1 | 0.04 | 3.91 | 0.06 | 0.09 | 0.49 |
| Training*Experience Level | 1.72 | 1 | 1.72 | 0.00 | 0.97 | 0.00 | 0.05 |
| Error | 0.37 | 39 | 0.01 | | | | |

## ANOVA tables for Hit Rate (HR)

**Test of Within-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Time | 0.00 | 1 | 0.00 | 0.07 | 0.80 | 0.00 | 0.06 |
| Time*Training | 0.00 | 1 | 0.00 | 0.31 | 0.58 | 0.01 | 0.08 |
| Time* Experience Level | 0.01 | 1 | 0.01 | 0.60 | 0.44 | 0.02 | 0.12 |
| Time*Training*Experience Level | 0.00 | 1 | 0.00 | 0.04 | 0.85 | 0.00 | 0.05 |
| Error (Time) | 0.50 | 39 | 0.01 | | | | |

**Test of Between-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Training | 0.06 | 1 | 0.06 | 1.99 | 0.17 | 0.05 | 0.28 |
| Experience Level | 0.03 | 1 | 0.03 | 1.07 | 0.31 | 0.03 | 0.17 |
| Training*Experience Level | 0.00 | 1 | 0.00 | 0.04 | 0.85 | 0.00 | 0.05 |
| Error | 1.22 | 39 | 0.03 | | | | |

## ANOVA tables for False Alarm Rate (FAR)

**Test of Within-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Time | 0.01 | 1 | 0.01 | 0.61 | 0.44 | 0.12 | 0.12 |
| Time*Training | 0.00 | 1 | 0.00 | 0.20 | 0.66 | 0.01 | 0.07 |
| Time* Experience Level | 0.00 | 1 | 0.00 | 0.03 | 0.86 | 0.00 | 0.05 |
| Time*Training*Experience Level | 0.00 | 1 | 0.00 | 0.27 | 0.61 | 0.01 | 0.08 |
| Error (Time) | 0.36 | 39 | 0.01 | | | | |

**Test of Between-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Training | 0.11 | 1 | 0.11 | 4.42 | 0.04 | 0.10 | 0.54 |
| Experience Level | 0.00 | 1 | 0.00 | 0.09 | 0.76 | 0.00 | 0.06 |
| Training*Experience Level | 0.00 | 1 | 0.00 | 0.12 | 0.73 | 0.00 | 0.06 |
| Error | 0.99 | 39 | 0.03 | | | | |

**ANOVA tables for Bias (*c*)**

**Test of Within-Subjects Effects**

| **Source of Variation** | **Sum of Squares** | ***df*** | **Mean Squares** | ***F*** | **Sig.** | **Partial Eta Squared** | **Power** |
|---|---|---|---|---|---|---|---|
| Time | 0.00 | 1 | 0.00 | 0.04 | 0.85 | 0.00 | 0.05 |
| Time*Training | 0.00 | 1 | 0.00 | 0.01 | 0.92 | 0.00 | 0.05 |
| Time* Experience Level | 0.01 | 1 | 0.01 | 0.15 | 0.70 | 0.00 | 0.07 |
| Time*Training*Experience Level | 0.02 | 1 | 0.02 | 0.26 | 0.61 | 0.01 | 0.08 |
| Error (Time) | 2.89 | 39 | 0.07 | | | | |

**Test of Between-Subjects Effects**

| **Source of Variation** | **Sum of Squares** | ***df*** | **Mean Squares** | ***F*** | **Sig.** | **Partial Eta Squared** | **Power** |
|---|---|---|---|---|---|---|---|
| Training | 0.73 | 1 | 0.73 | 3.82 | 0.06 | 0.09 | 0.48 |
| Experience Level | 0.03 | 1 | 0.03 | 0.17 | 0.68 | 0.00 | 0.07 |
| Training*Experience Level | 0.01 | 1 | 0.01 | 0.03 | 0.87 | 0.00 | 0.05 |
| Error | 7.42 | 39 | 0.19 | | | | |

