The influence of imagination, connectivity, and social context on the assessment and measurement of empathic accuracy using photographic stimuli

Tracey Woolrych
BPsych (Hons)

This thesis is presented in fulfilment of the requirements of the Degree of Doctor of Philosophy

2014

School of Law
Murdoch University
Western Australia
Declaration

I declare that this thesis is my own account of my research and contains as its main content work which has not previously been submitted for a degree at any tertiary education institution.

Tracey Woolrych

December 2014
Abstract

The ability to accurately interpret the emotions of others is known as empathic accuracy, and in this thesis is referred to as Affect Recognition-Empathic Accuracy (AR-EA). This ability can facilitate pro-social behaviours while deficits may result in anti-social behaviours. Research has demonstrated that imagination, connectivity, and social context can all influence our ability to accurately interpret the emotions of others; however, there has been little research investigating how these specific factors might be enhanced, or influence AR-EA abilities when using photographic stimuli. There were two aims to this thesis. The first aim was to investigate the possibility of inserting specific empathy related elements, imagination, connectivity, and social context, into a set of photographic stimuli to assess the potential influence on AR-EA. The second aim was to develop an original set of photographic stimuli for use in this thesis, and to conduct psychometric evaluations on said photographs in order to develop a new photographic measure for the assessment and evaluation of AR-EA. The photographs consisted of both male and female models expressing six different basic emotions (happy, sad, fear, anger, surprise, disgust) at three different levels of intensity (low, medium and high intensity), plus one neutral expression. Imagination and connectivity were both facilitated through the insertion of a silhouette (blacked out full body figure, male or female) into the photographic stimuli. Social context was manipulated through the use of different social setting backgrounds in the photographs: a kitchen, a bar (as in a tavern), and a neutral background. Results demonstrated the silhouette inserted into the photographs to facilitate imagination and connectivity not only enhanced empathic processes, but also produced photographic-based measure of AR-EA that was superior in both reliability and validity to other presentation modes (full body only, and head and shoulders only stimuli). The different social settings of the photographs also impacted AR-EA abilities facilitating the accurate interpretation of some emotions, whilst inhibiting others. The overall findings of this
thesis question past research methods as well as provide intriguing insights into the functioning of empathic accuracy processes which have not been previously reported. The testing and research also resulted in a new photographic measure for the assessment of AR-EA abilities, whilst the use of simple techniques to manipulate empathy-based elements within the photographs offers new opportunities for future research.
# Table of Contents

DECLARATION ................................................................................................................................. ii

ABSTRACT ......................................................................................................................................... iii

TABLE OF CONTENTS ....................................................................................................................... iv

LIST OF TABLES ................................................................................................................................. xii

LIST OF FIGURES ............................................................................................................................... xiii

ACKNOWLEDGEMENTS ..................................................................................................................... xv

CHAPTER 1 – INTRODUCTION ........................................................................................................ 1

1.1 Overview ...................................................................................................................................... 1

1.2 What is Empathy? ......................................................................................................................... 5

1.2.1 A brief history ......................................................................................................................... 5

1.2.2 Empathy – more than just sympathy ..................................................................................... 6

1.2.3 Theories of empathy ............................................................................................................... 7

1.2.3.1 The cognitive approach .................................................................................................... 7

1.2.3.2 The affective approach ..................................................................................................... 8

1.2.3.3 Bringing the two approaches together .............................................................................. 9

1.3 Davis’ Multi-dimensional Model of Empathy ............................................................................. 10

1.4 Defining Empathy in the Context of this Thesis ......................................................................... 12

1.4.1 Empathy and Behaviour ........................................................................................................ 14

1.4.1.1 Empathy and antisocial behaviours .............................................................................. 15

1.4.1.2 Empathy and offending .................................................................................................. 17

1.4.1.3 How empathy can increase antisocial behavior .............................................................. 18

1.4.1.4 Empathy: biases, misinterpretations and attributions ................................................... 20

1.5 Measuring Empathic Accuracy .................................................................................................. 25

1.5.1 Interpersonal interactions – Icke’s approach ....................................................................... 26
1.5.2 Movies, cartoons, vignettes, and stories ......................................................... 29
1.5.3 Photographs and images .................................................................................. 32
1.5.4 Summary of previous methodologies .............................................................. 38
1.6 Developing Original Photographic Stimuli .......................................................... 38
  1.6.1 Head and shoulders only, body only, or full body? ........................................ 39
  1.6.2 Task difficulty ................................................................................................. 41
  1.6.2.1 Facial expression variability as a function of display rules/norms .......... 43
  1.6.3 Which emotions to capture ............................................................................... 46
  1.6.3.1 Posed versus spontaneous emotional expressions ................................. 49
  1.6.3.2 Elicitation techniques and guidelines ....................................................... 50
1.7 Manipulation of Photographic Elements ............................................................. 52
  1.7.1 Enhancing imagination and connectivity ....................................................... 52
  1.7.2 The influence of social context on empathic accuracy............................... 54
  1.7.2.1 Background cues ....................................................................................... 54
  1.7.2.2 Social context and cognitive processes ................................................... 57
1.8 Study Aims and Chapter Outlines ...................................................................... 59

CHAPTER 2 – THE PHOTOGRAPHS: .................................................................... 64
  2.1 Overview ........................................................................................................... 64
  2.2 Composition of the Photographs ..................................................................... 65
    2.2.1 Presentation modes: full body versus head and shoulders only .............. 65
    2.2.2 Universal and Basic Emotions ................................................................. 66
    2.2.3 Emotional Expressions .............................................................................. 68
      2.2.3.1 Posed versus natural/spontaneous expressions ............................... 68
      2.2.3.2 Eliciting emotional expressions for recording ............................... 69
      2.2.3.3 Varying expression intensity levels .................................................. 72
2.2.3.4 Accuracy criterion ................................................................. 74
2.3 Taking the Photographs – Method ........................................... 74
2.3.1 Accuracy Criterion Outcomes of photographic sessions .......... 75
2.4 Testing the Photographs ................................................................. 77
2.4.1 Study 1: Open-ended responding ........................................... 77
  2.4.1.1 Method ........................................................................... 78
    2.4.1.1.1 Participants ............................................................... 78
    2.4.1.1.2 Materials ................................................................. 79
    2.4.1.1.3 Procedure ............................................................... 79
    2.4.1.1.4 Analysis .................................................................. 80
  2.4.1.2 Results ............................................................................ 80
  2.4.1.3 Discussion ...................................................................... 84
2.4.2 Study 2: Forced-choice sorting task ....................................... 87
  2.4.2.1 Method ........................................................................... 89
    2.4.2.1.1 Participants ............................................................... 89
    2.4.2.1.2 Materials ................................................................. 90
    2.4.2.1.3 Procedure ............................................................... 90
    2.4.2.1.4 Analysis .................................................................. 91
  2.4.2.2 Results ............................................................................ 91
    2.4.2.2.1 Emotion category (EC) and emotion intensity
      (EI) scores ................................................................. 91
    2.4.2.2.2 Overall accuracy percentages by model, emotion, and
      photograph ................................................................. 92
    2.4.2.2.3 Model sex .................................................................. 95
    2.4.2.2.3 Participant sex .......................................................... 97
2.4.2.3 Discussion – Study 2 ................................................................. 97

2.4.3 Discussion – Study 1 and Study 2 Comparison .......................... 100

2.5 Selecting the Photographs for Inclusion in the Empathic Accuracy Measure ................................................................. 101

2.5.1 Rationale and selection .............................................................. 101

CHAPTER 3 – IMAGINATION AND CONNECTIVITY ................................. 103

3.1 Overview ....................................................................................... 103

3.2 Enhancing empathic accuracy processes .................................... 104

3.2.1 Imagination .............................................................................. 104

3.2.2 Connectivity ........................................................................... 104

3.2.3 Facilitating both imagination and connectivity in photographic stimuli ........................................................................... 105

3.2.4 Presentation modes: full body versus head and shoulders only .......... 105

3.3 Study 1: Imagination and Connectivity ......................................... 107

3.3.1 Exploratory questions, hypotheses, and testing conditions ............ 107

3.3.2 Method ...................................................................................... 108

3.3.2.1 Participants ........................................................................ 108

3.3.2.2 Materials ........................................................................... 108

3.3.2.2.1 Creation of the ‘silhouette’ ............................................ 108

3.3.2.3 Procedure .......................................................................... 109

3.3.3 Results ..................................................................................... 110

3.3.3.1 Photographic conditions .................................................... 111

3.3.3.2 Model sex ......................................................................... 112

3.3.3.3 Participant sex .................................................................. 113

3.3.4 Discussion ............................................................................... 114

3.3.4.1 Photographic conditions .................................................... 115
5.4 Comparing and Testing Photographic Presentation Modes ................................. 160
5.5 Overall Findings ........................................................................................................ 162
  5.5.1 Photographic presentation mode ........................................................................... 162
  5.5.2 Imagination, connectivity, and social context ...................................................... 163
  5.5.3 The AR-EA photographic measure ...................................................................... 164
5.6 Limitations .................................................................................................................. 165
5.7 Conclusions ............................................................................................................... 166
REFERENCES .................................................................................................................. 169
APPENDICES .................................................................................................................... 213
  A. Procedures used with photographic models during photographic
     sessions ......................................................................................................................... 213
  B. Relaxation and stretching exercises ......................................................................... 219
  C. Cleansing exercises ................................................................................................... 220
  D. Model emotion expression capture scale ................................................................. 221
  E. Sample booklet – Study 1: Open-ended responding ................................................ 222
  F. Procedure used for selection of photographs to be used for
     further testing .............................................................................................................. 224
  G. Final photographs selected across six models, by emotion category
     and intensity level ....................................................................................................... 228
  H. Creation of the silhouette ........................................................................................ 232
  I. Example of testing booklets: Study 1 and Study 2 ................................................... 234
  J. Scales used for validity testing: Study 2 ................................................................. 240
  K. Social context – backgrounds ................................................................................ 245
  L. Example of testing booklets: Study 1 .................................................................... 248
  M. Example of testing booklets: Study 2 ................................................................... 254
LIST OF TABLES

Table 2.1. Mean Self-Report Ratings of How Well Each Model Felt They Captured Each Emotion.................................................................................................................................................. 76
Table 2.2. Mean and Standard Deviations of EC and EI Scores for Male and Female Models ......................................................................................................................................................... 92
Table 2.3. Inter-Rater Accuracy Percentages by Photograph and Model.......................................................... 93
Table 2.4. Final Counts for Numbers of Photographs from each Model Selected for Inclusion in the AR-EA Test .................................................................................................................................................. 102
Table 3.1. Reported Cronbach’s Alpha Levels for the 3 Scales Chosen for Validity Testing.................................................................................................................................................................................. 124
Table 3.2. Mean and Standard Deviation Figures for Entire Sample for Each Validity Scale in the FBSil Condition Only ........................................................................................................................................ 126
Table 4.1. Correlations Between T1 and T2 for Mean Accuracy Score Percentages by Emotion Category for the Full Body (FB) Photographic Condition ........................................ 150
Table 4.2. Correlations Between T1 and T2 for Mean Accuracy Score Percentages by Emotion Categories for the Silhouette (FBSil) Photographic Condition ........................................................................................................ 151
LIST OF FIGURES

Figure 2.1. Average accuracy percentage of photographs by emotion category .......... 81

Figure 2.2. Average accuracy percentage of photographs by intensity level
.............................................................................................................................................. 81

Figure 2.3. Comparison of average accuracy percentages of photographs between
male and female models by emotion ................................................................. 82

Figure 2.4. No. of photographs that scored 0.00% accuracy by model...................... 83

Figure 2.5. Mean percentage accuracy ratings for each emotion category (EC)......... 94

Figure 2.6. Average accuracy percentages for three levels of emotion intensity (EI) .... 94

Figure 2.7. Mean EC and EI scores by model sex .................................................... 95

Figure 2.8. Mean EC accuracy percentages by model ............................................. 96

Figure 2.9. Comparison of mean percentage accuracy EC scores for each emotion
category by model sex ............................................................................................. 96

Figure 2.10. Average percentage accuracy scores of different intensity levels by model
sex ........................................................................................................................................ 97

Figure 2.11. Comparison of mean percentage accuracy scores for each emotion
category of study 1 with study 2 .............................................................................. 100

Figure 3.1. Mean percentages by emotion category ................................................ 110

Figure 3.2. Mean percentages for intensity levels across all emotions categories
except neutral ............................................................................................................. 111

Figure 3.3. Mean percentage scores across the different photographic conditions ...... 112

Figure 3.4. Mean percentage accuracy scores comparing male and female models
cross all emotion categories except neutral .......................................................... 113

Figure 4.1. Mean percentage total emotion accuracy scores for each background by
photographic condition ............................................................................................. 138
Figure 4.2.  Comparison of male and female participants’ overall emotion category accuracy percentages by background ............................................................... 140

Figure 4.3.  Mean accuracy score percentages for each emotion category comparing T1 and T2 for the full body (FB) photographic condition ................................. 149

Figure 4.4.  Mean accuracy score percentages for each emotion category comparing T1 and T2 for the silhouette (FBSil) photographic condition ............................ 150
Acknowledgements

There are many people who have contributed towards this thesis and for all those who have given me words of encouragement and advice during this process I give my heartfelt thanks. There are some people, though, that I would like to extend especial thanks to. My thanks go to my supervisors Associate Professor Guy Hall and Dr Courtney Field for their endless patience and guidance. You have turned a pig’s ear into a silk purse – thankyou. To my photographic models (whom I will not name – but you know who you are) for allowing me to take photos of you in your most vulnerable states and happiest ones too. An enormous thankyou to Liv Stockley my photographer, your support did not end with taking the photographs for which I am grateful. Mum and Dad and Aaron for listening to my endless talk of studies and empathy – Mum thanks for letting me photograph your kitchen. Brianna for letting me bounce ideas around and sharing the PhD pain. The psychology honours girls – always there with a sympathetic ear. To all the staff in the School of Law administration office – each one of you has offered endless words of encouragement, thank you. To Pia Broderick, Bethanie Gouldthorp, Anne Pedersen, and Kathy Modecki from the School of Psychology for your words of advice, wisdom, and encouragement. Kate Makowiecka, Murdoch’s University Copyright Advisor for her generous help with rewording my consent forms into normal human language. To Dean Jurgen Brohmer for generously hearing out my research ideas and offering financial support, as well as ongoing encouragement. The Student Guild, in particular Bec Thompson (Guild President) for allowing me access to the tavern to take photographs. To Vanessa Caparas (MUPSA) for listening to me moan and groan. To Maria Gardener from Thinkwell for her guidance and support.

Thank you to everyone who helped in any way – you have my gratitude.
Chapter 1

1.1 Overview

In our everyday interactions we interpret each other’s emotions and thoughts and try to predict other’s reactions to us and the environment. This interpretation is known as empathy: the “the ability to understand and share in another’s emotional state or context” (Cohen & Strayer, 1996, p. 988). Empathy helps us with our interpersonal interactions, a social navigation tool which allows us to understand others thereby changing and adapting our own behaviours to others as needed (e.g.: de Vignemont, & Singer, 2006). Increased empathy has been linked to pro-social behaviours such as helping others and engaging in altruistic actions (e.g.: Batson, Polycarpou, et al., 1997). A decrease in empathy, however, has been associated with offending and antisocial behaviours (Jolliffe & Farrington, 2004). For this reason empathy and empathic processes have been an important area of research both in psychology and criminology.

Empathic Accuracy is the ability to accurately interpret another’s emotions, mental state, motivations and thoughts (Ickes, 1993) in the context of the social arena in which the event/behaviour occurs (Gesn & Ickes, 1999; Ickes, Robertson, Tooke, & Teng, 1986; Sze, Goodking, Gyrulk, & Levenson, 2012; Wyer & Srull, 1986). Affective empathic accuracy (referred to as Affect Recognition-Empathic Accuracy (AR-EA) in this thesis) is specifically the ability to recognise and interpret another’s emotional state (see Gesn & Ickes, 1999; Ickes, 1993. The ability to successfully interpret another’s emotion is important in interpersonal and social interactions (Gleason, Jensen-Campbell & Ickes, 2009; Howland & Rafaeli, 2010; Ickes, 1993). It allows us to not only interact appropriately with another person in terms of their emotional state, but also allows us to judge our own impact on that person and adjust our behaviour accordingly.
This thesis focuses on three factors that can influence empathic processes - imagination (Davis, 1980; Eisenberg & Miller, 1987; Hoffman, 1984), connectivity between observer and target (Batson et al., 1995; Batson et al., 2007; Krebs, 1975; Davis, 1996; Zagefka, Noor, & Brown, 2013), and social context (Batson, Eklund, Chermok, Hoyt, & Ortiz, 2007; Batson, Turk, Shaw, & Klein, 1995; Carroll & Russell, 1996; Marian, & Shimamura, 2011; Walbott, 1988). Imagination allows us to put ourselves into the shoes of another, and imagine a situation or event from their point of view, aiding the effort to understand and interpret their reactions (Hoffman, 1987). The importance of imagination in the empathy construct can be seen through various measures of empathy. Davis (1996) in his Interpersonal Reactivity Index (IRI) includes the imagination process in two of his subscales – Perspective-taking and Fantasy. Whilst in research investigating factors that may elicit empathy, “imagine-self/imagine-other” instructions are given to participants in order to enhance empathy mechanisms (see Batson et al., 2003; Davis et al., 1996; Davis et al., 2004; and Lamm, Batson & Decety, 2007 for some examples).

The connection between the observer and target can influence empathic processes. The greater the familiarity between the target and the observer the more likely the target will engage in empathy processes (Krebs, 1975; Davis, 1996; Zagefka, Noor, & Brown, 2013; Stinson & Ickes, 1992). Empathy is also increased the more we value the welfare of the target (Batson et al., 1995; Batson et al., 2007), the more we view the target as being similar to ourselves (Batson et al., 1995) and the more interaction we have with the target (Marangoni, Garcia, Ickes, & Teng, 1995; Simpson, Ickes & Blackstone, 1995).

Social context too is important, as different social situations give rise to different expectations as to what behaviours and emotional displays are appropriate, and therefore what responses are acceptable (Ekman & Friesen, 1969, as cited in Kupperbusch, 1999, also see Ekman, 1999b). The research in this area is particularly limited, but indicates that context will influence our interpretations of the emotions of others (Walbott, 1988). For this reason it
is important that with any measure of empathic accuracy, imagination processes can be facilitated, and that social context is taken into consideration.

A common vehicle for testing and assessing affective empathic accuracy abilities are photographs which offer an effective and practical way of assessing how well a participant can interpret the emotions of a target. As a general rule these photographs depict models, males and females, expressing emotions. A match between the expressed emotion and the participant’s responses indicates high empathic accuracy abilities (e.g.: Ekman & Friesen, 1974a). The problem is that many of the photographs used in research have not been developed specifically for empathy processes and as such have not included ‘imagination’ or ‘social context’. There appears to be no research into how the stimulus itself (the photographs) can be used to enhance imagination processes or induce social context and how this, in turn may impact empathic accuracy abilities.

The aim of this PhD, therefore, was to investigate if photographic elements could be added or manipulated to enhance and/or influence factors that would impact empathic accuracy abilities: namely imagination, connectivity and social context. Two different techniques were used to achieve this. Firstly an imagination component was added to the photographs by way of a blacked out silhouetted figure. Using the silhouette, participants imagined that they were in the photograph with the target, the purpose being to enhance both imagination processes in participants, as well as the connection between participant and target. The second involved manipulation of photographic backgrounds in order to situate the target, and therefore the participant, into different social contexts. In order to achieve this, however, an original set of photographic stimuli needed to be developed. The second aim of this thesis, therefore, was develop and test these photographs for use in a new photographic measure for assessing empathic accuracy.

In pursuit of these aims, the studies in the thesis deal with both goals concurrently, following the investigation of empathic factors in photographs alongside the development of
the new empathic accuracy photographic measure. The first tests dealt with the taking of the photographs and initial testing to establish baseline patterns of responding from participants. The following study investigated the effect of inserting a simple photographic device (i.e.: a blacked-out silhouetted figure) designed to enhance imagination processes and the connection between observer and target. During this study, the first validity and reliability testing was conducted of different versions of the photographs. The next study was designed to assess the influence of different social settings, in the form of different backgrounds in the photographs, upon participant AR-EA abilities. Embedded within this testing was a test-retest of the photographs to confirm their reliability, representing the final testing stage of the AR-EA photographic measure.

The results of these studies point to a new way of using photographic-based material for assessing empathic accuracy in individuals. The findings revealed that both imagination processes and connectivity can be enhanced through the use of silhouettes in the photographs. Results indicated that by adding this component, the photographic stimulus measure of empathic accuracy abilities became more reliable. It was also found that social context influenced empathic accuracy performance.

These results question the methods used previously with photographic-based stimuli, revealing that more can be done to improve empathic accuracy assessment and measurement through simple and cost effective means. These simple techniques have allowed for a more nuanced investigation of empathic accuracy abilities, raising intriguing questions about the influences on empathic accuracy abilities which will open up new areas for inquiry and research into both empathy and empathic accuracy processes.

The focus of the current chapter will be to briefly outline the evolution of the empathy concept before moving on to the different types of empathy. A distinction between empathic disposition (the tendency to empathise), and empathic accuracy (how accurately we interpret
others during the empathy process), will be made, highlighting the importance of empathic accuracy and its assessment.

The second part of this chapter will provide a review of some of the different methods for measuring and assessing empathic accuracy, in particular, affective empathic accuracy. It will include Ickes’ interpersonal interaction paradigm, the use of movies and stories, and lastly photographic stimuli. The aim here is to highlight the advantages of photographs as an assessment tool of empathic accuracy abilities, as they are practical, adaptable, cost effective, easily standardised, as well as being user friendly. Photographic stimuli offer an effective method for manipulating factors such as imagination processes, connectivity, and social context, compared to other assessment protocols.

Finally the chapter will map out the rest of the chapters in the thesis, providing a brief overview of the aims and content of each.

1.2 What is Empathy?

1.2.1 A brief history.

The origins of empathy can be found in the German word *Einfühlung* meaning to project oneself into an object of art or beauty, where one becomes ‘fused’ with the object, thus becoming the subject (Wispe, 1987). Lipps defined it as the process by which one perceived and experienced another object or person, subjectively and consciously, thereby coming to experience it/them (Hunsdahl, 2006). Empathy came into use around 1909 when Titchener coined the term to mean the ability to experience another’s thoughts or feelings (Hunsdahl, 2006). This process according to Titchener involved imagining yourself in the other’s position, acting out their experiences through motor-mimicry, copying their facial expressions, and body posture. In this way one could feel what the other was experiencing (Wispe, 1987).
1.2.2 Empathy – more than just sympathy.