**APPENDIX N**

**ANOVA tables – Study 2**

**ANOVA table for Single Happy Expressions**

**Test of Within-Subjects Effects**

| **Source of Variation** | **Sum of Squares** | ***df*** | **Mean Squares** | ***F*** | **Sig.** | **Partial Eta Squared** | **Power** |
|---|---|---|---|---|---|---|---|
| Time | 0.00 | 1 | 0.00 | 0.34 | 0.57 | 0.02 | 0.09 |
| Time*Training | 0.00 | 1 | 0.00 | 0.10 | 0.75 | 0.01 | 0.06 |
| Time* Experience Level | 0.01 | 1 | 0.01 | 1.93 | 0.18 | 0.08 | 0.26 |
| Time*Training*Experience Level | 0.00 | 1 | 0.00 | 0.18 | 0.68 | 0.01 | 0.07 |
| Error (Time) | 0.12 | 21 | 0.01 | | | | |

**Test of Between-Subjects Effects**

| **Source of Variation** | **Sum of Squares** | ***df*** | **Mean Squares** | ***F*** | **Sig.** | **Partial Eta Squared** | **Power** |
|---|---|---|---|---|---|---|---|
| Training | 4.69 | 1 | 4.69 | 0.01 | 0.94 | 0.00 | 0.05 |
| Experience Level | 0.00 | 1 | 0.00 | 0.04 | 0.85 | 0.00 | 0.05 |
| Training*Experience Level | 0.00 | 1 | 0.00 | 0.25 | 0.62 | 0.01 | 0.08 |
| Error | 0.15 | 21 | 0.01 | | | | |

**ANOVA table for Single Sad Expressions**

**Test of Within-Subjects Effects**

| **Source of Variation** | **Sum of Squares** | ***df*** | **Mean Squares** | ***F*** | **Sig.** | **Partial Eta Squared** | **Power** |
|---|---|---|---|---|---|---|---|
| Time | 0.10 | 1 | 0.10 | 3.69 | 0.07 | 0.15 | 0.45 |
| Time*Training | 0.18 | 1 | 0.18 | 6.87 | 0.02 | 0.25 | 0.71 |
| Time* Experience Level | 0.03 | 1 | 0.03 | 1.14 | 0.30 | 0.05 | 0.18 |
| Time*Training*Experience Level | 0.00 | 1 | 0.00 | 0.11 | 0.74 | 0.01 | 0.06 |
| Error (Time) | 0.54 | 21 | 0.03 | | | | |

**Test of Between-Subjects Effects**

| **Source of Variation** | **Sum of Squares** | ***df*** | **Mean Squares** | ***F*** | **Sig.** | **Partial Eta Squared** | **Power** |
|---|---|---|---|---|---|---|---|
| Training | 0.02 | 1 | 0.02 | 0.26 | 0.61 | 0.01 | 0.08 |
| Experience Level | 0.00 | 1 | 0.00 | 0.01 | 0.95 | 0.00 | 0.05 |
| Training*Experience Level | 0.01 | 1 | 0.01 | 0.06 | 0.81 | 0.00 | 0.06 |
| Error | 1.55 | 21 | 0.07 | | | | |

## ANOVA table for Single Disgust Expressions

**Test of Within-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Time | 0.22 | 1 | 0.22 | 15.99 | 0.00 | 0.43 | 0.97 |
| Time*Training | 0.00 | 1 | 0.00 | 0.06 | 0.81 | 0.00 | 0.06 |
| Time* Experience Level | 0.02 | 1 | 0.02 | 1.56 | 0.23 | 0.07 | 0.22 |
| Time*Training*Experience Level | 0.01 | 1 | 0.01 | 0.47 | 0.50 | 0.02 | 0.10 |
| Error (Time) | 0.29 | 21 | 0.01 | | | | |

**Test of Between-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Training | 0.00 | 1 | 0.00 | 0.00 | 0.95 | 0.00 | 0.05 |
| Experience Level | 0.03 | 1 | 0.03 | 1.10 | 0.31 | 0.05 | 0.17 |
| Training*Experience Level | 0.07 | 1 | 0.07 | 2.56 | 0.13 | 0.11 | 0.33 |
| Error | 0.58 | 21 | 0.03 | | | | |

## ANOVA tables for Masked Happy Expressions

**Test of Within-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Time | 0.05 | 1 | 0.05 | 1.69 | 0.21 | 0.07 | 0.24 |
| Time*Training | 0.01 | 1 | 0.01 | 0.32 | 0.58 | 0.02 | 0.08 |
| Time* Experience Level | 0.01 | 1 | 0.01 | 0.53 | 0.48 | 0.02 | 0.11 |
| Time*Training*Experience Level | 0.00 | 1 | 0.00 | 0.00 | 0.99 | 0.00 | 0.05 |
| Error (Time) | 0.56 | 21 | 0.03 | | | | |