Empathy was often viewed in the context of a person’s pain and suffering, resulting in empathy becoming almost synonymous with ‘sympathy’ (Kohut, 1980); however, the process of empathy encompasses all emotional experiences, not just pain, grief and distress. For instance, Freud posited that humour worked on the basis of empathy, as we were able to put ourselves into the position of the unfortunate person who’s misfortune we then find funny (Allport, 1937). Hoffman (1987) suggests that empathy can result in other emotions such as anger in the face of an injustice to another; or sadness at the joy when seeing a little girl playing with a puppy, knowing that the little girl has been diagnosed with a terminal condition and does not have long to live. Empathy, then, forms the conduit through which we can experience many different emotions.

Empathy can act as a facilitator for interpersonal relations (Kohut, 1980; Rogers, 1959 as cited in Wispe, 1987; Schwartz, 1994), social interactions (Davis, 1996; Levy, Freitas, & Salovey, 2002; Watt, 2005), as well as social coherence and communication (de Vignemont & Singer, 2006). The ability to appreciate what another person is going through, acts as a motivator for altruistic behaviours such as helping (de Vignemont & Singer, 2006; Hoffman, 1975). It also allows us to predict the behaviour of others, therefore informing our responses to them (Davis, 1996). In these ways, empathy processes assist us in our everyday social encounters, aiding in communication, interpersonal relationships, the helping of others, and in understanding others behaviours. Empathy, therefore, is more than just a sympathetic reaction; it is a social navigation tool, as suggested by Kohut (1980).

Empathy can facilitate our social interactions with others by allowing us to interpret and appreciate another’s emotional and cognitive processes from their point of view. When these cues are interpreted accurately, our insights can provide us with information about the other, guiding our behaviour and letting us know what is, and what is not, an appropriate response (e.g.: Hill et al., 2008; Salmivalli, Poskiparta, Ahtola, & Haataja, 2013). Ongoing empathic
processes then allow us to gauge the other’s reactions to our own behaviour, furnishing us with the opportunity to monitor social goals and adjust our behaviour accordingly (Hoffman, 1984; Monin & Schulz, 2009).

1.2.3 Theories of empathy.

Two main streams of empathy research became apparent in the literature, the cognitive approach and the affective approach. The cognitive approach looks at the ability to interpret another’s mental and emotional states (Allport, 1937; Cohen & Strayer, 1996; Dymond, 1949; Madzia 2013; Selman, 1971; Singer, 2006; Stotland, 1969; Strayer, 1987; Vollm et al., 2006), whilst the affective approach focuses on the emotional outcomes of ‘feeling for’ another (Aderman & Berkowitz, 1970; Feshbach & Roe, 1968; Stotland, 1969).

1.2.3.1. The cognitive approach.

Cognitive empathy is the ability to ‘think’ or transpose oneself into another’s situation in order to appreciate what that other person is experiencing (Allport, 1937; Cohen & Strayer, 1996; Dymond, 1949; Madzia 2013; Selman, 1971; Singer, 2006; Stotland, 1969; Strayer, 1987; Vollm et al., 2006; Wakabayashi, et al., 2006). This is also known as perspective-taking or role-taking. According to this perspective, an indicator of high empathy in an individual is their ability to correctly interpret aspects of another’s personality (Dymond, 1949), the ability to distinguish self from others (see Dymond, 1950; Batson, Sager, et al., 1997), or predicting behavioural choices (Selman, 1971; Iannotti, 1985). The historical perspective of empathy being a dominantly cognitive process, has now given way to more modern understandings of empathy, which is now often considered to be part of wider social cognition constructs. These include Theory of Mind models (see Baron-Cohen, O’Riordan, Stone, Jones & Plaisted, 1999; Dolan & Fullam, 2004; Gregory et al., 2002), social information processing concepts (Ickes, Stinson, Bissonnette, & Garcia, 1990; Stone, Baron-Cohen, Calder, Keane & Young, 2003), social cognition (Fiske, 1993; Gregory et al., 2002; Marton, Wiener, Rogers, Moore, & Tannock, 2003).
Knowing and understanding what another person is experiencing allows the observer to make predictions regarding the other’s behaviours and motives (Milgram, 1960; Selman, 1971; Stotland, 1969; Wakabayashi, et al., 2006). As already mentioned above, when we interpret these cues accurately, we are provided with information that guides our interactions with that person, letting us know what behavioural responses might be appropriate (e.g.: Hill et al., 2008; Hoffman, 1987; Lanyon, 1967; Salmivalli, et al., 2013). Empathic processes involve ongoing interpretations that allow us to monitor responses towards our behaviours, giving us opportunity to make adjustments, as we pursue our social goals (Hoffman, 1984; Monin & Schulz, 2009). This ability to appreciate another’s situation has been linked to altruistic and helping behaviours (Batson, 1987; Batson, 1991; Batson, Batson, Slingsby, Harrell, Peekna & Todd, 1991; Batson, Duncan, Ackerman, Buckley, & Birch, 1981; Batson, et al., 2007; Batson, O’Quin, Fultz, Vaderplas, & Isen, 1983; Batson, Sager, et al., 1997; Coke, Batson, & McDavis, 1978; Hogan, 1969), and the decrease of offending behaviours (Jolliffe & Farrington, 2004; Jolliffe & Farrington, 2007; Lisak & Ivan, 1995; Mehrabian, 1997).

1.2.3.2. The affective approach.

The other aspect of empathy are the affective outcomes that result from empathy processes, which some researchers felt had been largely ignored in earlier research (Feshbach & Roe, 1968). Clark (1980) argued that it was not the cognitive elements of empathy that were the most important, but rather the emotional elements. This differing approach to empathy involved the vicarious emotional responses that were elicited as a result of observing and feeling for another (e.g.: Stotland, 1969; Aderman & Berkowitz, 1970). Within this framework, however, several different views of affective empathy became apparent.
Empathy elicited emotional responses of the observer could be congruent with (matching) the emotion of the target (Davis, 1996; Feschbach & Roe, 1968; Hoffman, 1984; Hoffman, 1987; Staub, 1987; Stotland, 1969), or could be incongruent (not matching) the emotion of the target (Davis, 1996; Hoffman, 1984; 1987). According to Hoffman (1987) these elicited emotions could, at times, be quite intense.

It was these emotional experiences of the observer that lead researchers to theorise that certain emotions may motivate specific behaviours (see Rudolph, Roesch, Greitemeyer & Weiner, 2004). Batson’s Empathy-Altruism model is based upon the notion that in the face of another’s pain and distress, an observer will also experience that distress, through empathy, and hence be motivated to behaviours that would relieve that distress in the victim, such as helping (Batson, Fultz, & Schoenrade, 1987; Coke, et al., 1978). There are two main theories about how this occurs, the first through genuine concern for relieving the distress in the victim (Toi & Batson, 1982), and the second through the desire of the observer to relieve their own distress felt in the face of the victims’ suffering (Cialdini, Schaller, Houlihan, Arps, Fultz, & Beaman, 1987). Regardless of whether the helping behaviour is altruistically or egoistically driven (see Batson et al., 1989) the result is a reduction in the victim’s and therefore the observer’s distress. More on the interactions between empathy and behaviour will be discussed later.

1.2.3.3. Bringing the two approaches together.

Later research into empathy recognises the role that both cognitive and affective processes have to play. Krebs and Russell (1981) make the distinction nicely stating that empathy is feeling what another is feeling, whereas perspective-taking is understanding what that person is thinking or feeling. Gladstien (1983) clarified the interaction of both cognitive and affective aspects of empathy stating that empathy itself was a ‘multistage interpersonal process’. According to Gladstein this involved many sub-processes such as emotional contagion (a process whereby the emotion experienced by the target is ‘transmitted’ to the
observer); interpretation of the situation and the emotion; as well as role-taking. From this, then, there was recognition that empathy could involve a range of emotional experiences and was not limited to just emotions of sympathy and concern (Kohut, 1980), and could result in a range of behavioural outcomes, not all of them prosocial (Staub, 1987; Hoffman, 1984, 1987, 2000). Cognitive processes are prone to interpretation errors and biases, in what Staub (1987) referred to as ‘false empathy’, which could result in harmful instead of helpful behaviours (see Covell, Huss, & Langhinrichsen-Rohling, 2007 for an example of this process). For instance, a person misinterprets another as being angry rather than in pain. The resultant violent behaviour would seem inappropriate, from the victim’s point of view, but totally justified from the observer’s point of view. If, as Gladstein (1983) surmised, empathy was made up of multi-processes, then a break-down or deficit in one area could impact all the others, leading to unexpected, and even antisocial outcomes.

Through the viewpoints of these researchers empathy was now seen as a complex construct that involved both affective and cognitive components. Two seemingly disparate aspects of empathy were considered as one process, where both cognitive and affective aspects would influence behavioural outcomes. These behaviours were not limited to just prosocial either, but these new models entertained the possibility that antisocial behaviours could also result from empathic processes. The most complete model of empathy that includes both cognitive and affective empathic elements is that developed by Davis (1980).

1.3 Davis’ Multi-dimensional Model of Empathy

Davis (1980 & 1983), stated that empathy was a complex construct, involving perceptual, cognitive and affective processes, all working together to elicit an emotional reaction in an observer in response to another’s experience. Cognitive empathy processes were seen as active and included elements such as the ability to understand another’s situation and any feelings they may be experiencing (Davis, 1996). This type of cognitive empathy was also referred to as perspective-taking or role-taking. Davis (1996) outlined three
different types of perspective-taking: perceptual, taking the literal visual perspective of another; cognitive, imagining the other’s thoughts and motives; and affective; imagining the other’s feelings or emotions.

Davis (1996) described the affective components of empathy as the emotional responses that occur in response to another’s experience. These emotional reactions were considered passive, the result of active cognitive processes such as perspective taking. These elicited emotions could be congruent or incongruent with the target’s responses. Davis (1996) draws on the research of Staub (1987) and Hoffman (1987) suggesting that congruent emotional reactions were based in self-centred outcomes meaning that the observer focused on the way they themselves were feeling at the time, rather than the target. Conversely, reactive outcomes, such as anger or sadness, can result in personal distress, and are other-oriented (Davis, 1996). In these cases the observer is experiencing the event as though it was happening personally to them and therefore reacts on behalf of the victim.

Davis (1996) believed that both cognitive and affective aspects of empathy needed to be considered when measuring the empathy construct. To capture both the cognitive and affective aspects of empathy, Davis included four subscales in his Interpersonal Reactivity Index (IRI) a measure of dispositional empathy. Personal Distress (PD), the amount of anxiety and discomfort an observer feels as a result of empathising with another; and Empathic Concern (EC), the amount of compassion and concern one feels for another are aspects of affective empathy, are both affective processes. The cognitive aspects of empathy are captured by the Fantasy Scale, which is the likelihood that a person might identify and relate to characters in movies, plays and novels, as well as Perspective-taking (PT) which is the tendency to take another’s point of view. The theory behind these four subscales being that those with high PT tendencies will experience more EC for others and less personal distress for themselves, and therefore will demonstrate high empathy. According to Batson’s (Toi & Batson, 1982) model, this should result in helping behaviours. When personal distress is high,
helping will be done for the observer’s sake (to reduce their own distress) rather than the victim’s (Cialdini et al., 1987).

1.4 Defining Empathy in the Context of this Thesis

The definition of empathy to be used in this thesis should be explicitly made so that the research in the following chapters can be viewed in this context. As a part of this, it will be demonstrated that the empathic accuracy concept is made up of more than one process, of which this thesis will examine only one termed affect recognition-empathic accuracy (AR-EA). That is, the ability to accurately identify specific emotional experiences in others. The arguments for this are presented briefly below.

Empathy, in the context of this thesis, is considered to be a mixture of Davis’ multidimensional model of empathy, where both affective and cognitive elements are included, as well as empathic accuracy. Therefore the definition used includes both dispositional empathy and empathic accuracy. Specifically, this thesis defines empathy as: A response to another’s experiences through the ongoing interpretation processes of the other’s thoughts, feelings, motives, situation, as well as verbal, visual, social, and historical cues, that may result in an emotional reaction in the observer, and which may subsequently influence behaviour. The advantage to this definition will be that it views empathy as encompassing both processes and outcomes. By incorporating both aspects of empathy in this way the research questions that can be asked regarding empathy are broadened as a result.

The key concepts here are that someone observes another; interprets their situation; and has a reaction, emotional, behavioural or both, as a result. The interpretation processes involved in this definition of empathic accuracy are “ongoing”. This is in recognition of the fact that the empathy process is not a one off event, rather it involves a continual assessment and reassessment of others in any interpersonal and/or social situation. In this way we are able to adjust our responses and behaviour from moment to moment, seeing our successes and failures, in the pursuit of our social goals. Multiple interpretations of the other are made.
during this process, “other’s thoughts, feelings, motives, situation, verbal, visual, social…..cues”. As well as this we draw upon our own experiences and memories “social, and historical cues”, as argued by Hoffman (1984). All of these interpretations and evaluations inform our “emotional reaction” to the observed target/event. I use the term “may result in an emotional reaction” in recognition that some individuals may be able to suppress their emotional response to another (as seen in the literature regarding empathy and offending reviewed, section 1.4.1.2, below). Our emotional reaction may then “subsequently influence behaviour”. Again there is an acknowledgement here that even if an emotional reaction is elicited, this may not influence behaviour, as some individual may behave in a manner that is counter to the emotions that have been elicited (see literature regarding empathy and offending, section 1.4.1.2 below).

This definition of empathy also acknowledges the fact that the emotional responses and the behavioural outcomes will not always be congruent with, or appropriate to, the target in question. In other words, this definition does not presume a pro-social outcome as a result of empathic processes, rather it encompasses all possible responses by the observer. These outcomes are predicated upon the interpretations involved during empathy and are not governed by sympathetic mechanisms alone.

Empathic accuracy should therefore be considered as involving two major elements – cognitive processes (interpretations) and individual outcomes (emotional and/or behavioural). The two processes are linked whereby the accuracy of our interpretations of another’s thoughts, feelings, and situation will drive the emotional and/or behavioural outcomes that result. The appropriateness of an actor’s emotional and behavioural response is, however difficult to judge, as it may be the case of an appropriate reaction from the actor’s point of view, but inappropriate when observed by another. For this reason this outcome-based empathic accuracy will not form part of the scope of this thesis.
The interpretation processes involved during empathy consist of both cognitive and affective elements. Cognitive recognition-empathic accuracy involves the accurate recognition and identification of another’s thoughts and motives. Affect recognition-empathic accuracy (AR-EA) is the accurate interpretation of another’s emotions. It is much easier to assess affect recognition-empathic accuracy (AR-EA) abilities, as this involves emotional expressions, a concept that has been the subject of much research, resulting in several different methods of successfully standardised and valid apparatus. Thus cognitive recognition-empathic accuracy will not be dealt with in this thesis. Instead the focus of this thesis, and the studies contained in the following chapters will focus upon the affect recognition – empathic accuracy (AR-EA) construct.

In order to understand the intersection between empathy and behaviour, however, all aspects of empathy should be considered. Empathy consists of a series of processes and so its impact on emotional and behavioural outcomes cannot be understood by looking at only one of these processes. Therefore the following review of research includes these different aspects of the empathy construct in order to offer a fuller understanding of the important interactions that take place, and how these may impact emotional and/or behavioural outcomes.

1.4.1 Empathy and behaviour

Empathy is most notably known for its link with altruistic and pro-social behaviours. In Coke, et al.’s, (1978) model of helping behaviours it was proposed that empathy mediated the relationship between perspective-taking and helping. In this model perspective-taking increased empathic emotion which then increased helping, therefore those that experienced the most empathic emotion also offered the most help (Coke et al., 1978). This became known as the empathy-altruism hypothesis (Batson, et al., 1981). The research has consistently confirmed this link, finding that feelings of empathy can act as a motivator to help another (Batson, et al., 1991; Batson, Batson, Todd, Brummett, Shaw & Aldeguer, 1995; Batson, et al.,
Empathy has also been found to improve attitudes towards stigmatised groups (Batson, Polycarpou, et al., 1997). This theory of empathy and altruism led to other theories being proposed that empathy could also bring about a reduction in antisocial and offending behaviours.

### 1.4.1.1. Empathy and antisocial behaviours.

The research exploring the relationship between empathy and offending, aggressive and antisocial behaviours has found that low levels of empathy are linked to aggression in boys (Feschbach & Feschbach, 1969); bullying behaviour (Jolliffe & Farrington, 2006; Choi & Cho, 2012); verbal abuse (see Heilbron & Prinstein, 2008); cyberbullying (Schultze-Krumbholz & Scheithauer, 2009; Ang & Goh, 2010); conduct-disordered youth (Cohen & Strayer, 1996); sibling abuse (Graham-Bermann & Cutler, 1994); and aggressive attitudes (Anderson, et al., 2010; Cohen & Strayer, 1996; McPhedran, 2009). In children, empathy is developed over time (Rothenberg, 1970; Vaish, Carpenter, & Tomasell, 2009), and it is argued that this development of empathy forms an important part of aggression regulation in children (Feschbach, 1975).

In adult samples, empathy was found to be negatively correlated with: aggression (Richardson, Hammock, Smith, Gardner, & Signo, 1994; Stanger, Kavussanu, Willoughby, & Ring, 2012); alcohol induced aggression (Giancola, 2003); sexual aggression (Dean & Malamuth, 1997; Bernat, Calhoun, & Adams, 1999); sexually coercive men (Baumeister, Catanese & Wallace, 2002); and sexually aggressive attitudes (Bushman, Bonacci, Dijik, & Baumeister, 2003; Hall & Barongan, 1997).

This link between low levels of empathy and high levels of antisocial behaviours is not, however, always so straightforward. Miller and Eisenberg’s (1988) meta-analysis of the empathy-aggression research revealed that the relationship seemed to depend upon the types of measures used. Observational measures of empathy did not result in a relationship with
aggression, and whilst questionnaire-based measures of empathy fared better, the negative relationship between empathy and aggression was only weak (Miller & Eisenberg, 1988).

There are problems with some of this research. Some research (as cited above) involves children and juveniles. According to Hoffman (1984, 1987) and Feasbach (1975) many empathy processes, such as perspective-taking, are heavily reliant upon the cognitive development of the observer. For instance, Hoffman (1975; 1984) argued that there are different developmental levels of empathy. Younger children are only aware of simple arousal, and cannot discern between themselves and others. As the child matures, however, these processes become more complex: from separation of self and other; to role taking; to appreciation of another’s situation as well as emotion. Empathy, therefore, cannot be considered a static process developmentally, meaning that different developmental stages need to be taken into consideration when looking at the relationship between empathy and behaviour, especially antisocial behaviours. Many samples used in the research are non-offenders. For example some studies used college males (Rapaport & Burkhart (1984); Dean & Malamuth, 1997; Bushman et al., 2003), or had participants that self-identified as aggressive (Bernat et al., 1999; Richardson et al., 1994), therefore making generalisations about how empathy may impact offending behaviours may not be valid in these particular studies.

The definitions used for violence and aggression can also be called into question. Definitions vary considerably in the research making comparisons difficult. For instance in a study into sexual aggression by Bernat, et al. (1999), the percentage of males that had engaged in only verbal aggression was 43.7%, whilst those engaging in physical aggression was only 12.6%. The group labelled as sexually aggressive had a mean score of 3.69 (a measure of sexual aggression) out of a range of 1-8 from the selected 6 items from the Sexual Experiences Survey (Bernat et al., 1999; Koss & Gidycz, 1985). In many of these types of studies, definitions of aggression seem quite ‘soft’, making it difficult to draw adequate conclusions regarding the relationship between empathy and these types of behaviours.
1.4.1.2. Empathy and offending.

In offender samples, Mehrabian (1997) found that measures of emotional empathy were negatively correlated with aggression and violence in incarcerated juveniles. Low empathy has also been demonstrated in convicted child molesters (Chaplin, Rice, & Harris, 1995); self-reported rapists and rape tendencies (Lisak & Ivan, 1995; McDonel & McFall, 1991); and violent offenders (Jolliffe & Farrington, 2007). A meta-analysis conducted by Jolliffe and Farrington (2004) found a strong negative relationship between the two factors for violent offenders, whilst only weak negative relationship was evident for sex offenders. After controlling for participant intelligence and SES, however, the relationship between violent offenders and a lack of dispositional empathy disappeared. Furthermore, only cognitive empathy (perspective-taking) was found to be strongly and negatively related to offending, with affective empathy (emotional responses) having only a weak relationship (Jolliffe & Farrington, 2004). Jolliffe and Farrington (2004) concluded that the relationship between empathy and offending behaviours was not straight forward. To understand why this is we need to understand how, theoretically, empathy influences offending behaviours.

In earlier research, Baron (1983) suggested the negative relationship between empathy and aggression would be the result of the incompatibility of two conflicting processes. If empathy enables one to feel the emotional state of another, then those perpetrators with high empathy will be able to feel the pain they are causing to the victim, experiencing that pain as their own (called personal distress by Davis, 1996). Therefore the perpetrator experiences a negative consequence from their behaviour, resulting in the pain- causing behaviour to cease (Cialdini, et al., 1987; Davis, 1996; Feschbach, 1964; Hoffman, 1975). In cases of low empathy, no such emotional sharing will eventuate. The perpetrator will not feel the victim’s pain and so the behaviour will continue (Feshbach, 1964). Cialdini et al. (1987) argued that high empathy elicited a process of ‘negative state relief’, resulting in the
perpetrator seeking to reduce their own pain via the victim, thereby causing more harm. This type of theorising, however, may be overly simplistic.

There is an assumption here that needs to be explored further. If we accept the premise that experiencing the victim’s distress as their own will result the behaviour becoming an aversive event resulting in a reduction of the pain-causing behaviour, then we must assume that the victim’s pain has been accurately interpreted by the perpetrator as pain. What these basic theories of empathy and offending behaviour assume is that the interpretation processes involved (empathic accuracy processes) are themselves operating without bias or deficit. This is not always the case. In fact, where interpretation is inaccurate, empathy mechanisms may increase, and not reduce, the offending behaviour.

### 1.4.1.3. How empathy can increase antisocial behaviour.

There are several explanations for how empathic processes may increase antisocial behaviours. Firstly, perpetrators may misinterpret the victim’s distress for another reaction. Secondly, personal distress elicited through empathic processes may result in maladaptive mechanisms to cope with the unpleasantness, which may result in an escalation of the aggressive behaviour (Batson, 1990; Hoffman, 1981; Milner, Halsey, & Fultz, 1995). Thirdly, perpetrators may not be reacting in an emotionally appropriate way to the victim’s distress. That is, that they experience pleasure rather than pain in the face of the victim’s pain (Berner, Berger, & Hill, 2003; Holt, Meloy, & Strack, 1999; Staub, 1987, and also see Dietz, Hazelwood & Warren, 1990; Hunter & Becker, 1994; Sonnby-Borgstrom, 2002). Fourthly, offenders may lack the emotional engagement required to elicit appropriate emotional responses (see Ali, Anorim, & Chamorro-Premuzic, 2009; Cleckley, 1988; Frick & White, 2008; Hare, 1996; Hare, Hart, & Harpur, 1991; Kimonis, Frick, & Barry, 2004; Marshall & Barbaree, 1989; Paradini, Lockman & Frick, 2003). Finally, there is evidence to suggest that empathy may be quite specific towards some groups, but not others, with some offenders displaying victim-centred deficits in empathy (Baumeister, et al., 2002; Chaplin, et al., 1995; Fernandez, Marshall,
Lightbody, & O’Sullivan, 1999; Fisher, Beech, & Browne, 1999; Marshall, Hamilton, & Fernandez, 2001; McGrath, Cann, & Konopasky, 1998; Varker, Devilly, Ward, & Beech, 2008), or empathy deficits that are present only at the time of offending (Ward, Hudson, & Marshall, 1995).

This means that deficits can occur in the level of empathy experienced (disposition), as well as in the interpretations of others made during the empathy process (empathic accuracy). For this reason it is important that we consider both dispositional empathy and affective empathy as two separate, but closely related constructs. A high amount of dispositional empathy does not guarantee correct emotional interpretations (AR-EA) of others. Neither do high AR-EA abilities guarantee high levels of dispositional empathy. Research comparing these two different aspects of the empathy confirm this, finding little or no relationship between the two (Eisenberg-Berg & Lennon, 1980; Hall 1984; Zaki, Bolger, & Oschsner, 2009), and yet the prominent theories of empathy that explain offending behaviour seem to assume the opposite.

Although many studies cite the importance of accurate empathic processes (in particular perspective-taking)(e.g: Carich, Metzger, Baig & Harper, 2003), it is unclear exactly how this accuracy is assessed, if at all. Varker et al, (2008) point out that dispositional empathy measures are primarily used to assess empathy in offenders. Some combine the self-report questionnaire with scenarios designed to incite emotional reactions, which are judged as appropriate or not (e.g: Chaplin et al., 1995). Again, though, this is not the same as determining the accuracy of empathic processes.

Jolliffe and Murray (2012) recognise that empathy is comprised of different elements and recommend that empathy assessments of offenders should include several different measures to ensure that these different empathic elements are captured. They argue that different assessments of empathy are necessary in order to fully appreciate and understand what kind of empathy deficits may be present (Jolliffe & Murray, 2012). As will be
demonstrated below, deficits may be present in any part of the empathy process – dispositional deficits, interpretation deficits (including AR-EA abilities), and deficits in emotional and behaviour outcomes.

1.4.1.4. Empathy: biases, misinterpretations, and attributions.

It has already been established that the empathy construct is made up of several different concepts and processes. Errors, biases or deficits in any one of these processes can result in behaviours that may be anti-rather than pro-social. As outlined above, antisocial behaviours have been consistently linked with low levels of dispositional empathy. Deficits in the interpretations of others, too can lead to similar outcomes.

Faulty attributions of another’s situation can influence the level of help given to a victim (Davis, 1996; Hoffman, 1987). Where we believe that events are beyond the control of the observed other, and their distress is not of their own making, we are more likely to feel sympathetic distress (Hoffman, 1987) and will therefore act to reduce that distress. Where we believe that the situation has been caused by the victim then feelings of sympathy will be reduced, helping behaviours are less likely, and other reactions such as aggression can result (Betancourt, 1990; Betancourt & Blair, 1992; Meyer & Mulherin, 1980; Rudolph, et al., 2004). There is a solid body of research demonstrating that violent behaviours are strongly linked with faulty interpretations of others. Serin and Kuriychuk (1994) looked at the social and cognitive processing deficits of psychopaths. They suggest that attribution errors combined with other factors such as reduced impulsivity control and event/cue ambiguity can all combine to increase the probability of aggressive behaviours (1994). Dolan and Fullum (2006) also found that male criminals displaying psychopathy or antisocial personality disorder displayed deficits in decoding of sad faces.

These deficits are not just evident in general attributions or misinterpretations, but also in empathy-specific processes. According to the Bradbury and Finchman model presented by Davis (1996), both empathy and perspective-taking processes are involved when we
interpret and therefore react to another. That is, empathic processes will influence and be
influenced by attribution processes in regards to another’s behaviour. Davis (1996) includes
such processes as interpersonal judgements, attributional judgements (of the target’s
behaviour/situation), and the accuracy of perspective-taking processes. These faults in
attribution and interpretation accuracy have been demonstrated in the research. Clements,
Holtzworth-Munroe, Schwienle, and Ickes (2007) found that violent men demonstrated
empathic accuracy deficits towards their intimate partner’s thoughts and feelings, whilst
Hudson et al. (1993) reported that child molesters were less accurate at identifying the facial
emotions of either adults or children compared to non-offender controls. Similar findings have
been reported by Schweinle, Ickes, and Berstein (2002) where husbands demonstrated poor
empathic accuracy not only to their intimate other’s but also to female strangers.

Deficits too can be seen in violent populations, other than those involved in intimate
partner or domestic abuse. Research has found that various sexual offenders display empathic
deficits (Ward, et al., 1995; Ward, McCormack, & Hudson, 1997). Specifically these deficits
seem to be in the ability for offenders to accurately infer the mental states of others (Blake &
and in identifying the emotional states of others (Hudson et al., 1993). Difficulty recognising
fearful expressions was also found in antisocial samples (those that displayed high levels of
violence and aggression), across a range of studies (Marsh & Blair, 2008). Social and cognitive
processing deficits (empathic accuracy) were also found in violent offenders (Serin &
Kuriyuchuk, 1994).

Disturbingly the research also indicates that some offenders display normal or superior
empathic accuracy abilities. Blair (2005) argues that psychopaths do not display cognitive
empathic deficits, instead being able to read other’s thoughts and motives quite well. The
deficits instead lay with their affective empathic accuracy, the ability to discern emotions in
others. In particular these deficits seem to be restricted to disgusted (Kosson, Suchy, Libby, &
Mayer, 2002), fearful and sad faces (Blair, Budhani, Colledge, & Scott, 2005; Blair, Colledge, Murray & Mitchell, 2001). Other research has found that some offenders are actually very good at interpreting the emotions of others. Giannini and Fellows (1986) found that rapists, compared to controls, were better at interpreting the emotions, whilst others studies have found that violent non-sex offenders were better at this ability than convicted sex offenders, thieves and drug offenders (Hudson, et al., 1993). Babcock, Green, and Webb (2008) found that men diagnosed with non-clinical borderline/morphic disorder were more accurate at interpreting others compared to both interpersonally violent (intimate partner abusers) and non-violent men. Generally violent men, those that reported being violent both inside and outside of the home, did display deficits (Babcock et al., 2008).

Finally, psychologically and developmentally disordered populations also demonstrate deficits in empathic accuracy. Studies have found that participants diagnosed with schizophrenia displayed deficits in accurately identifying a range of emotions (Alfimova, Abramova, Barhatova, Yumatova, Lyachenko, & Golimbet, 2009; Carter & Neufeld, 2007), as well as deficits in perspective taking (Dentl et al., 2009; Langdon, Coltheart, & Ward, 2006). Schwartz, Vaidya, Howard Jr., and Deustch (2010) found this deficit to be evident only for the interpretation of fearful faces. Developmentally disordered adults who viewed a film of two people interacting, were less successful than controls at inferring the thoughts and feeling of the characters (Roeyers, Buysse, Ponnet, & Pichal, 2001). Empathic accuracy deficits have also been found in depressed individuals (Surguladze, Young, Senoir, Brebion, Travis, & Phillips, 2004), those with affective disorders (Ekman, Matsumoto, & Friesen, 1997), Autistic youth (Demurie, De Corel, & Roeyers, 2011) and conduct disordered youth (Cohen & Strayer, 1996). Participants who reported high levels of social anxiety were found to display a threat bias when interpreting emotions, misinterpreting ambiguous stimuli as threatening (Yoon & Zanbarg, 2007).
Bias and misattributions are not only present in specific populations like those above, but can occur in our everyday interactions as well. In a study by Maner et al., (2005) social goals of participants were manipulated, to be either mate-oriented, or to protect the self. They were exposed to pictures of male and females from various ethnicities displaying neutral expressions. It was found that when the social goal was mate-oriented, that is the goal was to successfully find a mate, male participants perceived sexual arousal in the neutral female faces they viewed. When the goal was self-protection, Black male and Arab neutral faces were interpreted as angry (Maner et al., 2005). Therefore social goals play a role in our perceptions of others, and will open up our interpretations of others to bias.

There is also evidence that these deficits may manifest from an early age. Main and George (1985) observed the reactions of abused and non-abused toddlers to the distress of other children around the same age. Whilst non-abused children responded with interest, concern or empathy, none of the abused toddlers showed any concern at all. Instead they seemed to act out in fear or anger towards the distressed targets (Main & George, 1985).

One cannot assume, however, that the decreased ability in empathic accuracy indicates a like inability in dispositional empathy. In fact it makes more logical sense to assume that empathic processes are present, but that attribution and cognitive errors result in a misinterpretation of the other’s motives, thoughts and feelings, which elicit inappropriate behavioural responses such as violence. Therefore the deficits reside in empathic accuracy, not necessarily dispositional empathy.

One of the implications for these findings is in the consideration of treatment and rehabilitation programs for offenders. Studies that have linked empathy with decreased aggression have been the catalyst for the inclusion of some form of ‘empathy training’ in many rehabilitation and treatment programs for offenders. Most of these programs revolve around ‘Empathy Groups’ which address issues such as accountability for the offence and victim empathy (Schwartz, 2003). Many empathy based treatments seem to be focused towards sex
offenders (see Carich, et al., 2003; Hanson, 2003), whilst empathy training has also been recommended for child molesters (Chaplin, et al., 1995), anti-social youths (Paradini, et al., 2003) and male domestic violence offenders (Scott & Wolfe, 2003). These programs, however, seem to rely only upon measurements of dispositional empathy, and not empathic accuracy (e.g.: Vaker et al., 2008).

Given that deficits may occur during empathic accuracy processes as well as in levels of dispositional empathy it seems short-sighted that only one type of empathy is being assessed before treatment. As has already been established, low levels of dispositional empathy do not equate to like deficits in empathic accuracy, nor vice versa. Without establishing if empathic accuracy deficits are present or not, we run the risk of giving empathy training to those who do not actually need it. In these cases we may be actually facilitating offending behaviour rather than reducing it.

The same argument applies for the assessment of these empathy training programs and the like. Again, dispositional measures seem to be relied upon here, but it is difficult to find literature that looks at the assessment of empathic accuracy or even perspective-taking at the end of these treatments. For example, Carich et al. (2003) offer a comprehensive overview of treatment techniques that are based on empathy and cognitive processes, however there is no mention of empathic accuracy assessment after treatment. Without assessing the different aspect of empathy processing, we cannot tell where these programs are succeeding in their aims or not.

Therefore as can be seen from the research above, empathic accuracy deficits are central to behavioural outcomes and can be used to explain violent behaviours in offenders. Empathic accuracy deficits have also been found in psychologically and developmentally disordered populations, as well as ‘normal’ populations including young children. The fact that rehabilitation programs and other empathy-based treatments may rely only on dispositional empathy measures has been demonstrated as short-sighted. In order to get a clear picture of
the deficits that are occurring we must assess both dispositional empathy and empathic accuracy.

1.5 Measuring Empathic Accuracy

The following review is intended to help inform the methodological approach to be used throughout the rest of the thesis. For this reason there are certain criteria that the chosen methodology must meet, and it is against these criteria that the following critique will be conducted. Firstly the method needs to be empirically sound, with a solid conceptual and theoretical grounding, as well as demonstrating good reliability and validity. Secondly, the measure should be able to assess both an individual’s empathic accuracy ability, and also sensitive enough to detect any biases or deficits that may be present. Thirdly, the method needs to be applicable in a wide variety of testing environments and to a wide range of populations. Lastly, it needs to be cost effective, practical, and easy to administer.

Three methods have been chosen which are used in empathic accuracy research currently: Icke’s interpersonal interaction paradigm; movies, cartoons, vignettes or stories; photographs, and images. Each approach will be described and advantages and disadvantages of each briefly discussed. Please note that this will include studies that refer to perspective-taking or role-taking. The definitional lines between empathic accuracy, perspective-taking and role-taking are often blurred, therefore for the purposes of this review these concepts are included under the empathic accuracy umbrella.

The research perspectives adopted to investigate empathic accuracy are varied depending upon definition. Some studies look at the interpretation of another’s thoughts and feelings as empathic accuracy (e.g.: Clements et al., 2007; Klein & Hodges, 2001; also see Ickes, 1990; 1993) whilst others focus on only one aspect, usually affective empathic accuracy (e.g.: Howland & Rafaeli, 2010; Keltner, Ekman, Gonzaga, & Beer, 2003; Kraus, Cote, & Keltner, 2010). Generally both cognitive recognition empathic accuracy and affect recognition-empathic accuracy (AR-EA) are investigated together, although they are not often
distinguished and are treated as one single construct or process (e.g.: Iannotti, 1985; Gadassi, Mor, & Rafaeli, 2011).

1.5.1 Interpersonal interactions – Icke’s approach.

Early research into empathic accuracy is often based on observations of subjects interacting with each other (see Dymond, 1949; 1950). Although these early forays are criticised for their lack of operationalization of the empathy construct (Bernieri, Zuckerman, Koestner & Rosenthal, 1994) and lack of control of potential confounds (Gage & Cronbach, 1955), the basic premise of interpersonal interaction is basis for Icke’s empathic accuracy research paradigm.

Based upon the original designs of Mehrabian (1971) Ickes devised a method of investigation that could take into account both the specific thoughts and feelings experienced by people when interacting with another (Ickes, Robertson, - et al., 1986). In this approach, two participants are videoed during an interaction. Participants are then be asked to fill in a questionnaire, or to view the video, relating what they were thinking and feeling at the time of the interaction, and what they believed the stranger/other was also thinking and feeling at that time (see Ickes & Barnes, 1977,1978; Ickes, Reidhead, & Patterson, 1986; Ickes et al., 1990). A match between the participant’s reports and responses for what each other was thinking/feeling is taken to indicate empathic accuracy – both cognitive (for thoughts) and affective (for feelings) (see Ickes, et al., 1990 for a detailed explanation of this method).

This method was an important step forward in empathic accuracy research and offers several advantages. Firstly, the interactions between the participants are unsolicited, and not directed in any way, avoiding demand characteristic effects thereby obtaining more ‘naturalistic’ data (Ickes & Barnes, 1977; Ickes, 1993). The other advantage of this approach is that participants use their own words to describe their own thoughts and feelings, and the inferred thoughts and feelings of the other (Ickes, 1993). It also allows empathic accuracy to be simply operationalized, being the degree to which a participant was able to match their
responses of how the other was thinking/feeling, with the other’s actual thoughts and feelings (Ickes, 1993). Unlike many other research designs for empathic accuracy, it involves both real-time and extended-time procedures allowing the process of empathic accuracy to be observed as a factor in the ongoing interaction between two people (Fiske, 1993, Ickes, 1993). Lastly, this approach allows for the assessment of both cognitive and affective empathic accuracy elements, which are normally difficult to capture with any real validity in any other research design.

Using this method, data can be obtained across a variety of interactions, individuals, and specific circumstances, both in terms of their relationship to each other, and in terms of the topic at hand. For instance other research projects include the interaction between empathy accuracy and support for intimate others in a relationship (Verhofstadt, Ickes, Davis & Devoldre, 2008), observing the effects on empathic accuracy of negative relationship thoughts; (Simpson, Orina, & Ickes, 2003) as well as effects of relationship threat (Ickes, Dugosh, Simpson, & Wilson, 2003; Simpson, Ickes, & Blackstone, 1995). This type of research supplies valuable insights into the types of empathic deficits that may occur and how empathic accuracy interacts with our intimate relationships. There are, however, some drawbacks to the method, both in terms of empirical concerns and a lack of practicality.

Firstly, although the method seems highly valid, the uniqueness of each interaction means that results cannot be standardised nor easily generalizable. The way one person communicates their own and other’s thoughts and feelings may be quite different to the next. This means that judgements regarding the coding and matching of responses may not be standard. Although inter-rater statistics are used, the uniqueness of each person’s responses will also require unique coding criteria, which may introduce a level of subjectivity. Therefore, whilst the procedure itself may be standard, the measures (i.e.: the responses) are not.

This method relies heavily upon both the memory and communication abilities of the participants involved. Not everyone is as adept at others at identifying or communicating their
own thoughts and feelings (e.g.: Rubin & Martin, 1994). There are also some circumstances, such as in domestic violence cases, where participants may be reluctant to share their own thoughts and feelings for fear of later retribution from their partner. The process also requires a good level of self-knowledge to enable participants to make responses about their thoughts and feelings. The method presumes that the ability to identify and communicate people’s own thoughts and feelings is similar throughout the population, when in fact this might not be the case.

The concepts of cognitive empathic accuracy and affective empathic accuracy are often blended into the one construct despite the fact that research tells us these two aspects of empathic accuracy may not be related. Ford (1979) found no significant correlations between perceptual, cognitive, and affective perspective-taking, concluding that the three constructs are independent of each other. Therefore it may not be wise to treat cognitive and affective empathic accuracy as the same construct. In fact, it might be more beneficial to separate the two constructs, as their separate contributions to relationships may well be important and insightful.

In terms of practicality the method is necessarily cumbersome. It requires specialist rooms to be set up for filming as well as booths for later viewing. The filming and audio equipment must be of a high quality so as not to undermine the viewing and reviewing procedures. It is time consuming, with participants having to be filmed, their responses recorded, twice, which requires the film to be paused for each response. This means that only a certain number of participants are able to be tested on any given day, at a rate of two per session.

The method also may not be suitable for all couple types, or for all populations, such as violent offenders. This can be seen in some of the research into domestic violence that uses this methodology. In much of this research the terms ‘violence’ and ‘aggression’ are given soft definitions. For instance, Schweinle, et al.’s (2002) study into husband to wife aggression,
recruited men that demonstrated ‘the respondent was indeed in a marriage or cohabitation relationship that was not entirely free of conflict’ (p146). Measures of aggression were usually done by self-report such as the Conflict Tactics Scale-Form A (see Schweinle, Ickes, Rollings & Jacquot, 2010), or the Propensity for Abusiveness Scale (see Schweinle et al., 2002). Whilst the criteria used by Clements et al., (2007) to describe physically violent couples consisted of self-reported evidence of a physically violent act having occurred in the past year, including being - “pushed, grabbed, or shoved” (p374). Many of the participants, necessarily, are from non-offender and/or non-clinical based samples, and with good reason. If a woman who was the victim of severe domestic physical abuse was to participate with her spouse, the interaction itself may become violent and out of control. The spouse may withhold responses, knowing that her husband/partner may not approve of her own thoughts and feelings. And there may be back-lash after testing, when the couple are back in their own home. A violent offender, also, poses some of the same problems. Under these circumstances, safety could not be guaranteed for either participant or offender. The potential harm committed, emotionally and physically, would make this research difficult to justify ethically.

This is the only paradigm, however, that allows for direct observation of empathic accuracy processes, and its interactive nature. It is also one of the few methodologies that incorporates both cognitive and affective empathic accuracy constructs. For this reason Icke’s approach is, and remains an important research paradigm within the field of empathy.

1.5.2 Movies, cartoons, vignettes, and stories.

Similar to the interpersonal interaction methodology of Ickes, the use of movie segments, cartoons/illustrated strips, vignettes and stories attempt to assess a person’s ability to infer the thoughts, feelings and motives of others. This is usually done with the participant adopting the perspective of a character in the story. Scenarios often involve a person in need talking to the ‘audience’ about their current situation, difficulties, choices to be made, or even some general background information (e.g.: Oswald, 1996). Others involve an interaction
between two people discussing a particular topic or problem (e.g.: Klein & Hodges, 2001). The participant views or reads the story, and then infers the thoughts feelings and/or motives of the character portrayed (e.g.: Demurie, et al., 2011; Wakabayashi & Katsumata, 2011; Weiss, Salloum, & Schneider, 1999), or tries to find an appropriate end to the story by picking out the next scene or slide (e.g.: Marton, et al., 2009). A match between the participant’s responses and the character denotes empathic accuracy. Like Ickes’ approach, these methods allow both cognitive and affective empathic accuracy to be assessed, although most studies do not separate the two constructs. They also allow for a variety of social contexts across a range of situations, cognitions, and emotions.

Unlike Ickes’ paradigm, the use of movies and slides can reduce language and communication barriers. These approaches have often been used with child subjects, as they are able to convey different situations without relying on complex and advanced language skills. For instance, Feshbach and Roe (1968) presented a set of slides to children depicting a story about a boy/girl in different emotional situations: happy, sad, afraid and angry. The slides are narrated with simple non-emotive language. After viewing the children are asked what the character in the story felt. If their response matches the character’s emotion, the child is said to be high in empathy (Feschbach & Roe, 1968, see also Feschbach & Feshbach, 1969). A similar approach has been used in other research (Iannotti, 1978; Marton et al., 2009). Others have reduced the need for verbal responses even further by getting children to point to a representation of a face that matched the character in the story (Eisenberg-Berg & Lennon, 1980). Variations on the method involve children labelling pictures of facial emotions expressed by other children, after hearing a story about them (e.g.: Schultz, Izard, & Bear, 2004). The number of items a child answers correctly is added up to give an emotion attribution accuracy score (Schultz et al., 2004).