**Test of Between-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Training | 0.08 | 1 | 0.08 | 1.51 | 0.23 | 0.07 | 0.22 |
| Experience Level | 0.13 | 1 | 0.13 | 2.52 | 0.13 | 0.11 | 0.33 |
| Training*Experience Level | 2.51 | 1 | 2.51 | 0.00 | 0.98 | 0.00 | 0.05 |
| Error | 1.06 | 21 | 0.05 | | | | |

## ANOVA tables for Masked Sad Expressions

**Test of Within-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Time | 0.30 | 1 | 0.30 | 6.97 | 0.02 | 0.25 | 0.71 |
| Time*Training | 0.01 | 1 | 0.01 | 0.22 | 0.65 | 0.01 | 0.07 |
| Time* Experience Level | 0.19 | 1 | 0.19 | 4.54 | 0.05 | 0.18 | 0.53 |
| Time*Training*Experience Level | 0.11 | 1 | 0.11 | 2.54 | 0.13 | 0.11 | 0.33 |
| Error (Time) | 0.89 | 21 | 0.04 | | | | |

**Test of Between-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Training | 0.12 | 1 | 0.12 | 2.12 | 0.16 | 0.09 | 0.29 |
| Experience Level | 0.00 | 1 | 0.00 | 0.02 | 0.89 | 0.00 | 0.05 |
| Training*Experience Level | 0.01 | 1 | 0.01 | 0.11 | 0.74 | 0.01 | 0.06 |
| Error | 1.23 | 21 | 0.06 | | | | |

## ANOVA tables for Masked Disgust Expressions

**Test of Within-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Time | 0.00 | 1 | 0.00 | 0.03 | 0.86 | 0.00 | 0.05 |
| Time*Training | 0.00 | 1 | 0.00 | 0.10 | 0.76 | 0.01 | 0.06 |
| Time* Experience Level | 9.42 | 1 | 9.42 | 0.00 | 0.95 | 0.00 | 0.05 |
| Time*Training*Experience Level | 0.00 | 1 | 0.00 | 0.02 | 0.90 | 0.00 | 0.05 |
| Error (Time) | 0.45 | 21 | 0.02 | | | | |

**Test of Between-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Training | 0.00 | 1 | 0.00 | 0.00 | 0.97 | 0.00 | 0.05 |
| Experience Level | 0.01 | 1 | 0.01 | 0.09 | 0.77 | 0.00 | 0.06 |
| Training*Experience Level | 0.10 | 1 | 0.10 | 1.41 | 0.25 | 0.06 | 0.21 |
| Error | 1.43 | 21 | 0.07 | | | | |

## ANOVA tables for Leaked Happy Expressions

**Test of Within-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Time | 0.00 | 1 | 0.00 | 0.65 | 0.43 | 0.03 | 0.12 |
| Time*Training | 0.01 | 1 | 0.01 | 2.08 | 0.16 | 0.09 | 0.28 |
| Time* Experience Level | 0.11 | 1 | 0.11 | 18.42 | 0.00 | 0.47 | 0.98 |
| Time*Training*Experience Level | 0.00 | 1 | 0.00 | 0.23 | 0.64 | 0.01 | 0.07 |
| Error (Time) | 0.12 | 21 | 0.01 | | | | |

**Test of Between-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Training | 0.01 | 1 | 0.01 | 0.26 | 0.61 | 0.01 | 0.08 |
| Experience Level | 0.00 | 1 | 0.00 | 0.03 | 0.87 | 0.00 | 0.05 |
| Training*Experience Level | 0.01 | 1 | 0.01 | 0.24 | 0.63 | 0.01 | 0.08 |
| Error | 0.65 | 21 | 0.03 | | | | |

## ANOVA tables for Leaked Sad Expressions

**Test of Within-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Time | 0.04 | 1 | 0.04 | 1.41 | 0.25 | 0.06 | 0.21 |
| Time*Training | 0.13 | 1 | 0.13 | 4.66 | 0.04 | 0.18 | 0.54 |
| Time* Experience Level | 0.13 | 1 | 0.13 | 4.67 | 0.04 | 0.18 | 0.54 |
| Time*Training*Experience Level | 0.04 | 1 | 0.04 | 1.35 | 0.26 | 0.06 | 0.20 |
| Error (Time) | 0.59 | 21 | 0.03 | | | | |