The technique allows for a wide range of variability in social contexts (Schultz et al., 2004), as well as manipulation of target characteristics such as mood (Oswald, 2002), disorders
such as ADHD (Demurie et al., 2011), emotional situations (Feshbach & Roe, 1968); or the ability to manipulate which characters participants were to take the perspective of (Iannotti, 1978). The stimuli can also be standardised across different testing samples. Reliability for many of these tests has been shown to be good, with levels from .50 for the ACES test (Schultz, Izard, Ackerman & Youngstrom, 2001), to .80 for the method used by Demurie et al. (2011) as reported by Marangoni, Garcia, Ickes and Teng (1995). Participant’s instructions can be varied with participants asked to specifically take the perspective of a character (perspective-taking condition) or told to pay attention to other elements such as the background or the lighting (neutral condition; e.g.: Oswald, 1996). These methods also allow for the observation of only one aspect of empathic accuracy as in Iannotti’s (1978) study of affective empathic accuracy in children.

All of these approaches are easy to administer with only movies requiring any particular equipment, making them cost effective. As has already been demonstrated above, language load can also be adapted to cater for various age groups and reading abilities, with the use of cartoons or illustrated strips being particularly effective for children or other populations with language barriers or difficulties. There are, however, two main drawbacks with the use of these particular approaches.

Firstly, some language barriers may still exist, for all forms of this method. The actor in the movie may talk about their dilemma in English; the vignette might be written in English; and the characters in the comic strip may be speaking to each other in English. This means that participants need to have a relatively strong grasp of the language in which testing is conducted. Thoughts and feelings can be communicated in many different ways. In the English language alone, there are hundreds of words for the various emotions a person may feel. Therefore the method used must take into consideration the language their participants are familiar with and their level of language development.
The other complication is in the form of culture. Display rules, the problems one communicates about, or the social dilemmas one comes across, may vary and have varying significance from culture to culture (Glikson & Erez, 2013). The way one responds to any given situation, a person crying, or a public display of aggression, may be dealt with differently within varying cultures. The tests, then, need to be culture specific, and take into consideration the various social and cultural norms that may be in place for participants. Cultural idiosyncrasies also mean that comparing like research from different cultures may be difficult. Therefore generalisations from culture specific research to other cultures would be difficult.

1.5.3 Photographs and images.

The final method to be reviewed is the use of photographs and images. In these studies participants are presented with photographs or images of people displaying various emotions and asked to identify the emotion being displayed. This is done through either forced choice of emotion-based words (e.g.: Feleky, 1914), or open-ended responses (e.g.: Munn, 1940). Accurate interpretation of the emotions is indicative of high empathic accuracy.

The earliest use of this method was by Darwin in his attempt to ascertain the similarity or dissimilarity of emotional expressions across cultures (Darwin, first published, 1872, 3rd edition, 1998). Other representations consisted of drawings or sketches of human facial expressions, such as Rudolf sketches (Langfeld, 1918A; 1918B) or Piderit profiles (Buzby, 1924, Fernberger, 1927; 1928; Jarden & Fernberger, 1926). There is, however, some criticism levelled at these illustrations. There are some concerns with the images in Darwin’s work where the artist took some (artistic) liberties with the interpretation from photograph to illustration (see Forward by Ekman, Darwin, 1998). The Piderit pictures are in profile, meaning that certain expressions may be confused or misinterpreted because only part of the face is available for observation (O'Sullivan, 1982). For instance, a grimace and a smile may look quite similar in profile. Frois-Wittman (1930) describes the Rudolf sketches (as used by Langfeld, 1918a) as
‘awkward’, observing that distinguishing facial hair such as moustaches and beards in the images obscured some facial features making interpretation difficult.

Some of the first photographs used as stimuli for emotion interpretation ability were from Feleky (1914) and Frois-Wittman (1930) who used themselves as models; however, these early studies were problematic with little statistical analysis and the subjective interpretation of what was deemed an ‘appropriate’ representation of the emotion. For instance, Frois-Wittman (1930) practised various emotions in a mirror. When he deemed that he had captured the emotion sufficiently, he then had an assistant take a photograph of him. Coleman (1949) used various techniques to elicit ‘natural’ expressions from his models, such as surprising them by placing a large snake in their laps. Whilst this might produce highly valid expressions, today’s ethical guidelines would no longer allow such procedures to be used.

Since these early attempts to capture and measure human emotion, there has been considerable development in the types of measures used. Primary among these are the development of photographic-based stimuli. Much of the research in this area has come from the work of Ekman and colleagues, whose aim was to investigate the mechanisms behind people’s identification of other’s facial emotions.

The Brief Affect Recognition Test (BART) was developed as a way of assessing people’s abilities to decode the emotional expressions of others (Ekman & Friesen, 1974b). The slides consist of people facially expressing a range of six different emotions: happiness, sadness, fear, disgust, anger, and surprise. Each image is shown only briefly to the participant (1/100th to 1/25th of a second), whereupon the participant attempts to identify the emotion expressed by the target (Ekman & Friesen, 1974b).

From this Ekman and Friesen (1986) developed the facial affect slides. These photographs were developed for the study of basic emotions and general emotion identification. In an effort to devise a reliable and valid way of assessing emotion interpretation abilities, Ekman and colleagues developed the Facial Affect Scoring Technique,
otherwise known as the Facial Action Coding System (FACS) (Ekman, Friesen & Tomkins, 1971). This is a system based on the actual measurements of facial muscle movements during the expression of emotion. A wide range of models were chosen to pose six different emotions at varying degrees of intensity; however, only pictures that achieved at least a 70% concurrence from the judges for each single emotion were selected. The end product was 10 pictures, each of a different person, selected for happiness, sad, and surprise. Fewer pictures were selected for fear, anger and disgust (Ekman, et al., 1971). The testing used to produce these photographs was extensive and has been shown to have good reliability, especially for spontaneously produced facial expressions (Sayette, Cohn, Wertz, Perrott, & Perrott, 2001). The FAC system has been used to verify new photographs developed for the study of emotion. Tracy, Robins, and Schriber (2009) recently developed a set of photographs of basic and self-conscious emotions (including pride, shame and embarrassment) of African and White males and females using the FAC system to validate the separate emotions.

Many other photographic based measures have been developed, some of which come under the label of emotion interpretation, empathic accuracy, nonverbal decoding and interpersonal sensitivity. Many of these measures incorporate more than one communication mode, usually visual (moving), audio, and still photographs. Both the Profile of Nonverbal Sensitivity (PONS; Rosenthal, Hall, DiMatteo, Rogers, & Archer, 1979) and the Multimodal Emotion Recognition Test (MERT; Banziger, Grandjean & Scherer 2009), use different nonverbal communication channels to assess ability to read nonverbal cues. These include audio and video stimuli, as well as still pictures presented together or separately (the MERT; Banziger et al., 2009) and also focus on different areas of the body such as the face or hands (Banziger, Scherer, Hall, & Rosenthal, 2011; Rosip & Hall, 2004; Rosenthal et al., 1979). The Emotion Recognition Index (ERI) developed by Scherer and Scherer (2011) also incorporates facial expression and vocal emotion recognition elements. In a slight variation of this approach, the Diagnostic Analysis of Nonverbal Accuracy (DANVA) looks at the accuracy of
both sending and receiving of emotional expressions of happiness, anger, fear, and sadness and nonverbal information specifically for children subjects (Nowicki & Carton, 1993; Nowicki & Duke, 1994).

The reliability and validity for these tests seem sound with the DANVA reporting test-retest reliability levels as high as .84 for adults and .74 for young children (Nowicki & Carton, 1993). Reliability for the full PONS and the MiniPONS, as measured by Intraclass Correlation Coefficients were reported as .750 and .566 respectively. Test-retest reliability for these measures were correlated with each other, with individual levels not reported (Banziger et al., 2011). The MERT demonstrated a reliability level of .78 for the total test (Banziger et al., 2009). Scherer and Scherer (2011) argued the veracity of reporting reliability levels for their Emotion Recognition Index due the varying level of difficulty of the items, therefore no reliability was reported.

Other methods for assessing empathic accuracy use only photographs as a stimulus. Many of these are available free for use via internet databases (see Anitha, Venkatesha, & Suryanarayan 2010 for an overview), however, not all have been developed specifically with emotion research in mind. In fact a large number of databases have been designed specifically for gaming and computer simulation programming, in an endeavour to enrich and make more realistic the lives of on-screen avatars (Anitha, et al., 2010; Caridakis et al., 2007). For this reason these photographic databases have not tested the reliability or validity of their emotion expression stimuli. For instance the University of Texas video database contains various images of expressions described as being happiness, fear, anger, puzzlement, laughter and boredom, and yet none of these emotions have been tested or validated (Anitha, et al., 2010).

There are more reliable photographic stimuli available, including Ekman and colleagues’ facial affect slides as mentioned above (Ekman, et al., 1971), and more recently, Matsumoto with the help of Ekman and colleagues developed a set of photographic stimuli
explicitly designed to test emotion recognition abilities in Japanese subjects called the Japanese and Caucasian Facial Expressions of Emotion (JACFEE) (Biehl et al., 1997; Matsumoto, 1992; Matsumoto & Ekman, 1989). This set of stimuli was developed in recognition of the fact that some responding may be culturally influenced (Matsumoto, 1992; Matsumoto & Ekman, 1989). The Japanese and Caucasian Brief Affect Recognition Test (JACBART) was developed as a shorter alternative to the JACFEE, and contains photographs of both American and Japanese models expressing the seven universal emotions (Matsumoto et al., 2000). The photographic set demonstrated reliabilities between .86 and .92 across five different studies (Matsumoto et al., 2000).

Photographs offer a highly reliable and cost effective way of assessing and measuring empathic accuracy, in particular, measures of affect recognition-empathic accuracy (AR-EA). The procedure is simple enough that it can be used in a wide variety of testing environments. It is easy to administer and code/score with only a few specific instructions given to participants for the task. Photographs can be used with a wide variety of populations. Although there may be some culture specific issues, these can be overcome with the use of culture-specific targets in the photographs. A wide range of targets can be used, male and female, young and old. The intensity of the stimulus can also be manipulated, from highly intense emotional expressions to more neutral or ambiguous ones. This means that many different aspects of empathic accuracy can be assessed. These include, but are not limited to, empathic accuracy generally; the ability for individuals to identify subtle emotional expressions compared to more intense expressions; and the assessment and identification of potential empathic accuracy deficits.

There are some difficulties with using photographs. These mainly revolve around the limitations encountered when using other’s photographic material. Firstly most of these are of the head and shoulders of the target only, limiting the information channels available for interpretation. This particular limitation will be discussed in more detail in the next section.
(see section 1.6 below), but simply put there is evidence to suggest that in a social situation the whole body of the target is used by observers to interpret thoughts and feelings, and not just the face (Aviezer, Bentin, Dudarev, & Hassin, 2011; Aviezer, Trope & Todorov, 2012).

Secondly, as Matsumoto (1992) discovered in his research, there are differences in the way people interpret the emotions of faces from cultures other than their own, which can impact results. As a general rule most photographs are of Caucasians, or contain faces only from a single culture, such as the Korean Face Database (KRDB) which is made up only of Korean subjects (Hwang, Byun, Roh, & Lee, 2003). Whilst the Radboud Faces Database (RaFD) was developed specifically as stimuli for emotion studies, and consists of only Caucasian faces of adults and children (Langner, Dotsch, Bijlstra, Wigboldus, Hawk, & van Knippenberg, 2010). The lack of variety of ethnicities or indeed the number of models may limit, again, the research that can be conducted with these materials.

Thirdly, photographs, empirically speaking, are not effective in measuring cognitive empathic accuracy, only affective (recognition-based) empathic accuracy. Via a photograph, one can see the person’s facial expressions, one of the primary modes through which we convey our emotions to others (e.g.: Ekman, 1965). Our thoughts, however, remain hidden. A single photograph, without social context, or a narrative upon which to base inferences, does not communicate enough information to the observer regarding the target’s thoughts and motivations. Therefore photographs are not used for research into cognitive empathic accuracy.

Finally, as is evidenced with the limitations described above, problems arise when using photographs developed by someone else. Using databases or photographs developed for other purposes means that these materials lack the flexibility and controllability necessary for conducting research into specific areas such as affective empathic accuracy (Roesch, Tamarit, Reveret, Grandjean, Sander, & Scherer, 2011). O’Sullivan (1982) makes the point that many measures designed to investigate empathy have not been developed specifically with
the empathic concept in mind, and have oft times not been thoroughly tested for validity and reliability, a sentiment echoed by Scherer and Ekman (1982). So the disadvantage here is that by using other’s photographs, there is no control over the actual stimulus material itself, such as content, lighting, composition, models, etc., meaning that photographic elements of the pictures themselves cannot be controlled nor manipulated.

1.5.4 Summary of previous methodologies.

The methodological approaches reviewed above show the wide variety of techniques used to investigate empathic accuracy, and as already outlined, each has advantages and disadvantages. At the beginning of the section it was stated that the methodology chosen for measuring empathic accuracy would have to meet certain criteria. To reiterate the method chosen should: demonstrate good reliability and validity; be sensitive enough to allow for the detection of empathic accuracy biases as well as measuring empathic accuracy abilities; thirdly must be adaptable to a variety of testing situations and testing populations; and finally be cost effective and easy to administer. Only photographs as stimulus material for the measurement and assessment of AR-EA abilities meet all of these criteria. There is, however, a need to develop an original set of photographs for use in this thesis due to the problems outlined above in using other photographs developed for other uses. The following section will look at how this should be achieved, including which emotions to capture, the mode of presentation (full body versus head and shoulders only), and variations in emotion expression intensities.

1.6 Developing Original Photographic Stimuli

In the above review it was argued that photographs developed by other researchers had some limitations. The first of these is that most photographs only show the target’s head and shoulders, narrowing research to only one communication channel. The second involves the intensity of the emotions expressed in these photographs, many of which show only intense emotions and do not deal with more subtle expressions. There are also other
considerations that should be dealt with when developing your own stimulus materials. These include which emotions to capture and how these emotions should be elicited. The following literature review will deal with these issues, informing the underlying principles for the development of the photographic stimuli, the procedure and testing of which will be presented in chapter two.

1.6.1 Head and shoulders only, body only, or full body?

A review of available photographic databases revealed that most of these portrayed only the head and shoulders of targets (Anitha, et al., 2010; Biehl et al., 1997; Ekman, et al., 1971; Hwang, et al., 2003; Langner et al., 2010; Matsumoto, 1992; Matsumoto & Ekman, 1989; Matsumoto et al., 2000). The focus on the face as the major source of emotion interpretation has been mainly due to the research conducted by Ekman who has stated that the face is the main source of our emotion interpretation information in our daily interactions (1965; 2004; Ekman & Friesen, 1967; Ekman & Friesen, 1974a; Ekman et al., 1980; Kline & Johansen, 1925). Comparing the face and body to investigate which was superior in supplying observers with information regarding the emotion of the target, Ekman (1965) found that both provide quite different affective information. According to Ekman (1965) the head provides information as to the type of emotion being experienced, whilst the body reveals the strength or intensity of that emotion. In a further study of this, Ekman and Friesen (1967) concluded that when shown only the head and shoulders of the target, participants were more accurate in their identification of the emotion expressed, compared to viewing only the body. According to other research the face automatically draws the focus of our attention (Eastwood, Smilek, and Marikle, 2001) and is considered the main source of deception detection in others due to phenomenon such as emotional leakage (Ekman & Friesen, 1969). One of the problems with these studies, however, is that the face and body have been tested separately, and have not been compared with whole-person stimuli.
The isolation of the face from the rest of the body in these studies is an oversimplification of the whole emotion identification process. This was highlighted in Ekman, Friesen, and Ancoli’s (1980) study that compared single communication channels (face, body and speech) with multiple channel combinations (face, body and speech, or face and speech – note that the face and body channel combination was not tested). The results were unclear, with judgement accuracy dependent upon the context in which the behaviour occurred, and the part of the body being assessed.

Research investigating the ability to interpret emotions from body cues only have found that basic emotions can be quite successfully and accurately interpreted (Atkinson, Ditrich, Gemmel, & Young, 2004; Coulson, 2004). Body cues alone can also be used to successfully identify deception in a target (Ekman & Friesen, 1974a). Research has found that body positions (Dael, Mortillaro, & Scherer, 2012) and patterns of body movement are strongly related to particular emotions (Fawver, Amano, Hass, & Janelle, 2012), making it possible to identify emotions through only the use of body cues. The presence of attribution and emotion identification biases can also be found in body poses alone (Atkinson, Heberlein, & Adolphs, 2007; Munoz, 2009). Therefore, using only body cues, participants can accurately identify the emotion experienced by the target and whether the target is being deceptive, as well as researchers being able to identify emotion identification deficits and biases in participants.

Given that accurate emotion identification can be made from either face only or body only cues, it is interesting that studies comparing the whole person (full body) with head only or body only are far and few between. Research has found that only the face is necessary for accurate emotion identification when dynamic (moving) stimuli are used (Nelson & Russell, 2011; App, McIntosh, Reed & Hertenstein, 2011); however when the stimulus is static, as in photographs, findings suggest that both head and body are needed for accurate emotion identification (Nelson & Russell, 2011; Tracy & Robins, 2004). Indeed research has found that when both the face and body are presented together, participants process them as one whole
(Aviezer, et al., 2012) and find it difficult to separate the two, even when directed to (Aviezer, et al., 2011). Therefore the research evidence would strongly suggest that both the face and body (full body) presented together should be used as stimuli in photographs to assess affect recognition-empathic accuracy abilities in participants.

Logically, a whole-body stimulus would also represent stronger construct validity (De Vellis, 2012). In the real world, people are not only presented with the facial expressions of others, but the whole body. In social settings, we will generally see the whole person, and thus using full-body stimuli should represent a more ‘real world’ assessment of empathic abilities.

### 1.6.2 Task difficulty.

In order for any instrument to be effective it should not only assess abilities, but also be able to detect the presence of any deficits or biases in the participant. The aim, therefore, should be to develop a photographic-based set of stimuli that can achieve both of these objectives. In order to do this, the difficulty of the task must be altered in such a way that these biases can show through. There are two ways of potentially achieving this. The first involves making the task of interpretation more difficult by limiting the time participants are exposed to the stimuli. Ekman & Friesen (1974b) displayed black and white images of facial emotion expressions to participants for 5 seconds, finding that accuracy between participants was very high. In another part of the study, this viewing time was limited to 1/100th and 1/25th of a second in the belief that in the social world, emotion recognition and interpretation is done very quickly (Ekman & Friesen, 1974b, p. 221). They found that increasing the difficulty in this way impacted the interpretation of specific emotions, such as fear for depressed persons, and disgust for schizophrenics (Ekman & Friesen, 1974b), demonstrating that deficits may become more apparent for specific emotions when the difficulty level is altered. This, however, would require specialist equipment in the form of computers and screens, limiting the adaptability of the testing instrument.
Another way of achieving the same end is to adjust the subtlety of the emotions being expressed in the photographs. Research consistently shows that biases become more evident in the face of ambiguous stimuli (Burt, Mikolajewski, & Larson, 2009; Dodge, 1980; Dodge et al., 1986; Dodge & Tomlin, 1987; Halligan, Cooper, Healy, & Murray, 2007; Matthews & Norris, 2002; Orobio de Castro, Veerman, Koops, Bosch, & Monshouwer, 2002; Takarangi, Polaschek, Hignett, & Garry, 2008), including the unclear intentions of others (Dodge, 1980; Feldman & Dodge, 1987). This occurs because people with biases, such as hostile attribution bias, use fewer cues to interpret others (Dodge & Newman, 1981; Milich & Dodge, 1984), and inaccurately recognise those cues when they are used (Dodge & Frame, 1982). Therefore, when the stimulus is ambiguous, such as in subtly expressed emotions, interpretation accuracy is reduced and the inherent bias of the individual shows through.

Many studies using photographs have made their final selection of photographic stimuli based on their reliability. As a result only those photographs attracting the highest accuracy ratings are selected, meaning that only the most intensely expressed emotions are included (e.g.: Ekman, 1965). This many increase reliability of the instrument, but does so at the cost of detecting any potential mistakes or biases, or the finer nuances that may be at work during empathy processes. For instance, Hoffman, Kessler, Epple, Rukavina, and Traue (2010) found that there were no differences in emotion recognition between males and females when the expressions were intense, but when assessing more subtly expressed emotions, females were found to be superior. Therefore more nuanced differences between the sexes were only evident with the use of more subtle stimulus materials.

It is surprising then that there is little research into the functioning of AR-EA abilities in participants using both intensely and more subtly expressed emotional stimuli within the same instrument. The ability to express emotions varies from person to person (Dimberg, Andreasson, & Thunberg, 2011; Ekman & Oster, 1982; Hess, Senecal, Kirouac, Herrera, Philippot, & Kleck, 2000) as well as from context to context (Ekman & Friesen, 1974b; Glikson...
& Erez, 2013; Hess, Adams & Kleck, 2004; LaFrance, Hecht, & Paluck, 2003; Santiago-Mendendez & Campbell, 2013). Logically then it would make sense to conduct research into people’s empathic interpretation abilities at varying levels of expression intensity, as this would be the actual task they would need to tackle in the social world. Using stimulus materials that vary the intensity of emotional expressions would then be a far more accurate assessment of actual AR-EA abilities.

Those studies that have endeavoured to vary expression intensity levels have found that less intensely expressed emotions are interpreted less accurately than more intensely expressed emotions (Hoffman, et al., 2010; Montagne, Kessels, De Haan, & Perret, 2007; Palermo & Coltheart, 2004). It has also been found that sex differences become apparent at more subtle levels of expression intensity levels (Hoffman et al., 2010). Beyond this, little has been investigated regarding the more nuanced workings of AR-EA abilities in the face of different expression intensity levels.

1.6.2.1. Facial expression variability as a function of display rules/norms.

In the real world, people do not express each emotion to the same degree. People do not walk around expressing emotions at their most intense levels all the time. Instead, the degree to which we express our emotions, and indeed which emotions we express at all, are governed by the social context in which we find ourselves, in what Ekman referred to as display rules (Ekman & Friesen, 1969, as cited in Kupperbusch, 1999, also see Ekman, 1999b).

Ekman and colleagues (1969 as cited in Kuppberbusch, 1999; 1999b) proposed that emotions were displayed according to cultural rules of appropriateness as well as the learned behavioural consequences to the display of certain emotions. Expression intensity can be influenced by different social situations, which will also influence whether an emotion is expressed at all. For instance, one might be comfortable displaying an intense level of grief with a close friend in the confines of one’s own home (Fridlund, 1991), but will be less likely to express that same grief so overtly in a restaurant in the presence of strangers (e.g.: Jakobs,
Manstead, & Fischer, 2001). Obviously, the more intensely an emotion is expressed the more easily it will be interpreted. When we suppress or mask these emotions, they become more difficult to identify.

Ekman’s work on facial emotions, although focusing on universal emotions, acknowledged that the expressivity of emotions varied between cultures (Ekman, Sorenson, & Friesen, 1969). Glikson and Erez (2013) looked at the different display rules in teams of multicultural and culturally homogenous members. In teams made up of participants from different cultures, it was more acceptable to express positive emotions and suppress negative emotions compared to teams that were culturally homogenous. ‘National identity’ was found to influence display norms more strongly in the homogenous teams than the multicultural (Glikson & Erez, 2013), implying that within each culture different display norms may be at work.

Variations in the displays of emotions are not limited to cultural influences. Gender roles also impact displays of emotion. For instance Santiago-Menendez and Campbell (2013) found that girls admitted to crying more often than boys when sad, and also that only girls, and not boys, admitted to crying when angry. Differences in expression can also be seen in affective disorder samples, both between different diagnoses and between individuals within each diagnostic group (Ekman, Matsumoto, & Friesen, 1997).

One of the biggest influences on our displays of emotions comes from whether others are present to witness these displays or whether we are alone. Fridlund (1991) investigated the effects of an audience on smiling. Participants viewed a video under different social conditions that included the participant being alone, with a friend, a friend in the room (but not watching the video), or with the belief that the friend was viewing the same video in another room. The study found that participants smile more when the friend is in the same room, compared to when participants are alone. Therefore smiling was predicated more upon the social context of the situation rather than any emotion cues. Jakobs, et al. (2001) looked
at negative emotional displays rather than positive ones. Participants viewed a sad eliciting video either alone, with another person present, or another person present in another room. They found that negative emotions, in this case sadness, were expressed less frequently in the presence, and believed presence, of others, compared to when the participants were alone. Also sad expressions were observed more frequently when the ‘other’ was a friend compared to a stranger. Unexpectedly, however they also found that participants smiled more in the ‘other present’ conditions, specifically in the presence of a friend compared to a stranger.

There are some questions regarding significance levels in this study, as the authors reported significant findings at the $p = .10$ level (Jakobs et al., 2001). Despite this, the study does raise some interesting questions about the purpose of facial displays when in the presence of others. It would seem that the ‘smile’ in these contexts, can convey many different meanings when the shared experience is negative (sad) (Jakobs et al., 2001).

The social context and our relationship with anyone present can also influence what emotions we display. Hess, Banse, and Kappas (1995) looked at the different factors that might influence emotion intensity in expressers. In this study they varied the social context of the situation, the intensity of the emotionally eliciting stimuli, as well as the relationship between the target and the observer. They found that one of these factors alone was not enough to predict the intensity of expressions, but instead concluded that the extent to which we express emotions relies on an interaction between all three of these factors. Fischer, Becker, and Veenstra (2012) also found that these factors influenced the extent to which participants engaged in unconscious mimicry of others, regarded as another form of empathy.

An interesting study by Mendolia (2002) found that under stimuli of self-threatening emotional events, participants actively suppressed their facial emotional expressivity. Facial muscle and autonomic activity were reduced under these conditions when compared to control conditions (Mendolia, 2002). This finding has some relevant ramifications. If under threat a person will repress their emotions, then the provocateur will a) not be able to read
the response of the other to their actions and b) in the case where the threatening behaviour
is instrumental, may even be unsatisfied with such a low level of emotional response, and
therefore increase the intensity of their subsequent actions.

The above evidence indicates that many different factors influence our display of
emotions. The cultural norms and socialised standards of emotional display and behaviour
play a substantial role in our everyday expression making, in what Ekman referred to as display
rules. These rules seem to be implicit for different social contexts and even whether another
person is present or not. Our relationship to that person, whether friend or stranger will also
have an impact not only on the strength of our emotional displays, but the kinds of emotions
displayed as well. Finally, we will endeavour to control the expression of our own emotions
when we are trying to protect ourselves in the face of threatening stimuli. Therefore the study
of empathic accuracy, the ability to interpret another’s emotional state, needs to acknowledge
that there are influences that might impact the types of emotion expressed and the intensity
of those expressed emotions, and that these will, in turn, impact the ability to accurately
identify those emotions.

1.6.3 Which emotions to capture.

Whilst the human face can express a number of different emotions not all of these
may be readily discernible to the same degree by everyone. Therefore which emotions are
expressed in a photographic-based measure of empathic accuracy becomes important. The
following represents an argument for the use of what Ekman called basic emotions (1992a,
1992b, 1999a). These represent a range of emotions that are the most readily recognisable
across different cultures, and therefore are considered the best emotion constants to be used
in empathic accuracy study. There are several criteria that these emotional expressions should
meet.

Firstly that the emotions expressed can be identified by the majority of people – what
Ekman referred to as universal emotions (Ekman & Friesen, 1986; Ekman et al., 1987). Darwin
was one of the first to ask if human emotional expression was the same for particular emotions regardless of culture (1872/1998). If this was found to be the case, then emotions, Darwin argued, were the result of biological evolution and were not solely sourced from individual cultures. Darwin used photographs and drawings of natives expressing various emotions and showed them to others to see if the emotions expressed could be readily identified (1872/1998). He concluded that many emotions were recognisable across cultures (1872/1998). Over many years of research, the evidence for a specific set of universally recognisable emotions has grown (Ekman, Friesen, & Ellsworth, 1972; Ekman et al., 1987; Ekman, et al., 1969). Reviewing cross-cultural studies Ekman (1999a) concluded that there was evidence for fear/surprise, anger, happiness, disgust and sadness. Ekman also included contempt arguing that this emotion too was recognised by many different cultures (1999a).

Recent research by Barrett (2006) has argued that there are no universally recognised emotions at all. Barrett proposed that facial expressions of emotions were not necessarily true representations of what people were actually feeling, but only represented notions of those emotions which individuals related to specific experiences. Barrett conducted research looking at how people interpreted the emotions of others, reporting that participants had difficulty distinguishing depression and fear. One obvious problem is that it is debatable whether ‘depression’ could be classed as an emotion, being rather an experience that encompasses mental, physical and emotional elements. Also, whereas fear is considered a single emotion (or basic emotion, see discussion below), depression is arguably a composite of many different emotions including anxiety, fear, anger, sadness and lethargy. Although sadness is a universal emotion and could be used to compare to fear, another universal emotion, it is unclear why Barrett has not done this. When looking at a complex issue such as emotion, it is important that one is comparing apples with apples.

The research on basic emotions has come under much scrutiny and criticism, with debate ongoing. Barrett (2006) does point out that not all people express their emotions at
the same levels as others, nor do they express their emotions to the same degree in every social situation. Ekman points out that display rules, those rules that culturally and socially govern the appropriateness of expressing certain emotions in specific social settings, and to what intensity, plays a large role in our everyday emotional expression experiences and our ability to interpret these expressions in others (Ekman & Friesen, 1974b). The use of varying intensities of emotional expressions as stimuli will, therefore, help shed light on this debate.

Secondly basic emotions should be singular in nature. That is that they are discreet emotions and not a result of several different emotions blended together such as you might find with anguish or suffering (Ekman, Friese, & Ellsworth, 1982). Blended emotions are problematic as different people might be interpreting different aspects of emotions, or recognise one emotion over another in the same expression. Therefore, it is important that the emotions to be used as stimuli be singular and discreet. Ekman (1992a) proposed the existence of basic emotions, those that were expressions of one pure emotion, and identified seven emotions that fit this criterion: happiness, surprise, fear, anger, sadness, disgust/contempt and interest (Ekman 1992a; 1992b; 1999a). In spite of different methodologies and theoretical approaches the same seven emotion categories seem to be common to most emotion research (Ekman et al., 1982; Izard, 1984; Munn, 1940; Osgood, 1966; Plutchick, 1962; Tomkins & McCarter, 1964; Woodworth & Scholsberg, 1955) although Ekman later found little support for the emotion ‘interest’ as a basic emotion (1999a), leaving six basic emotions.

There is still some debate around which emotions should be considered basic. Jack, Garrod, and Schyns (2014) argue that there are only four, not six basic emotions. This view is based on the argument that some basic emotions actually share similar facial features in their expressions. Jack et al. (2014) found that anger/disgust and fear/surprised shared many of the same facial elements in their expressions, and served to communicate information in terms of fight or flight responses to different types of danger. They go on to argue that the distinction
between these sets of emotions (anger/disgust and fear/surprised) are only made by participants when the expressions are more intense, suggesting that these emotions are less discernible to participants when more subtly expressed. More study needs to be done to ascertain if this finding holds across other research, as was pointed out earlier, very little has been done into research of varying levels of emotional expression intensity.

According to Ekman’s research, universal emotions and basic emotions encompass the same emotion categories. This is not surprising as singular emotional expressions are much easier to successfully interpret than expressions composed of blended emotions (Ekman, 1992a). This is important as singular emotional expressions should provide a degree of reliability for any photographic–based measure in that participants will be able to successfully and consistently identify the emotions being expressed in them.

1.6.3.1. Posed versus spontaneous emotional expressions.

The development of any testing instrument requires that at each point the most valid and reliable methods have been employed (O’Sullivan, 1982). One of the issues to be considered is whether the emotions to be captured should be spontaneous or posed. The following review will demonstrate that posed expressions are superior to spontaneous expressions for several reasons.

Spontaneous expressions are those captured in the moment. Whilst more valid in many respects, they are less valid in others. Firstly, from a practical standpoint it is difficult to capture purely spontaneous emotional expressions, as to do so would require participants with no knowledge, initially, that their image was being captured. If they were to know that their image was being captured, then, for intents and purposes, these images would be the same as those acquired from posed expressions. Secondly, apart from the ethical issues that this then raises, it would be difficult to ascertain the purity of the emotion being expressed. Many of our emotions are actually a blend of several emotions (Ekman et al., 1982). Without asking the participant before-hand, or during the expression of the emotion, it would be
impossible to ascertain the specific emotions being expressed. This would mean a high possibility of blended emotions. Lastly, some sort of accuracy criterion needs to be established to ensure which specific emotions are being captured. That is, that the person expressing the emotion needs to confirm which emotion they were experiencing. This is then used as the ‘label’ of that emotion, and is the criterion against which participants’ accuracy would be later determined.

Posed emotions, on the surface, seem less desirable than spontaneously expressed emotions, as posed emotions may be seen as fake or unrealistic. This speaks to the issue of validity. Research has demonstrated, however, that for the purposes of an emotion identification task, there seems to be little difference in accuracy between spontaneous and posed expressions (Buck, Reuben, Goodman, & Shapiro, 1980; Motley & Camden, 1988; Sayette et al., 2001: Tucker & Riggio, 1988; Wagner, MacDonal, & Manstead, 1986; Zuckerman, Hall, DeFrank, & Rosenthal, 1976). It has also been found that a person’s ability to successfully communicate an emotion spontaneously is positively related to their ability to also pose those same emotions (Riggio, Widaman, & Friedman, 1985; Tucker & Riggio, 1988). Ekman et al. (1982) make the argument that posed facial expressions are actually very closely related to their spontaneous versions. Therefore it seems that, in terms of testing purposes, there is little difference between the two. The advantage for posed emotions is that specific, and basic, emotions can be requested and posed by models, as opposed to the danger of capturing blended emotions that would be possible in a more ‘natural’ setting. Posed emotions allow for a strict accuracy criterion to be set and tested. Potential models can not only record exactly which emotion they were expressing at the time, they can also record how well they believed they captured that emotion.

1.6.3.2. Elicitation techniques and guidelines.

Studies have used many different techniques to elicit emotional expressions from their participants. Coleman (1949) subjected participants to different types of stimuli, including a
live snake, to elicit the desired emotions. Ethical considerations aside, it has been acknowledged that eliciting specific facial emotions can be problematic (O’Sullivan, 1982). One problem is ensuring that only one emotion is expressed at a time. As mentioned above, blended emotions, combinations of several emotions, are difficult to interpret compared to pure singular emotions such as happy or sad. The problem is in ensuring that the emotion being expressed by the model is in fact only that required emotion.

There are two main approaches recommended by Ekman et al. (1982). The first involves the models themselves drawing upon personal experiences, memories, or events, to re-experience that specific emotion, and therefore express it. The second option is to precisely direct models to move and hold their facial muscles in particular combinations. For instance, for surprise, the eyebrows would be raised, the mouth slightly open and the eyes opened wide (Ekman, 2007). The first technique is preferable to the second as it cannot be guaranteed that all the models chosen would have the amount of control needed over their own facial muscles to adequately form the required emotion. Having models draw upon their own personal resources, with the aid of scripts or scenarios, in combination with the accuracy criterion of reporting what emotion they captured and how well they did so, should result in highly valid and realistic emotional facial expressions.

One other consideration with this type of stimuli is the danger of emotions being blended across categories during the photographic session. Expressing one emotion after another may lead to emotional bleeding of one emotion into the next (Ekman et al., 1982; O’Sullivan, 1982). This has been found in previous research (Davis, 1934). Ekman and others recommend ‘cleansing exercises’ to combat this. Cleansing exercises are designed to eradicate any residual emotion the model may be feeling, allowing them to tackle the next emotion without any trace of the previous emotion (O’Sullivan, 19982). Most of these exercises incorporate tasks to engage cognitive rather than emotive processes (O’Sullivan, 1982). The exercises should be conducted between each emotion set to ensure that each emotion is
1.7 Manipulation of Photographic Elements

It has already been established that, despite the amount of research into emotion interpretation that use photographs, no studies have looked specifically at how photographs might be better used to promote empathy processes, and therefore investigate more closely the empathic accuracy construct. Neither has research delved deeply into the effect social context may have on empathy processes, especially not specifically using photographs to manipulate the setting. The following section looks at how photographs might be used as an effective means of investigating both of these issues.

1.7.1 Enhancing imagination and connectivity.

As stated in the introduction of this chapter, there has been little research into how the imagination component of empathy might be enhanced, beyond simple participants instructions of ‘imagine self/imagine other”, nor how connectivity between target and participant can be enhanced when using photographic stimuli. Photographs provide an opportunity to easily manipulate specific elements of the image, which may impact AR-EA abilities via imagination and connectivity.

According to many researchers, the empathy process requires a strong imagination component (Davis, 1980; Eisenberg & Miller, 1987; Hoffman, 1984). During empathy, the observer takes the perspective of the target, appreciating the specific situation or circumstance from their point of view (Hoffman, 1987). In this way, it is proposed, the observer can come to understand the thoughts, motivations and emotions being experienced by the target (de Vignemont & Singer, 2006; Hoffman, 1987; Ickes, 1993).

Many studies have manipulated the imagination processes of empathy by having participants either imagine themselves in the target’s situation, or by imagining how the
situation from the target’s perspective (see Batson, Lishner, Carpenter, et al., 2003; Davis, Conklin, Smith, & Luce, 1996; Davis et al., 2004; and Lamm, et al., 2007 for some examples). In a study that looked at adequacy of these imagination instructions to participants, Davis et al. (2004) found participants who received no specific instructions regarding the target except to ‘watch the interview’ demonstrated other-related thoughts similar to those participants who had been specifically instructed to imagine the situation from the target’s point of view. Davis et al. (2004) concluded that imagining situations from another’s point of view was actually a natural response, and so imagine-other instructions may be a little superfluous.

These types of studies often use photographs as their stimulus materials, however the photographs themselves have not been used as a means to manipulate or facilitate imagination processes. Beyond participant instructions of imagine-self/imagine-other, or variations of this, no other techniques have been used to manipulate imagination. This is a surprising oversight considering the important role imagination processes are theorised to play during empathy.

Empathy studies have also revealed that the connection between the target and the observer is also important. Two mechanisms seem to be at play with this. Firstly is the level of familiarity or similarity the observer has with the target. The more familiar we are with the target and/or their situation, the more that empathic-related behaviours (including helping behaviours) tend to increase (Krebs, 1975; Davis, 1996; Zagefka, Noor, & Brown, 2013). Secondly, participants are more likely to engage in perspective-taking based behaviours (which increases empathy; Coke, et al., 1978; Toi & Batson, 1982), if they value the welfare of the target (Batson, et al., 1995; Batson et al., 2007). What is more, the more similar the target to ourselves, the more we will value their welfare, resulting in an increase in empathy (Batson et al., 1995).

Accurate interpretation of another’s state is also enhanced with increased familiarity. Stinson and Ickes (1992) found that male friends displayed higher empathic accuracy for each
other than for male strangers. One of the best predictors for empathic accuracy is the amount of interaction between strangers. With increased interaction comes increased empathic accuracy (Marangoni, et al., 1995; Simpson, et al., 1995).

Taken all together, the research indicates the better the connectivity between target and observer, be it through familiarity, interaction, similarity or increasing their value, the more that empathy processes are facilitated. The process of interpreting the emotional expression of a stranger in a photograph, however, is not innately conducive to these mechanisms.

There appears to be no research on how, specifically, both imagination processes and connectivity might be enhanced through the use of photographic-stimuli. This thesis will propose that a simple photographic element (a silhouette) may be able to accomplish these aims. The chapter outlines at the end of this chapter will go into more detail regarding this particular mechanism and the rationale behind its construction.

1.7.2 The influence of social context on empathic accuracy.

1.7.2.1. Background cues.

Empathy is not only an interpersonal process, but is also contextual in nature. The process of interpreting another person’s emotions occurs in relation to many cues, not just the emotional cues of the target. These include the situation, the event, and the environment in which this has occurred. Not only this, but the types of emotions expressed and the degree to which they are expressed are also reliant upon the social situation (see section 1.6.2 above). Therefore our interpretation of others will necessarily include processing of social cues.

Silverthorne, Gibson, Micklewright, and O’Connell (1975) tested the idea that our interpretations of others could be influenced by the colour of the background behind the target. Male head and shoulders pictures portrayed against different coloured backgrounds (blue, green, red, yellow, and white) were shown to participants. Yellow and white
backgrounds produced the most positive reactions to the faces; green and blue produced both positive and negative responses; whilst red elicited the most negative responses. Young, Elliot, Feltman, and Ambady (2013) found that red coloured backgrounds facilitated interpretation of angry faces. Therefore it is not only the physical cues of the target that we read when we are interpreting others, but also contextual cues such as the environmental background.

In a study by Marian and Shimamura (2011) the positioning of the characters within a simple drawing heavily influenced the interpretation of emotions felt by those characters. The drawing was ‘Terror Subterra’ by Shephard, and depicts two figures running down a dark tunnel. The background figure is larger than the foreground figure, but otherwise both figures are exactly the same, with exactly the same facial ‘expression’. Participants often identified the background figure as being angry and the foreground figure as afraid. Therefore judgements about the emotions experienced by the characters were dependent upon context (Marian & Shimamura, 2011).

Some research suggests that contextual information may be even more important than facial cues alone. Walbott (1988) showed film clips to participants, which was followed by an emotional facial expression of the actor responding to the event in the clip. Clip presentation varied with some only showing the event, others only the facial expression of the actor, and some clips showing both the event and the emotion expression in reaction. Results indicated that the contextual information of the event was more of a key factor in identifying the emotion of the actor than the information portrayed by the face alone (Walbott, 1988). Other research has found that biases, such as hostile attribution biases, may be influenced by social context (Goldeski, Ostrov, Houston, & Schlienz), which may also impact accurate emotion interpretations in others (see section 1.4.1.4 above).

There has been very little research carried out in this area, and for some of these the quality of the research is questionable. Russell and Fehr (1987) made the argument that interpretation of facial emotions is relative, depending upon what emotion cues have been
previously presented. In this study participants were presented with an ‘anchor’ face, which displayed an emotion. This was then followed by the target face which participants then labelled with an emotion. The study came under heavy criticism from Ekman and O’Sullivan (1988) stating that the target faces used were neutral and ambiguous, forcing participants to rely on other cues. In reality this study was probably closer to investigating the effects of emotional priming, demonstrating only that the interpretation of ambiguous emotional cues are vulnerable to biases and influence from other stimuli.

Carroll and Russell (1996) also looked at the effects of context on emotion interpretation. Participants were presented with different social scenarios, and then presented with a face expressing an emotion. Their results showed that social and situational cues, and not the facial cue, were the primary source for interpreting the emotion; however, there are many problems with this study.

The authors seemed to confound the factors of social context and emotion. The social scenarios presented to participants were emotionally primed. For instance, waiting overly long for a table at an expensive restaurant, when you had made the reservation weeks ago, watching others being seated (without a reservation) before you. Most people reading this scenario would interpret the characters in the story as being frustrated, bitter, angry, etc. Therefore the participants have already been ‘emotionally primed’ before looking at the stimulus faces. It seems that this is confounding social context with primed emotions. The social situations presented here were not emotionally neutral, which impacted the interpretation of the emotions portrayed by the target faces.

The expressions displayed by these faces were also problematic. The faces presented to participants after reading the stories were often very similar to the emotion elicited in the scenario just read. One scenario was designed to elicit ‘determination’, whilst the stimulus face that followed expressed ‘anger’. These two expressions have very similar features including a pressing together of the lips, and a slight downturn of the eyebrows (Ekman, 2007;
Ekman, Friesen, & Ellsworth, 1972). According to Darwin the act of frowning is often accompanied by effortful reflection, as well as feelings of determination and anger (1872). Ekman (2007) points out that the facial movements that happen with slight anger are also very similar to those of other emotions such as determination. Therefore the selection of the emotion elicited and the facial expressions presented, potentially confound each other in Carroll and Russell’s (1996) study, calling into question both methodology and findings.

In conclusion, although some of the studies are questionable, the limited research available shows that the interpretation of emotions is not done in the vacuum of the target’s face alone. The environment in which the observation is made can also influence judgements, whilst social context and the presence of other emotion cues can influence accurate emotion identification in others. Empathic accuracy, then, involves the interpretation of more than just facial cues, but is influenced by a number of different contextual factors.

1.7.2.2. Social context and cognitive processes.

Hoffman (1984; 1987) proposed that higher order empathy processes involved cognitive functions that included memory and experience. That is to say that our own experiences can colour how much empathy we may feel for another person’s situation. Likewise, the interpretations we make regarding that experiences of the other person is driven by contextual cues, as well as the emotion cues, of the target.

Wyer and Srull (1986) presented a model of cognitive processes that took social context into account. In this model, social cognition (input) is influenced by the observer’s prior knowledge and their goal-driven motives in that particular social context. The resultant behaviour (output) takes into account the social judgements made about the target, the emotional responses to those judgements, the situation, and memory (Wyer & Srull, 1986). In the case of affect recognition – empathic accuracy abilities, interpretations of others emotions may be influenced by any or all of these processes. Gesn and Ickes (1999) found that
cumulative knowledge of the context and target facilitated empathic accuracy abilities in participants, highlighting the importance of these other contextual cues.