**Test of Between-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Training | 0.01 | 1 | 0.01 | 0.54 | 0.47 | 0.03 | 0.11 |
| Experience Level | 0.05 | 1 | 0.05 | 1.84 | 0.19 | 0.08 | 0.25 |
| Training*Experience Level | 5.26 | 1 | 5.26 | 0.00 | 0.99 | 0.00 | 0.05 |
| Error | 0.52 | 21 | 0.03 | | | | |

## ANOVA tables for Leaked Disgust Expressions

**Test of Within-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Time | 0.28 | 1 | 0.28 | 4.06 | 0.06 | 0.16 | 0.49 |
| Time*Training | 0.09 | 1 | 0.09 | 1.25 | 0.28 | 0.06 | 0.19 |
| Time* Experience Level | 0.14 | 1 | 0.14 | 2.08 | 0.16 | 0.09 | 0.28 |
| Time*Training*Experience Level | 0.05 | 1 | 0.05 | 0.79 | 0.38 | 0.04 | 0.14 |
| Error (Time) | 1.43 | 21 | 0.07 | | | | |

**Test of Between-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Training | 0.22 | 1 | 0.22 | 1.56 | 0.23 | 0.07 | 0.22 |
| Experience Level | 0.00 | 1 | 0.00 | 0.01 | 0.91 | 0.00 | 0.05 |
| Training*Experience Level | 0.08 | 1 | 0.08 | 0.58 | 0.45 | 0.03 | 0.11 |
| Error | 2.92 | 21 | 0.14 | | | | |

## ANOVA tables for Sensitivity *P(A)*

**Test of Within-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Time | 0.00 | 1 | 0.00 | 0.94 | 0.34 | 0.04 | 0.15 |
| Time*Training | 0.00 | 1 | 0.00 | 0.96 | 0.34 | 0.04 | 0.16 |
| Time* Experience Level | 0.02 | 1 | 0.02 | 6.07 | 0.02 | 0.22 | 0.65 |
| Time*Training*Experience Level | 0.00 | 1 | 0.00 | 1.12 | 0.30 | 0.05 | 0.17 |
| Error (Time) | 0.07 | 21 | 0.00 | | | | |

**Test of Between-Subjects Effects**

| Source of Variation | Sum of Squares | *df* | Mean Squares | *F* | Sig. | Partial Eta Squared | Power |
|---|---|---|---|---|---|---|---|
| Training | 0.02 | 1 | 0.02 | 1.54 | 0.23 | 0.07 | 0.22 |
| Experience Level | 0.01 | 1 | 0.01 | 0.83 | 0.37 | 0.04 | 0.14 |
| Training*Experience Level | 0.02 | 1 | 0.02 | 1.47 | 0.24 | 0.07 | 0.21 |
| Error | 0.30 | 21 | 0.01 | | | | |

**ANOVA tables for Bias (*c*)**

**Test of Within-Subjects Effects**

| **Source of Variation** | **Sum of Squares** | ***df*** | **Mean Squares** | ***F*** | **Sig.** | **Partial Eta Squared** | **Power** |
|---|---|---|---|---|---|---|---|
| Time | 0.00 | 1 | 0.00 | 0.03 | 0.87 | 0.00 | 0.05 |
| Time*Training | 0.23 | 1 | 0.23 | 3.21 | 0.09 | 0.13 | 0.40 |
| Time* Experience Level | 0.02 | 1 | 0.02 | 0.29 | 0.59 | 0.01 | 0.08 |
| Time*Training*Experience Level | 0.00 | 1 | 0.00 | 0.04 | 0.85 | 0.00 | 0.05 |
| Error (Time) | 1.48 | 21 | 0.07 | | | | |

**Test of Between-Subjects Effects**

| **Source of Variation** | **Sum of Squares** | ***df*** | **Mean Squares** | ***F*** | **Sig.** | **Partial Eta Squared** | **Power** |
|---|---|---|---|---|---|---|---|
| Training | 0.02 | 1 | 0.02 | 0.10 | 0.75 | 0.01 | 0.06 |
| Experience Level | 0.04 | 1 | 0.04 | 0.24 | 0.63 | 0.01 | 0.08 |
| Training*Experience Level | 0.25 | 1 | 0.25 | 1.66 | 0.21 | 0.07 | 0.23 |
| Error | 3.11 | 21 | 0.15 | | | | |