Research has demonstrated that the motives of others can influence both emotion recognition and our reactions to those emotions. De Melo, Carnevale, Read, and Gratch (2013) found that the same emotional display could elicit different responses from participants depending upon the perceived motive of the target. When the participant was considered to be competitive, smiles from the target were interpreted negatively; however, when the target was deemed to be cooperative, their smiles were interpreted more positively. The author’s concluded that appraisals about other’s motives influenced the effect emotional displays had on participants, which in turn would influence behavioural decisions (De Melo, et al., 2013).

Interpretations of others are also predicated upon the cognitive schemas or scripts that we use to process social information. A study by Gesn and Ickes (1999) found that cognitive schemas were in operation during the interpretation process which at times facilitated, and at other times actually impaired, participants’ empathic accuracy. This finding suggests that social schemas regarding events and social context can influence interpretation abilities, although more research is needed in this area to fully understand the implications.

Children, as they grow older, take more social cues into account when identifying the facial emotions of others (Reichenbach & Masters, 1983). Older adults too, demonstrate better accuracy for interpreting other’s emotions when they have more social contextual information, compared to just the facial cues of the target (Sze, Goodkind, Gyurak, & Levenson, 2012). Such contextual information may include the sex of the target, which has been demonstrated to influence emotion identification (Vrana & Gross, 2004), as well as more simple processes such as if the target is looking directly at the observer or not (Mumenthaler & Sander, 2012).

Munn (1940) is one of the few studies that looked at the differences in empathic accuracy when social (background) information is supplied, compared to head and shoulders
only stimuli. In this early study, Munn chose photographs from magazines and then presented these pictures either intact (with the backgrounds left in) or with only the head and shoulders of targets showing (background left out). Munn (1940) found that leaving the social context in the photograph increased the accuracy with which participants were able to interpret the emotion being expressed by the target. Although this study does have problems in regards to the accuracy criterion set for the emotions expressed by the targets, it does enhance the point that other’s emotions are not interpreted in isolation, but are done in a wider context. Therefore knowledge about the social situation, memory, cues from the other person (both non-emotional and emotional), can all influence how we read another’s emotion.

1.8 Study Aims and Chapter Outlines

The above literature review demonstrates the importance that mechanisms such as imagination, connectivity, and social context can have on empathy processes. By imagining a situation from another’s perspective, empathy is achieved. Connection between the observer and target is also important, as we are more likely to engage in empathy processes if we feel connected to the target in some way. Social context influences our emotional expressions, changing the types of expression and/or intensity with which they are expressed dependent upon the social arena we are in at the time, and therefore impacting accurate emotion identification.

Although very little research has been conducted to investigate these elements individually, their importance in empathy processes can be seen in the instructions given to participants (imagine self/imagine other) as well as the methodologies used in empathy studies. A prime example of this is Icke’s paradigm that uses direct interpersonal interaction (connectivity), imagination (what was the other person thinking, feeling) and context (what was the other person thinking/feeling at that point in time) to investigate empathy processes. It was demonstrated, however, that photographs represent the most cost effective and adaptable method of assessing AR-EA abilities. Photographs can be manipulated to include
imagination/connectivity and social context elements within them for very little cost and effort; however, in order to do this, original photographs need to be developed.

The aims of this thesis, therefore, are twofold. Firstly to investigate the potential impact that imagination, connectivity, and social context, might have upon AR-EA abilities in participants. It is proposed that both imagination and connectivity can be enhanced in photographic stimuli through the use of a simple element – namely a blacked out silhouette. Participants would imagine that they are the silhouette in the photograph with the target. It is hoped that this simple mechanism will facilitate both imagination and connection, thereby enhancing empathy processes. Social context can be manipulated by inserting different social backgrounds into the photographic stimuli in order to investigate the effect this might have on participant’s AR-EA abilities.

In order to achieve this first aim, an original set of photographic stimuli will need to be developed and tested. This is the second aim of the thesis. Development of new stimuli for use in a measure of AR-EA abilities will allow for standardisation of images, as well as ensure that manipulation of the images can occur at a high standard, maintaining the integrity (resolution, etc.) of the original photographs. This will involve developing methodology and procedures for the taking of the photographs. These photographs then need to be tested for accurate interpretation by participants, as well as for their validity and reliability. To achieve this, some of this testing will be done concurrently with the studies into imagination, connectivity, and social context. Some chapters, therefore, will contain two sections: one section containing the study for the main inquiry (imagination and connectivity; social context), and then a second section dedicated to the psychometric evaluation of the photographs. In this way not only are the issues of imagination and connectivity, and social context, addressed, but a new measure for the assessment of AR-EA abilities will have also been developed. The following outlines the aim and content of each of the chapters in the
thesis, beginning with the taking of the original photographs for the AR-EA photographic measure.

Chapter two will begin with outlining the specific emotions and types of expressions (spontaneous versus posed) to be included in the photographs. Elicitation techniques will also be briefly outlined followed by the procedure used for taking the photographs. The second part of the chapter is divided into two studies testing the accuracy with which participants were able to interpret the emotions expressed in the photographs. The first study involved an open-ended response methodology for testing. The method and results for this are presented followed by a short discussion. The open-ended responses proved to be problematic, which negatively impacted accuracy ratings, therefore a second study was conducted using a forced-choice, sorting-task methodology. The method and results of this second study are presented, followed by a discussion. The chapter finishes with a comparison of the results of Study 1 and Study 2. From these, the final photographs to be used for testing as the AR-EA photographic measure were selected. This resulted in 40 black and white full body photographs that were used in testing in the following chapter.

The focus of chapter 3 was to investigate the possibility of inserting a photographic element (a silhouette figure) into the photographs in order to facilitate imagination and connectivity. The aim was to see if this mechanism would influence participant AR-EA abilities. In order to assess this, three different photographic presentation modes, were compared: head and shoulders only (H&S) which is the most common presentation mode to be found in empathy research; full body only (FB); and full body with silhouette (FBSil). The full body modes were introduced based on research (reviewed in section 1.6.1 above) demonstrating that when present, participants will use the full body and not just the face of a target to interpret emotions. The method and results from this study are presented, followed by a brief discussion. The second part of this chapter involved validity and reliability testing of the actual photographic measure, conducted as part of the ongoing development of the AR-EA
photographic measure. Procedure and results are presented, along with a brief discussion. From this testing it was concluded that the full body photographs containing the silhouettes should be included as the AR-EA photographic measure, to be used as a part of the study into social context conducted in the following chapter.

The first study in chapter 4 was designed to investigate the potential impact of social context on AR-EA abilities. Using the results from Chapter 3, the most reliable presentation modes, being full body (FB) and full body with silhouette (FBSil) were doctored to include different social backgrounds. These included a bar, a kitchen and a neutral (blank) background. The three different backgrounds were tested against each other to investigate if AR-EA abilities overall and/or for specific emotion categories were impacted. Procedure and results are presented followed by a discussion and explanation of the findings. A second study in chapter 4 presents the method and results for further reliability testing of the AR-EA photographic measure. Results from chapter 3 were marginal and so a different methodology was engaged in the form of a test-rest study, to confirm the original reliability testing results. Procedure and results are presented, along with a brief discussion of the findings.

The final chapter (5) presents a general discussion and summary of results for the studies conducted throughout the thesis. Firstly the findings from the studies investigating the impacts of imagination, connectivity, and social context, on AR-EA abilities in participants are discussed. This will includes a comparison of the different photographic modes (head and shoulders only (H&S); full body only (FB); and full body with silhouette (FBSil)), summarising the results and producing suggestions for further testing and development of the measure. Following this a discussion bringing together the findings of the impact of imagination and connectivity, and social context, upon AR-EA abilities will be presented. Secondly, the results of reliability and validity testing of the AR-EA photographic measure will be discussed. This will include a brief discussion of the difficulties in developing a measure that includes varying task difficulty levels, and the impact this can have on empirical testing. This will be followed by an
acknowledgement of limitations of the studies in the thesis, ending with an overall conclusion and discussion of future directions for study.
Chapter 2

2.1 Overview

The focus of chapter two was the development of original photographic stimuli designed to assess affect recognition – empathic accuracy (AR-EA) abilities in participants. To that end the chapter has three main sections. The first involves the procedure for the taking of the photographs, the methods used for elicitation and a short discussion regarding the outcome. The second section involves the initial testing of the photographs to make sure that participants are able to accurately identify the emotions expressed by the models. The testing occurred over two studies, the results of which were used to determine which photographs were selected for the AR-EA photographic measure to be used in the studies in the chapters to follow. The final section presents the procedure used for selecting these photographs with a brief discussion of the outcomes. To summarise, this chapter will deal with the taking of the photographs, initial testing of accuracy ratings, and selection of photographs to be used in further studies.

As outlined in chapter one, there are five issues that need to be considered when regarding the composition of the photographs. Firstly, which emotions these photographs should capture; secondly whether these emotional expressions should be posed or spontaneous; thirdly, if posed, which elicitation techniques should be used; fourthly how many levels of expression intensity should be included (only the most intense expressions or more subtle versions); and finally, the accuracy criterion to be used. This chapter will outline the processes adopted in this study with reference to the literature to determine the content and structure of the photographs.

The main portion of this chapter will focus on the initial testing of the photographs, using two different methodologies. This is to determine whether participants are able to successfully interpret the intended emotions expressed by the photographic models. Based on
these findings, decisions about which photographs should be selected for the next series of testing were made. This process will be briefly discussed at the end of the current chapter.

### 2.2 Composition of the Photographs

**2.2.1 Presentation modes: full body versus head and shoulders only.**

The use of photographic stimuli to assess AR-EA type abilities is not new, however, there are potential problems with some of these measures. As discussed in chapter 1, many existing photographic measures of empathic accuracy use only the head and shoulders of participants as the stimuli (see chapter 1, section 1.6.1). The reason for focusing on the face as a stimulus stems from research that has found that emotions are more accurately identified using the face only, compared to the body only (Ekman & Friesen, 1967), that facial emotions act to draw attention to the face (Eastwood, et al., 2001), and the information rich properties of the face for communicating social information such as our emotions (Ekman, 2004; 2007). Therefore much of the research has relied upon the assumption that the face is superior to the body in communicating affective information, and that the body has no role to play in emotion interpretation processes; however, this is not necessarily the case.

Firstly, in judgements of the whole person, Ekman, Friesen, O’Sullivan and Scherer (1980) found little support for the facial channel being superior to either the body or speech channels, when compared to judgements using the whole person. Instead it was found that no one channel elicited more accurate judgements than any other.

Secondly, the body has been found to be information rich, not poor, in regards to emotion. The body can provide information about the intensity of emotion (Ekman, 1965), and even reveal which specific emotions are being experienced (Atkinson, et al., 2004; Coulson, 2004; Dael, et al., 2012; Fawver, et al., 2012). For instance, Coulson (2004) found that participants were able to accurately the emotions sadness and anger using only static body cues. Interpretation biases have also been elicited from body only cues (Munoz, 2009), and in
whole-body pictures (Atkinson, et al., 2007). Munoz (2009) found that the using body cues only to interpret the target’s emotions, some participants exhibited hostile attribution biases, which had previously only been associated with interpretations of facial emotion cues (when using static images).

Thirdly, more recent studies reveal that not only are both the head and body necessary for accurate emotion identification (Nelson & Russell, 2011; Tracy & Robins, 2004), but that participants actually find it difficult to separate the face from the rest of the body when interpreting emotion cues (Aviezer, et al., 2011; Aviezer, et al., 2012). Consequently, these findings suggest that the best way to assess people’s ability to interpret the emotions of others is to use both the head and body stimuli together. For this reason the photographs will not be confined to only the faces of the models but include the full body of the models.

2.2.2 Universal and Basic Emotions

One consideration regarding photographic content is which emotions should be selected for capture. Firstly, the emotions should ideally be recognised by most people, limiting any cultural idiosyncrasies in expression and interpretation, and making the photographs applicable throughout a wide range of settings. Ekman (1999a) conducted a review of cross-cultural studies and found evidence for universal emotions, those that are recognisable by people from a range of different cultures. These emotions were fear/surprise, anger, happiness, disgust and sadness; however, Ekman (1999a) concluded that contempt should be included as many cultures also recognise this emotion.

Secondly, the emotional expression should be of one single emotion and not a blend of several emotions. Expressions need to be of one single emotion to ensure high validity (O’Sullivan, 1982). Blended emotions are those that combine the expressions of several different emotions at once, such as anguish, which might include pain, sadness, anger, and grief. Expressions of this type would be difficult for participants to identify, with different participants identifying different emotions, jeopardising both validity and reliability.
The studies on universal emotions led Ekman (1992a) to propose the existence of basic emotions. Ekman (1992a; 1992b; 1999a) considered these to be facial expressions of purely one single emotion as opposed to blended emotions. Ekman, et al., (1982) argued that Happiness, Surprise, Fear, Anger, Sadness, Disgust/Contempt and Interest could be considered basic affect categories, whilst blended emotions such as anguish and suffering, could be considered secondary affect categories. Despite different methodologies and research perspectives, these seven emotion categories seem common to much of the research (Ekman, 1992a; Ekman et al., 1982; Izard, 1994; Keltner et al., 2003; Munn, 1940; Pultchick, 1962; Tomkins & McCarter, 1964; Osgood, 1966; Woodworth & Scholsberg, 1955). Through a review of the research, however, Ekman (1999a) found no support for the emotion ‘interest’. Since then, Ekman and colleagues have considered only the remaining six emotions to be basic emotions.

There is some research that suggests that there may be fewer than six basic emotions. A recent study by Jack, et al. (2014) suggests that there are in fact only four basic emotions, arguing that anger/disgust and fear/surprise share many of the same facial elements. Jack et al. (2014) argue that these emotions have very similar facial features, and so can be easily confused. These similarities, however, seemed to be confusing at the more subtle levels of expression compared to more intense expressions. Recent research by Barrett (2006) argues against the existence of any universally recognised emotions at all. The debate for Barrett’s findings and conclusions is ongoing (see Barrett, et al., 2007; Izard, 2007; Panksepp, 2007a; 2007b), with many of the underlining assumptions of Barrett’s ideas being questioned. For instance the assumption that emotional expression is only based on individual conceptual notions of those emotions and related experiences, contradicting the idea of universally recognised emotions. As both Izard (2007) and Panksepp (2007a & 2007b) point out, there is ample evidence of basic emotions both across and within species, as well as evolutionary and neurobiological bases for basic emotion production. Despite these challenges to Ekman’s six
basic and universal emotions, the majority of research evidence strongly supports both the existence and the consistency of recognisability of these emotions (Ekman, 1992a; Ekman, 1999a; Panksepp, 2007b; Tracy et al., 2009). For this reason it was decided that these six emotions – happy, sad, angry, fearful, surprised, and disgust should be the emotions captured in the photographic stimuli. Studies of emotion recognition also include a ‘neutral expression’ (Ekman, 1992a; Ekman et al., 1982; Izard, 1994; Keltner et al., 2003; Tomkins & McCarter, 1964; Osgood, 1966). This is an ‘emotionless’ expression that does not portray any emotion at all. It is standard practice to include a non-emotion stimulus (e.g.: Matsumoto, et al., 2000; Blackwell, et al., 2012).

2.2.3 Emotional Expressions

2.2.3.1. Posed versus natural/spontaneous expressions.

Although most of the research into emotion interpretation employs posed expressions, the question of whether these are superior to photographs of natural expressions should be examined. Natural expressions occur spontaneously without direction from another on how to express a particular emotion, and therefore should be more valid than posed expressions. It is, however, difficult to apply an accuracy criterion to a spontaneously expressed emotion due to its fleeting nature (Ekman & Oster, 1982; O’Sullivan, 1982). These emotions would be difficult to capture and may be vulnerable to blending of several different emotions. Posed expressions, on the other hand, can be more controlled by the model, and more accurately rated for intent. That is, that the model can more easily assess which emotion they intended to express, which is then used as an accuracy criterion for comparison against participant responses. Therefore whilst naturally expressed emotions may be more valid assessing reliability is difficult, whilst posed expressions are more controllable, yet may be considered less valid.
Chapter one (section 1.6.3.1) reviewed the empirical support for the use of posed emotions in an emotion identification task. Research comparing posed with spontaneous emotions revealed that posed expressions attract the higher accuracy ratings compared to more natural expressions (Ekman et al., 1982; Motley & Camden, 1988; Zuckerman, et al., 1976). Naturally expressed emotions have also been found to be positively related to the ability to express posed emotions (Ekman, et al., 1982; Riggio, et al., 1985; Tucker & Riggio, 1988). Based on the weight of research, the photographs for this thesis will be comprised of models posing/acting the required emotions.

2.2.3.2. Eliciting emotional expressions for recording

Many different methods have been used to elicit specific emotions for study. These have included the reciting of poems (Feleky, 1914), practice of facial expressions (Frois-Wittman, 1930), live eliciting stimuli (Coleman, 1949), and other eliciting stimuli such as pictures and narratives (Feschbach & Feschbach (1969). There are, however, problems with many of these techniques including subjective judgements, inconsistent emotional outcomes, and ethical considerations.

O’Sullivan (1982) argues that elicitation of facial emotions can be problematic, such as ensuring that the emotion expressed, then captured and interpreted are all one and the same. In other words, it can be difficult to ensure that the emotion intended by the actor is the emotion that is captured, and is also the same emotion that can be distinctly and accurately interpreted by participants viewing the photograph during the task. These problems can be overcome by ensuring safeguards are put in place to help ensure a valid and reliable test. Firstly an accuracy criterion must be established (O’Sullivan, 1982). This is generally given as a self-report measure from the models who rate their success at capturing the desired emotion, or by simply reporting what emotion they felt at that time (Ekman et al., 1982). This then ensures that any judgements made by participants regarding the photographs can be accurately compared to the intended emotion.
Secondly, elicitation of each singular emotion needs to be carefully thought through (O’Sullivan, 1982). Humans experience a vast range of emotions, many of which can blend and merge into each other. Emotions such as suffering can incorporate more than one emotion or experience – pain, anguish, grief, anger, etc. According to Ekman (1992a; 1999a) basic emotions are singular in nature, therefore it is important that any elicitation be designed specifically for one and only one emotion at a time.

Ekman et al., (1972) propose several ways in which singular emotions can be elicited. They recommend the use of scripts, scenarios, or static poses, as well as the memories and experiences of the models themselves. Another technique is to direct the actor in which facial features should be posed, or which set of facial muscles should be used in specific ways (Ekman et al., 1982). For instance, for surprise, the eyebrows should be raised high and the eyes held wide (Ekman, 2007). Ekman’s (1971) Facial Affect Scoring Technique has been used in this way to obtain specific emotions from models. The technique relies upon the physiology of facial muscles, with emotions being mapped according to these specific muscle movements. The training and cost to use such a technique, however, can be prohibitive and not all models have the capacity to adequately control their facial movements to ensure that each emotion would be accurately represented, especially for more subtle levels of expression production. Therefore in this thesis, a set of elicitation exercises were developed for models to produce specific emotions. Models were first asked to draw upon their own memories or experiences from which they could reproduce the required emotion and therefore facial expression. If models had difficulty with this, then a scenario was given to them to enable them to imagine the specific emotion. For example, for fear, “you are walking down a dark alley. It is night time and unnaturally quiet, when you suddenly hear running footsteps behind you. You feel very frightened.” Alternatively, for some emotions physical stimuli was also available (e.g.: toy slime to elicit feelings of disgust). (See appendix A for more detail on the procedures and techniques used).
Thirdly, the process of producing several different emotions in a row, which models are often asked to do, can lead to a ‘bleeding’ of one emotion into the next (Ekman, et al., 1972; O’Sullivan, 1982). Davis (1934) conducted a reanalysis of Landis’ (1924, 1929) data and found evidence for one emotion carrying over to the next emotion category, resulting in a cumulative build-up of blended emotions. The researcher, therefore, needs to consider a way to cleanse any residual emotions between one category and the next. There seems to be little in the way of directives to address this particular problem in the research.

Although authors have highlighted the dangers of emotional leakage between emotion sets (Ekman, et al., 1972; O’Sullivan, 1982), much of the research involving the expression of emotions does not make mention of techniques to stop this from occurring. For example, Tomkins and McCarter’s (1964) study required participants to display several different basic emotions; however, no mention is made of the problems of emotional leakage from one emotion set to the next. In Ekman’s 1965 study photographs were taken during real-time interactions designed to elicit specific emotions, again there seemed to be no consideration given to the blending of emotions. Ekman et al. (1971) avoided this problem by asking models directly how to place their various facial features, such as turning the mouth down with raised eyebrows, as opposed to using emotion-words to guide them. The problem, however, is not confined to older studies. More recently, a photographic instrument designed specifically to test the facial emotion interpretation abilities of schizophrenics using actors and six basic emotions (Behere, Raghunandan, Venkatasubramanian, Subbakrishna, Jayakumar, & Gangadhar, 2008) does not deal with the issue of emotion leakage across emotion categories. Logic dictates that if the danger lies in the remnants of one emotion influencing the expression of the next emotion, then some sort of emotional cleansing would be necessary. This could be achieved with time between emotion sets, or some sort of exercise (physical, mental or both) designed to distract the model/actor, allowing them an opportunity to ‘reset’ emotionally.
The cleansing exercises designed for this thesis included physical exercises (jumping jacks, push-ups, etc); mental problem solving puzzles (Sudoku, crosswords); trivia games; and relaxation techniques (slowed breathing, concentrating on the breathing, and relaxing the body). The exercises were designed to be mentally stimulating and/or demanding, in order to take the model’s attention away from their emotions. (See appendix C. for more detail on the procedures and techniques used).

2.2.3.3. Varying expression intensity levels

One of the major aims of this thesis is that the photographic test should be able to not only assess AR-EA abilities, but also identify potential deficits or biases that may occur during interpretation. Biases in interpretation most often manifest through the use of ambiguous stimuli (Burt et al., 2009; Dodge, 1980; Dodge et al., 1986). Hostile attribution biases, for example, have been found to arise when other’s intentions are unclear (Dodge, 1980; Feldman & Dodge, 1987). People who attribute hostile intent in others use fewer cues (Dodge & Newman, 1981; Milich & Dodge, 1984), are more selective about which cues they pay attention to (Dodge & Frame, 1982), and are less accurate in recognising cues (Dodge & Frame, 1982), factors that would be exacerbated in the face of ambiguous stimuli. Many studies have successfully used ambiguous stimuli in their investigations of hostile attribution bias (e.g.: Burt et al., 2009; Dodge & Tomlin, 1987; Halligan, et al., 2007; Matthews & Norris, 2002; Takarangi, et al., 2008), with some using pictures as stimuli (see Orobio de Castro, et al., 2002). Therefore some of the photographic stimuli need to be ambiguous enough to allow these biases to surface. The best way to insert ambiguity into an instrument is to adjust the difficulty level of the task. There are two ways in which this can be done with photographic stimuli – vary the time which participants are exposed to the image, or vary the intensity with which each emotion is expressed. The following offers a brief overview of the literature regarding methods to vary task difficulty, as well as linking the final solution of expression intensity to bias detection.
One technique for making tasks of this nature more difficult is by shortening the time participants are exposed to the stimulus. Ekman and Friesen (1974b) used this technique with black and white photographs of the six basic emotions (happiness, sadness, anger, fear, disgust and surprise). When shown to participants for a period of 5 seconds, agreement between participants was high. In their Brief Affect Recognition Test (BART) participants were only shown images for 1/100th to 1/25th of a second in a simulation of ‘normal interpretation conditions’ (Ekman & Friesen, 1974b, p221), making the task more difficult. They found that the shortened exposure time did reduce accuracy ratings. Whilst this approach may be empirically effective (in terms of standardisation and reliability) it is not a technique that lends itself to larger testing populations, or in areas where technology may be limited. This is because the technique requires computer screens and programs where the exposure time of each image can be controlled.

Another way to alter the difficulty is to vary the intensity with which each emotion is expressed by the target. In the social world, not everyone expresses emotions to the same degree. This may be because of specific social circumstances (LaFrance, et al., 2003; Santiago-Menendez & Campbell, 2013), cultural display rules (Ekman & Friesen, 1974b; Glikson & Erez, 2013; Hess, et al., 2004) or indeed the person’s actual ability to express emotions facially (Dimberg et al., 2011; Ekman & Oster, 1982; Hess et al., 2000). Varying the intensity with which emotions are expressed will have the effect of varying the level of ambiguity for stimulus items. The less intensely expressed emotions will be more ambiguous than expressions of higher intensity, and the more ambiguous the item, the more difficult the task to interpret it.

Unfortunately, there are few studies that have used variation of emotional intensity to investigate empathic accuracy. Those studies that have report consistent results with participants obtaining lower accuracy ratings for less intense emotions and higher accuracy for more intense expressions (Hoffman, et al., 2010; Jorgensen & Howell, 1969; Montagne, et al.,
Females were found to be better at interpreting more subtle levels of expressions compared to males, although this difference disappeared at higher intensity levels (Hoffman, et al., 2010). One study found a negative correlation between intensity ratings and reaction time, meaning that participants took longer to identify emotions when they were more subtly expressed (Palermo & Coltheart, 2004). This clearly demonstrates that less intense emotional expressions increased the difficulty of the identification task for participants.

Therefore, the evidence indicates that varying emotional expression intensities can be used to vary the ambiguity of the stimuli and therefore the difficulty of the emotion identification task. Varying the intensity levels of the emotional expressions should allow any interpretation deficits or biases to emerge.

2.2.3.4. Accuracy Criterion

According to Ekman (1972), Ekman et al., (1982) and O’Sullivan (1982) it is important that any instrument assessing emotion decoding abilities use accuracy criterion for the stimulus items. That is, that the emotion stimulus used actually portrays the emotion it is purported to. As recommended by O’Sullivan (1982) and Ekman et al. (1982), accuracy criteria for these studies was established by having the photographic models rate themselves on how well they thought they captured each intended emotion, and at each level of intensity. The scale used for this in the current study can be found in Appendix D.

2.3 Taking the Photographs - Method

The procedures for taking the photographs in this current study are described in detail in Appendix A. To summarise, four male and four female models, ranging in age from 28 to 47 years of age (m = 34.38), some with acting experience and some without, were chosen according to specific selection criteria. Models were chosen on the basis that they were visually free of culturally or individually distinguishing features (i.e.: all Caucasian, absence of
tattoos and piercings, absence of any other physical and/or facial uniqueness, etc.). Previous research has found that people’s judgements about others are often based on stereotyped categories (Fiske, 1993) which can result in ethnicity-driven biases (Lott & Saxon, 2002). Models were also instructed to dress in a non-descript style in order to also avoid socio-economic biases (Lott & Saxon, 2002; O’Sullivan, 1982).

Each model was photographed separately with only the experimenter and photographer present. They were asked to demonstrate a neutral pose (no emotion) and then six different emotions – surprised, disgusted, fearful, angry, sad, and happy – at three different levels of intensity – slightly (low), very (medium), and extremely (high). Models were told before the photographic session that they were to capture each emotion at three different levels of intensity. For example, they would start with slightly happy – a little happy, with a slight smile; then increase their intensity to very happy – more than slightly happy but not the happiest you could be; and then extremely happy – which is the happiest they could be. This was done so that models could judge their own expressional intensity for each level. The final photographs consisted of full body shots of the models to allow for manipulation of photographic elements in later testing (see chapters 4 and 5).

### 2.3.1 Accuracy Criterion Outcomes of photographic sessions

The mean scores for models’ self-report ratings of how well they captured each emotion are presented in Table 2.1 below (also see Appendix D). Each model reported how successfully they believed they captured each emotion using a Likert scale from 1 “did not capture the emotion at all” to 5 “completely captured the emotion”. Most models reported a level 3 or better for nearly all emotion categories, with scores as low as two only reported on three occasions. No model reported a self-score lower than 2. Some models reported difficulty in capturing and expressing disgust, as well as anger, and even extreme happiness for one model. In most cases models reported that they had captured emotions better at the ‘extreme’ (high) levels of intensity compared to the ‘slightly’ (low) levels of intensity. When
the models gave low self-ratings for a particular emotion they were asked if they wished to capture the emotion again, but no model wished to do this. They did confirm, however, that they thought they had captured the intended emotion, albeit not very well.

Table 2.1

*Mean self-report ratings, out of a maximum of 5, of how well each model felt they captured each posed emotional expression.*

<table>
<thead>
<tr>
<th>EMOTION CATEGORY</th>
<th>INTENSITY LEVEL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Slightly</td>
</tr>
<tr>
<td>SURPRISE</td>
<td>3.75</td>
</tr>
<tr>
<td>DISGUST</td>
<td>3.50</td>
</tr>
<tr>
<td>FEAR</td>
<td>3.63</td>
</tr>
<tr>
<td>ANGER</td>
<td>3.94</td>
</tr>
<tr>
<td>SAD</td>
<td>3.75</td>
</tr>
<tr>
<td>HAPPY</td>
<td>4.00</td>
</tr>
</tbody>
</table>

One emotion category ‘disgust/contempt’ became problematic. One model believed that disgust and contempt were in two distinct emotions, and so could not be displayed at the same time. Therefore each model was asked which emotion, disgust or contempt, they felt more confident in portraying. All models chose disgust. In studies by Ekman disgust and contempt are often categorised together (see Ekman et al., 1969; Ekman et al., 1987) due to the fact that both emotional expressions shared many of the same facial movements - a slight uprising of and at times tightening of the lip, with the eyebrows pulled down (Ekman, 2007). Other studies, however, have considered disgust on its own (Ekman et al., 1971, Ekman, et al., 1980; Ekman et al., 1987). In 1986, Ekman and Friesen made the argument that contempt was distinctly recognisable from disgust and that the two emotions should be treated separately. Although Ekman (2007) argued that the two involved similar facial features, the motivation behind their expression was different. He described contempt as an emotion generally directed at others, whilst disgust was related more to taste, touch, and smell (Ekman, 2007). Indeed each model indicated that the motivations behind the expressing of disgust or
contempt were quite different. Based on these considerations the rest of this thesis will deal only with the emotion disgust and not contempt.

In some cases the models reported that they ‘acted’ the emotion rather than ‘felt’ it, although they did not think this interfered with their ability to express the desired emotion. When the model indicated that they had captured the emotion, the photograph was taken, but was not shown to the model at that point. This is particularly important in terms of the more subtle emotions where models may have believed, after seeing the photograph, that they had not captured the emotion, resulting in them intensifying their expressions in order to do so, undermining the intensity levels trying to be achieved.

When more than one photograph was taken to capture an emotion, the model themselves chose which photograph they believed captured the emotion best. This was done only after the entire session had been completed to avoid any ‘self-editing’. In this way models could not continually re-adjust their expressions based on the photographic product.

2.4 Testing the Photographs

Having taken the photographs, testing then needed to be conducted to ascertain if participants would be able to accurately identify the emotions portrayed by the models in the photographs. This was done over two studies, due to methodological problems in the first round of testing. The first study had participants identify the emotions in the photographs using open-ended responses. In the second study participants chose from a group so labels, assigning one label to each photograph. This was done via a forced-choice sorting-task. The methods, results, and discussion for each of these studies are presented below.

2.4.1 Study 1: Open-ended responding.

The aim of this first study was to ascertain if the emotions expressed by the models could be accurately identified by others. In this study participants viewed each photograph and then gave an open-ended response identifying the expressed emotion. Open-ended
responses were chosen in the belief that accurate identification of emotions with no emotion labels to guide responses (as in forced-choice responding) would indicate high validity of the photographs. The following hypotheses were made based on the literature.

Firstly it was expected that the emotion happy would be the most readily recognisable emotion (see Hall & Matsumoto, 2004; Hoffman, et al., 2010; Montagne, Kessels, Frigerios, de Haan, & Perrett, 2005; Montagne et al., 2007; Slepian, Wisbuch, Adams, & Ambady, 2011; Thompson & Meltzer, 1964) whilst the least accurate scores were expected for fear (Hall & Matsumoto, 2004; Hoffman et al., 2010; Montagne et al., 2005; Montagne, et al., 2007; Slepian et al., 2011).

Secondly, female models will be more expressive than their male counterparts and therefore attract higher accuracy scores (Buck et al., 1980; Ekman & Oster, 1979; Hess et al., 2000; Wagner, et al., 1986; Wagner, Buck, & Winterbotham, 1993; Zuckerman et al., 1975).

Thirdly, we expect the performance between male and female participants to be different, with females performing better (Ekman & Oster, 1979; Hall, 1978; McClure, 2000; Montagne, et al., 2005; Zuckerman et al., 1975), especially for the more subtle emotional expressions (Hall & Matsumoto, 2004; Hampson, van Anders, & Mullin, 2006; Hoffman, et al., 2010).

2.4.1.1. Method.

2.4.1.1.1 Participants.

Participants were a convenience sample of adult university students and other volunteers sourced through social network sites and snow-ball techniques. A total of 31 participants were approached to take part, however, only 25 returned their booklets. These participants ranged from 21 to 67 years of age, and consisted of 17 females and 8 males. Anonymity was preserved by assigning each participant a testing/booklet code (e.g.: 1AB; 2GH).
2.4.1.1.2 Materials.

The photographic session (see section 2.3 above) resulted in 152 photographs across 8 models and 6 emotion categories plus neutral. Each of the six emotions was presented at three different intensity levels, low, medium, and high. There was concern that having each participant rate all 152 photographs may result in testing fatigue which could result in lower task performance (van der Linden, Frese, & Meijman, 2003), therefore each participant rated only two models (one female and one male).

Booklets were made up of 38 full body black and white photographs (A4 size), with one photograph on each page. Each booklet contained 19 male and 19 female photographs (see Appendix E. for an example booklet). In order to reduce model interaction effects, each model was paired randomly with a model of the opposite sex, resulting in 16 different sets of model combinations, and therefore 16 different booklet sets. In each booklet set, the photographs were presented in random order.

2.4.1.1.3 Procedure.

Participants were briefed on the nature of the study via an information sheet. In order to avoid undermining the purpose of the study, participants were told only that it involved expression of the human face and body. The participants signed a consent form at the beginning of each booklet. Booklet instructions directed participants to look at each photograph and answer the following questions in their own words: “What do you think this person is feeling?” and “What do you think this person is thinking?”. Participants were also instructed not to overthink their responses but to rather rely on their first instincts. This was done to simulate the way this type of processing would be conducted in the social world, which for the most part would be quickly and almost automatically (see Ekman & Friesen, 1974a; Knyazev, et al., 2010). On return of the booklets, participants were fully debriefed regarding the nature of the study.
2.4.1.4 Analysis.

Observations were recorded from both participant feedback and booklet responses. Many participants reported difficulty with the task, especially for particular models. Participants also said that when they could not interpret the expressed emotion via the face of the model, they relied on the body. The implications of this will be considered in the discussion section later in this chapter.

Participant responses were recorded verbatim and then categorised according to common meanings and definitions of words/phrases into one of eight categories: neutral, surprised, disgusted, angry, fearful, sad, happy, and other. The first seven categories represented the basic emotions originally portrayed by the models. The ‘other’ category contained any words/phrases that did not fit into the other emotion categories and included responses such as ‘curious’, ‘pride’, and ‘apologetic’. Where the word/phrase for the emotion response was unclear, responses to the second “thinking” question were used for clarification. For example, participant 1A-E gave the response ‘rejective’ for the emotion anger, which was unclear. The participant’s response to the second question, however, was “more than likely telling the person to piss off”, which indicated the category ‘anger’.

In total participants provided 272 unique words and/or phrases to describe the 6 intended emotions and the one neutral expression. Matches between participant responses and the model’s intended emotion attracted a score of 1, whilst incorrect matches were scored 0. Therefore scores could range from 0 to 38. Not all photographs were assessed the same number of times, therefore total scores for each emotion category and intensity level were converted to percentages.

2.4.1.2. Results.

The overall percentage accuracy score across all categories of emotion and intensity levels were above 50% for the entire sample ($m = 55.68$, $SD = 21.95$). The average accuracy
percentages by emotion category (see figure 2.1. below) show that participants were able to interpret happiness the most accurately, whilst neutral expressions were the most inaccurate.

![Figure 2.1. Average accuracy percentage of photographs by emotion category](image1)

The most intense levels of emotion elicited the most accurate responses with both ‘medium’ and ‘high’ intensity levels averaging above 40% accuracy (see figure 2.2 below).

![Figure 2.2. Average accuracy percentage of photographs by intensity level](image2)
In all emotion categories, female models attracted higher percentage accuracy ratings than male models (see figure 2.3 below).

Of the total 152 photographs available, 71 photographs attracted an accuracy percentage of 50% or above, with 25 of these coming from male models and 46 from female models. Another 19 photographs received between 30% and 49% accuracy, 10 male models and 9 female. A total of 31 photographs received 0% accuracy (see figure 2.4 below), meaning that no participant was able to identify the expressed emotion correctly. Male models produced 23 of these, and female models only 8. Male models C and F obtained the lowest accuracy scores overall and also had the highest number of incorrect matches. Model C attracted 0%’s on all intensity levels for both disgust and fear, and two levels of angry (low and medium). Model F had 0%’s recorded against all intensity levels of surprise and disgust, whilst Model H had no matches for sadness at any level of intensity.

![Figure 2.3. Comparison of average accuracy percentages of photographs between male and female models by emotion](image-url)
Happiness obtained the most consistent accuracy percentages with 20 out of the possible 24 photographs attracting 50% or better ratings. Totalling all the photographs that attracted 30% or higher accuracy ratings the following number of photographs were deemed useable from each emotion set: 21 Happiness; 17 Surprise; 16 Anger; 13 Sadness; 11 Fear; 10 Disgust; and 2 Neutral. The lack of performance for the neutral category is not surprising given that participants were asked to identify an ‘emotion’ and were not told that neutral or ‘no emotion’ was a possible response.

Female participants had a higher mean accuracy rating ($m = 57.43$, $SD = 22.45$) than male participants ($m = 51.97$, $SD = 19.27$). Although sample sizes were small and raw scores converted into percentages, inferential statistics such as t-tests can still be performed due to their robust nature (Haslam & McGarty, 2010). An independent samples t-test revealed that this difference between male and female participant scores was not significant ($t(23) = .57$, $p = .573$; cohen’s $d = 0.26$), meaning that males and females demonstrated similar emotion identification abilities for these photographic stimuli. Although this was a non-significant results statistically, it should be noted that the effect size is still quite high indicating that, with larger sample sizes a difference between the sexes may be observed.

![Figure 2.4. Number of photographs that scored 0.00% accuracy by model](image)
2.4.1.3. Discussion.

Overall participants were able to identify most of the emotional expressions portrayed by the models. More than half the photographs attracted accuracy ratings of 30% or higher. As expected happiness was the most accurately identified emotion, whilst neutral was the least accurate. The result for happiness concurs with previous research that found this emotion is the most easily identifiable of the basic emotions (see Adolphs, Damasio, Tranel, & Damasio, 1996; Hall & Matsumoto, 2004; Hoffman, et al., 2010; Montagne, et al., 2005; Slepian, et al., 2011; Thompson & Meltzer, 1964). Ekman (2007) makes the point that happy is the only emotion where there is a distinctive upturning of the mouth, making it quite different and therefore more easily recognised compared to other emotions. The performance of participants for the neutral category was not expected. This may be due to the fact that participants were not informed that neutral poses would be in the booklet, and that ‘no emotion’ was an acceptable response, resulting in participants erroneously assigning an emotion where there was none expressed.

As expected, participants were able to identify the emotions of female models more accurately than male models, which is in line with previous research (Buck et al., 1980; Ekman & Oster, 1979; Hess et al., 2000; Wagner, et al., 1986; Wagner et al., 1993; Zuckerman et al., 1975). Hess et al., (2000) suggest that this is partly due to gender stereotypes, where it is expected that women will experience emotions more intensely and therefore express them more intensely as well. Display rules might also explain why females seem to be more emotionally expressive than males. Santiago-Menendez and Campbell (2013) found that whilst both boys and girls were likely to react with crying in angry and sad circumstances, girls reported more instances of doing so than boys. The authors state that some reasons may be more acceptable than others when it comes to crying in boys. For instance, depression was a strong predictor for crying in boys, but not so anger. Therefore crying may be ‘ok’ but only for specific reasons or under certain circumstances. La France et al., (2003) also propose that the
frequency of emotional displays is regulated through gender norms, social circumstances, age, and culture. Therefore the research suggests that it is more acceptable for females to be more emotionally expressive than males, which could explain the differences in accuracy scores between male and female models.

There were no significant differences in performance between male and female participants, against expectations, although effect sizes indicate that there may be a difference observable with a larger sample size. Whilst some research has found that women are superior to men in discerning the emotions of others (Ekman & Oster, 1979; Hall, 1978; McClure, 2000; Montagne, et al., 2005; Zuckerman et al., 1975), the literature is inconsistent in this area. Hoffman et al., (2010) found that the advantage for women in identifying emotions does not necessarily hold for all levels of emotional intensity. Instead it seems to be focused on more subtly expressed emotions (Hall & Matsumoto, 2004; Hampson, et al., 2006; Hoffman, et al., 2010).

Differences between male and female AR-EA abilities, seems to be predicated upon more nuanced factors than just wholesale gender differences. For instance, research has found that women are faster than men in correctly identifying emotion expressions in others (Hampson et al., 2006; Knyazev et al., 2010); however, if the time factor is removed, there is no difference in actual ability (Hampson et al., 2006). Knyazev, et al., (2010) argue that this is due to different processing routes used by the different sexes. Other factors include the sex of the target, where opposite-sex targets are processed more quickly than same-sex targets (Hoffmann, et al., 2006), or the type of incentive or motivation involved (Klein & Hodges, 2001). Therefore differences in these abilities may not be a simple process, but a result of varying factors including processing mode, target sex, and motivation.

Participants reported that when the emotion expressed by the face of the model was unclear, that they relied on cues from the body. The literature concurs with this perspective finding that participants utilise more cues than only those provided by the face to interpret
others’ emotions (Nelson & Russell, 2011; Tracy & Robins, 2004). More than this though, when presented together, participants will process the whole of the image and not just the individual components, such as only the head or only the body (Aviezer, et al., 2011; Aviezer, et al., 2012). Although only anecdotal, the participant feedback supports this notion that more cues, when available, will be used to decode emotions.

Participants also reported that body position helped identify the expressed emotion, especially where the model’s face was difficult to interpret. Whilst this might be the case, there may well have been instances where this mislead rather than helped the participant. For example, Model G represented ‘very disgusted’ with the body bent over and leaning forward. This elicited many responses of pain such as “my back hurts” (1AG). Model A’s ‘very fearful’ pose saw the model put their hands to their lower back/hip area, to which one participant responded “confusion – where’s my wallet”(2AE)? In order to ascertain if these were isolated responses, more testing of the photographs will need to be conducted.

The different intensity levels of emotions resulted in accuracy ratings that reflected the varying difficulty levels. That is that the more subtle expressions attracted the lowest accuracy ratings whilst the most expressive expressions received the highest accuracy ratings. This is not unexpected as more subtly expressed emotions would be more difficult to interpret (see Hoffman et al., 2010). This indicates that the different intensity levels of emotions have successfully influenced accuracy scores.

One of the major limitations of this study was the absence of more detailed instructions to participants to let them know that a ‘no emotion/neutral’ response was possible, and that emotions would be presented at varying levels of intensity, and therefore repeated. Overall accuracy percentages for the photographs were quite low, with less than half attracting an accuracy rating of 50% or more. A total of 272 words and/or phrases were used by participants to label just six emotions and one neutral expression. The ‘other’ category contained 103 unique descriptors demonstrating the range of emotions ‘seen’ by the
participants, which were far more than the original six emotions intended. Part of the reason for this may have been the booklet design and instructions. Participants were instructed to describe the emotion being expressed by the photographed target, but were not informed that there were only six different emotions presented at three different levels of intensity. This may have set up a ‘demand effect’, where participants tried to come up with a new word or identify a new emotion for each and every photograph, rather than repeating responses they had already given. Inclusion of this information may have resulted in fewer varied responses outside of the original seven emotion categories.

This along with the evidence that participants were misled by some of the body positions of the models means that further testing is necessary. It was therefore decided that the photographs should be tested again, this time using a forced-choice design, in which participants would select from a range of emotion category labels. The specific methodology and procedure for this is presented in the following section.

2.4.2 Study 2: Forced-choice sorting task.

Due to the expansive nature of the responses given by participants in the first study, accuracy rating results may not have reflected the true accuracy with which the photographs could be interpreted. Open-ended responses were first chosen due to the high validity this method offered (Russell, 1994); however, with over 270 different responses to just seven different emotion categories, low accuracy percentages resulted. Therefore the open-ended response methodology, although valid, did not seem to work for this kind of stimulus material.

Another approach is to use a forced-choice design which has been found to increase the accuracy with which participants identify emotional expressions. Both Woodworth and Scholsberg (1995), and Tomkins and McCarter (1964) found that when specific emotion categories were introduced, accuracy ratings rose considerably. The problem here though is that participants no longer bring their own interpretations to the identification of emotions, having only restricted categories or labels to choose from. There is an assumption then, that
the targets' emotion can only be interpreted in specific ways, according to the categories given. This makes the task of identifying emotions much easier, but severely limits the range of possible responses. What the previous study uncovered was that participants were able to attribute a wide range of emotion words to a limited set of emotions.

Cronbach (1950) argues that participants have habitual ways of responding and these habits can become more evident with the use of open-ended responses, negatively impacting validity. Rosenberg and Ekman (1995) tested free-choice responding against force-choice methods and found that the two were comparable. More recently Matsumoto et al. (2000) stated that forced-choice responding was more user-friendly than open-ended responses, and did not pose a threat to the validity or reliability of a measure, validating the decision to use a forced-choice design.

A new methodology was also chosen for the current study in order to assess not only the accuracy with which participants could identify emotion categories, but also the extent to which they recognised varying levels in intensity of those emotions. This was achieved through a sorting task. This method is loosely based upon the Q-Sort methodology used in studies of personality (Knight, 2002; Mercier, Piat, Peladeau, & Degenais, 2000; Niec & Russ, 2002; Taft, 1966; Western, Muderrisoglu, Fowler, Shedler, & Koren, 1997). The method requires participants to physically sort items into appropriate categories. The forced-choice component consists of a limiting the number of items that can be sorted into each category. In this study participants were asked to sort photographs into appropriate emotion categories, and then, for each model within a set category, sort the model's photographs according to the intensity level of the emotion expressed.

In the previous study intensity levels were assumed to be correct because the more subtle expressions attracted the least accurate scores, and the most expressive, and highest accuracy scores; however, the intensity levels within each emotion category were not directly tested. Having participants sort the photographs into order according to intensity level will
allow for direct analysis of how participants are performing for each emotion category and at each intensity level. A literature review has uncovered no research that has used the q-sort procedure, or any other sorting method, for identifying and categorising expressed emotions in photographs. Nor have these methods been used to then rate photographs by emotional intensity. Having found the many, and time consuming, pitfalls of open-ended responses to the original photographs, it seems surprising that this type of sorting technique has not been employed more often.

Since the expectation is that the pattern of responding by participants will be consistent with the previous study, although with higher accuracy ratings, the following predictions have been made. Happy is expected to be the most accurately identified emotion and fear the least accurate. Although neutral attained the least accurate scores in the last study, the inclusion of a neutral category in the present study should see an improvement in the number of correct responses by participants. It is expected that more subtle expressions will be classified as the least intense, whilst strongly expressed emotions classified as the most intense, which will be reflected in the accuracy scores for each level as in the previous study. Female models, again, are expected to be more expressive than male models, and therefore attract higher accuracy ratings. The previous study found no differences in performance between male and female participants, therefore it is predicted that this performance trend will continue.

2.4.2.1. Method

2.4.2.1.1 Participants.

There were a total of 16 participants, 7 males and 9 females, which included both university students and members of the general public. Participants ranged in age from 19 to 58 years, with a mean age of 34.25 years.
2.4.2.1.2 Materials.

Full body black and white A4 laminated photographs were used, with models expressing 3 different levels of one of 6 emotions, plus one neutral pose. Four male and four female models were included in the photographs, however, participants only sorted two male and two female models in one session, a total of 76 photographs, in order to avoid task fatigue. The emotion categories set up for sorting were labelled: Neutral, Surprised, Disgusted, Fearful, Angry, Sad, and Happy. Each set of photographs were randomised before sorting, and emotion labels were also placed in random order for each participant.

2.4.2.1.3 Procedure.

Participants were tested in an office at a desk large enough for the photographs to be sorted upon. They were then briefed regarding the study, although information that might alter their responses, such as the true purpose of the research, was withheld until debriefed at the end of the task. Participants filled in a form that included demographic information such as age, sex, income level, and education level, as well as consent. Participants provided their names to ensure that they could not participate in any future studies of the same photographs. If participants wished to be informed about the outcomes of the study, they supplied their email address.

The sorting task was described to participants as a two stage process. In the first stage participants were asked to sort the randomised photographs into six emotion categories and one neutral category provided. Each category had a maximum number of photographs it could contain: a maximum of four for the neutral category, and 12 photographs for every other emotion category. This was the forced-choice element of the task where participants would be required to make decisions about which photographs to include or exclude in each category if they had too many or too few photographs in each.
Stage two involved participants then sorting the photographs for each model in order of intensity level, for that specific emotion category. Participants were advised to work on just one emotion category and one model at a time within each category. Participants could choose if they arranged the photograph in ascending or descending order of intensity and simply advised the researcher as to which order they chose.

2.4.2.1.4 Analysis.

Two different types of scores were calculated. Emotion Category (EC) scores were calculated on the basis of 1 mark for each photograph sorted into the correct emotion category, for each model. Emotion Intensity (EI) scores were calculated on the basis of 1 mark for each photograph sorted into the correct intensity level order within each category, for each model. This scoring method meant that each model (per participant) could attract a range of 0 to 3 points for each emotion category and each group of intensity levels within that emotion category, except for neutral which did not have any intensity levels. This meant that the category of neutral could only attract a score of 0 or 1. A participant’s EC scores could therefore range from 0 to 76 and EI scores from 0 to 72.

EI scores had the potential to be confabulated by the EC score as the EI score contained both. A perfect EC score of 3 for a model could result in an EI score ranging between 0 – 3 for the same emotion category, whilst an EC score of 0 would result in an EI score of 0. For this reasons, EC and EI scores were treated separately. Analysis included scores for each model, each emotion category, each level of emotion intensity, and a total score for each participant. Each model was assessed 8 times by the various 16 participants. Both EC and EI raw scores were calculated and entered into SPSS for analysis.

2.4.2.2 Results.

2.4.2.2.1 Emotion category (EC) and emotion intensity (EI) scores.

Participants scored an average of 48.06 out of 76 (63.23%; SD = 6.22) for EC scores and 38.06 out of 72 for EI scores (52.86%; SD = 8.13). The total possible male/female model EC
score was 38 and for EI was 36. Female models had higher mean EC and EI scores than male models (see Table 2.2 below).

### Table 2.2

<table>
<thead>
<tr>
<th></th>
<th>EC Score</th>
<th></th>
<th></th>
<th></th>
<th>EI Score</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Min.</td>
<td>Max.</td>
<td>Mean</td>
<td>SD</td>
<td>Min.</td>
<td>Max.</td>
</tr>
<tr>
<td>Male Models</td>
<td>20.44</td>
<td>3.98</td>
<td>12</td>
<td>25</td>
<td>16.50</td>
<td>3.97</td>
<td>6</td>
<td>21</td>
</tr>
<tr>
<td>Female Models</td>
<td>27.62</td>
<td>4.14</td>
<td>21</td>
<td>34</td>
<td>21.56</td>
<td>5.43</td>
<td>9</td>
<td>28</td>
</tr>
</tbody>
</table>

**2.4.2.2.2 Overall accuracy percentages by model, emotion, and photograph.**

Table 2.3 below shows the accuracy percentages obtained by each photograph and model. These values varied widely according to the emotion category, the intensity level of that emotion, and the model themselves. Accuracy percentages included the full range of possible scores – 0 to 100%. Zero percentages involved only the low and medium intensity photos and none of the extreme intensities. Surprisingly, most photographs that attracted 0% ratings belonged to male models with only one female model attracting a 0% for fear at low intensity.

Each model was assessed 8 times, meaning that each of their photographs could attract a maximum 8 out of 8 with an overall maximum score of 8 x 19 = 152 across all emotion categories and intensity levels. Mean percentage ratings demonstrated that happy attracted the highest accuracy ratings followed by neutral (see figure 2.5 below). Sad, surprise, disgust and anger were very close together with fear having the lowest accuracy ratings.
Table 2.3
Inter-rater accuracy percentages by photograph and model

<table>
<thead>
<tr>
<th>MODEL</th>
<th>NEUTRAL</th>
<th>SURPRISE</th>
<th>DISGUST</th>
<th>FEAR</th>
<th>ANGER</th>
<th>SAD</th>
<th>HAPPY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model</td>
<td>Sex</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>A</td>
<td>M</td>
<td>87.50</td>
<td>62.50</td>
<td>100.00</td>
<td>75.00</td>
<td>75.00</td>
<td>100.00</td>
</tr>
<tr>
<td>B</td>
<td>F</td>
<td>100.00</td>
<td>25.00</td>
<td>87.50</td>
<td>87.50</td>
<td>12.50</td>
<td>100.00</td>
</tr>
<tr>
<td>C</td>
<td>M</td>
<td>87.50</td>
<td>25.00</td>
<td>100.00</td>
<td>87.50</td>
<td>62.50</td>
<td>0.00</td>
</tr>
<tr>
<td>D</td>
<td>F</td>
<td>87.50</td>
<td>0.00</td>
<td>87.50</td>
<td>75.00</td>
<td>50.00</td>
<td>100.00</td>
</tr>
<tr>
<td>E</td>
<td>F</td>
<td>62.50</td>
<td>37.50</td>
<td>75.00</td>
<td>62.50</td>
<td>62.50</td>
<td>87.50</td>
</tr>
<tr>
<td>F</td>
<td>M</td>
<td>62.50</td>
<td>25.00</td>
<td>75.00</td>
<td>75.00</td>
<td>25.00</td>
<td>0.00</td>
</tr>
<tr>
<td>G</td>
<td>F</td>
<td>87.50</td>
<td>100.00</td>
<td>87.50</td>
<td>100.00</td>
<td>100.00</td>
<td>0.00</td>
</tr>
<tr>
<td>H</td>
<td>M</td>
<td>37.50</td>
<td>12.50</td>
<td>25.00</td>
<td>75.00</td>
<td>87.50</td>
<td>50.00</td>
</tr>
</tbody>
</table>
The most intense levels of emotion attracted the highest percentage accuracies (see figure 2.6 below), whilst the lowest levels of intensity were the least accurate. A repeated measures ANOVA revealed that the differences were significant ($F(2,30) = 17.90, p < .001; \eta^2 = .54$), a moderate effect. Post-hoc analyses revealed significant differences between all three intensity levels (at $p < .05$).
2.4.2.2.3 Model sex.

Paired samples t-tests were used to explore if there were any differences between the EC and EI scores of male and female models. This was chosen as each participant reviewed both a male and female model. Although the sample size for this is quite small, Winter (2013) states that if effect sizes are not required then t-tests for samples as small as 2-5 can be used (also see Bray & Maxwell, 1985). Significant differences were found between both EC ($t(15) = -5.50$, $p < .001$; estimated $d = 1.15$) and EI scores ($t(15) = -4.11$, $p = .001$; estimated $d = 0.62$)(see figure 2.7 below). Effect size calculations indicated that these differences were moderate to large, however since the sample sizes are small, these should be interpreted with caution (Winter, 2013). In both cases the expressions portrayed by female models were more easily discerned by participants than those of the male models.

![Figure 2.6. Average accuracy percentages for three levels of emotion intensity (EI)](image-url)
Figure 2.8 below demonstrates that most models averaged 50% accuracy or better across all emotion categories, except for models C and F both of which were male. The best performance was by Model E, a female model. All female models, except model G, attained higher accuracy scores than the male models.

Figure 2.8. Mean EC accuracy percentages by model

This pattern of female models attracting higher EC accuracy scores was seen for every emotion category except anger where male models scored higher (see figures 2.9 below). Due to
the small sample size a MANOVA could not be run (Bray & Maxwell, 1985), however, looking at the graph below we can see that fear, disgust, and sad revealed the greatest differences between male and female models, whilst neutral, surprise and happy showed the least differences in mean accuracy scores.

Analysis of model sex and emotion intensity levels revealed a significant interaction between the two at $p = .015$. Responses to emotion intensities for both male and female models followed the predicted pattern of least accurate with slight emotions and most accurate with extreme emotions (see figure 2.10 below), however, female models attracted more accurate scores than male models.
2.4.2.2.4 Participant sex.

In order to see if sex differences existed between male and female participants, an independent samples t-test was conducted. Winter (2013) argues that t-tests are quite robust with even very small sample sizes. Both EC and EI scores revealed no significant differences ($t(14) = .75, p = .464$; cohen’s $d = 1.77$ for EC; $t(14) = .57, p = .576$; cohen’s $d= 1.06$ for EI), meaning that male and female participants demonstrated similar abilities in identifying both categories of emotions and the intensity levels of those emotions. In both cases, cohen’s $d$ indicates a large effect, however, small sample size also means that interpretation should be done with caution (Winter, 2013).

2.4.2.3 Discussion – Study 2.

Overall, the use of a forced-choice methodology that provided participants with emotion labels, increased accuracy ratings across all emotion categories. As expected happy was again the emotion that participants found easiest to identify, whilst accuracy ratings for neutral rose dramatically. Participants found fear the most difficult emotion to correctly identify. As already discussed in the previous study, other research confirms these findings that happy is the most

Figure 2.10. Average percentage accuracy scores of different intensity levels by model sex.
readily recognisable emotion (Hall & Matsumoto, 2004; Hoffman, et al., 2010; Montagne, et al., 2005; Slepian, et al., 2011; Thompson & Meltzer, 1964), and fear the least recognisable (Hall & Matsumoto, 2004; Hoffman et al., 2010; Montagne et al., 2005; Slepian et al., 2011). Many studies reflect this same pattern of responding, except for in the case of the hardest emotion to identify, with some studies finding these to be sadness or contempt (e.g.: Langner, et al., 2010; Matsumoto et al., 2000). Therefore the pattern of accuracy ratings for this current study is consistent with the literature.

The more subtle levels of expressed emotion attracted less accurate responses compared to the more intense levels, as expected. There is very little in the literature regarding variations in intensity levels for individual emotions. Hoffman et al., (2010) found that more intensely expressed emotions were more accurately identified by participants compare to more subtly expressed emotions, as did Palermo and Coltheart (2004), which is consistent with the current findings. Therefore the current results suggest that the three different levels of intensity are having the expected influence on AR-EA abilities and are therefore working as desired.

Female models attracted higher scores on nearly all emotion categories and all levels of emotion intensity compared to male models, which is consistent with both the previous study and other literature. The moderate to large effect sizes gained indicate that the differences between the male and female models are strong. Palermo and Coltheart (2004) found that participants read female faces more accurately than male faces, a finding that is consistent across many studies (Buck et al., 1980; Ekman & Oster, 1979; Wagner et al., 1986; Wagner et al., 1993; Zuckerman et al., 1975). A comprehensive review of studies prior to 1990 by Kring and Gordon (1998) reveals that this superior expressivity in females seems to be evident across all categories of emotion. Again as already discussed in the study 1 of this chapter, these differences may be the result of display rules, social constraints, social context, culture, or age (Ekman & Oster, 1979; Gilkson & Erez, 2013; Hess et al., 2000; LaFrance et al., 2003; Santiago-Menendez & Campbell, 2013).
The superior expressivity in the female models of this current study held for all emotion categories except anger. Male models’ anger expressions were recognised more accurately by participants compared to female models’ expressions. It could be that the particular male models used for these photographs happen to be more expressive in this one particular emotion, or it could be that males more easily express anger compared to females. There is only one study that seems to lend support for this finding. Rotter and Rotter (1988) found women in identifying all emotions, except for anger where males performed better. The authors theorised that this was due to the social roles that males and female inhabit, in which anger is considered a more acceptable emotion for males to express compared to females. This notion is backed by other research which suggests that the stereotypes we associate with others can impact emotion recognition (Bijlstra, Holland, & Wigboldus, 2010). For instance Becker, Kenric, Neuberg, Blackwell, and Smith (2007) argue that there is a gender bias for the recognition of angry faces which could be due to the gender roles and stereotypes applied to men and women. Brescoll and Uhlmann (2008) found that in the workplace, expressions of anger were deemed much more acceptable coming from men compared to women, indicating a gender stereotyping bias which may explain why females do not facially express anger well. Only further studies of these photographs will reveal if this particular finding is consistent.

There were no differences found in accuracy scores between male and female participants. As in the previous study, there are studies both for (Hampson et al., 2006; Limbrecht, Rukavina, Scheck, Walter, Hoffman, & Traue, 2012; Wagner et al., 1986;) and against this finding (Ekman & Oster, 1979; Hall, 1978; Hall & Matsumoto, 2004; Hampson et al., 2006; Klein & Hodges, 2001; McClure, 2000; Montagne, et al., 2005; Rotter & Rotter, 1988; Tracy & Beall, 2011; Zuckerman et al., 1975). With the literature split the only conclusion that can be drawn at this stage is that for these specific photographs both males and females perform equally. It should be noted however, that the effect sizes here were quite large at a difference of more than one standard deviation between male and female participants. This implies that a large sample may reveal differences. Further testing will enable later confirmation of this conclusion.
2.4.3 Discussion – Study 1 and Study 2 Comparison

Accuracy ratings increased noticeably when a forced-choice methodology was used compared to open-ended responses. Of note was the increase in accuracy of the neutral emotion category. As already discussed above, the initial low ratings may have been due to poor design. Participants in Study 1 were not informed that a ‘no emotion’ option was available, and therefore assigned an emotion to the neutral photographs where there was none. In the Study 2, participants were given ‘neutral’ as a response category, and as a result accuracy increased.

The differing methodologies and score calculations make comparisons difficult, however, it is worthwhile acknowledging the similarities in results between the two studies. Figure 2.11 below shows the accuracy percentages reached by each study side by side for each emotion category. Note that this is not a statistical comparison, rather a demonstration of the responding patterns that were evident for both studies.

In both Study 1 and Study 2 Happiness was the most accurately identified emotion category, a finding which is consistent with the literature as discussed above (see section 2.4.1 above).

![Figure 2.11. Comparison of mean percentage accuracy scores for each emotion category of study 1 with study 2](image_url)
In the first study the disgust category was the next lowest scoring emotion followed by fear. Sadness, surprise and anger were closely grouped together (see figure 2.11 above). In the second study the lowest accuracy scores were obtained by fear and disgust, the same two emotions as in the first study (ignoring neutral), but in reverse order. Sadness, surprise and anger were again closely clustered in the second study. Overall the second study revealed less variation in accuracy scores across the different emotion categories compared to the first study.

The two studies showed similarities in responding patterns, indicating some consistency in how the photographs were being interpreted. This implies that the photographs provide a consistent measure of AR-EA abilities. The accuracy rating information obtained in these two studies provided the information to be used to make decisions about which photographs should be included for use in further studies. The photographs selected will also be tested for validity and reliability with the aim of formulating an AR-EA photographic measure to be used for the assessment of empathic accuracy abilities.

2.5 Selecting the Photographs for Inclusion in the Empathic Accuracy Measure

The initial testing of the photographs demonstrated that participants were able to identify the six basic emotional expressions and one neutral expression (as per the research of Ekman (1992a; 1992b; 1999a), as portrayed by the male and female models. The next step was to select which of these photographs should be included in for further testing in the development of the AR-EA photographic measure. The following provides an overview of this process.

2.5.1 Rationale and selection.

The full procedure for the selection of the final 40 photographs included for further testing can be found in appendix F. To summarise, accuracy percentage ratings were calculated for all photographs. Initial selection was based upon two considerations. Firstly that the photographs with the highest accuracy ratings be included, and secondly, that each emotion category and intensity level have one male and one female model included. This resulted in a total of 38 photographs.
being selected (see table 2.4 below). Two more photographs were also included, both of which attracted accuracy ratings of 0%. These photographs were deemed ‘ambiguous’ and were added in order to facilitate the identification of biases or misinterpretations.

Table 2.4

*Final counts for numbers of photographs from each model selected for inclusion in the AR-EA test.*

<table>
<thead>
<tr>
<th>Male Models</th>
<th>Total</th>
<th>Female Models</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>C</td>
<td>F</td>
<td>H</td>
</tr>
<tr>
<td>B</td>
<td>D</td>
<td>E</td>
<td>G</td>
</tr>
</tbody>
</table>

If selection of the photographs has been done successfully, it would be expected that the patterns of responding by emotion category and intensity level seen in the above studies should be repeated in the studies conducted in the following chapters. Refer to Appendix G for the photographs selected.
Chapter 3

3.1 Overview

Having established a set of photographic testing stimuli that can be used to assess affect recognition-empathic accuracy (AR-EA) abilities in participants, the focus shifts to the investigation of the impact of imagination and connectivity on AR-EA abilities. As was discussed in Chapter one (Section 1.7.1) although both imagination processes (Davis, 1980; Eisenberg & Miller, 1987; Hoffman, 1984) and connection between target and observer (Batson, et al., 2007; Batson et al., 1995; Carroll & Russell, 1996; Marian, & Shimamura, 2011; Walbott, 1988) are important factors in empathy processes, there has been little research into how these factors may be facilitated in a measure of AR-EA abilities. It was argued in the opening chapter that photographs could provide an effective and practical means via which insertion of an element to enhance these processes could be done, namely a silhouette (see Chapter 1, section 1.8).

It was also argued that inclusion of specific imagination and connectivity elements, which are central to empathic processes, would also enhance the photographs themselves as a measure of AR-EA abilities. In other words, including the silhouette would increase the reliability and validity of the photographic stimuli for measuring AE-RA. This chapter therefore has two main aims. The first aim is to test the impact of imagination and connectivity on AR-EA scores. The second aim is to continue the development of the AR-EA photographic measure by conducting validity and reliability testing, on different types of presentation modes of the photographs. To accommodate this, the chapter is split into the two sections.

The chapter will begin with a brief overview of the literature concerning the role of imagination and connectivity in empathic processes. This will also include a discussion regarding the different photographic presentation modes to be employed for the studies. The chapter then moves onto the first study: Study 1 Imagination and Connectivity. This section presents the investigation of enhancing imagination and connectivity and the impact this may have on participant AR-EA abilities. This will include the method and results of the testing conducted and a brief discussion of the
outcomes. The results from this study revealed that enhancement of imagination and connection processes is possible, and that their inclusion did impact participant AR-EA scores.

The second section presents the reliability and validity testing conducted on the photographic measure, reporting the method and results, as well as providing a brief discussion. This is done through comparing the photographic measure against selected scales designed to test different aspects of validity, as well as calculations of Cronbach’s alpha to assess reliability. Results demonstrate that the inclusion of the silhouette to enhance imagination and connectivity improved both validity and reliability compared to the other photographic presentation modes.

3.2 Enhancing empathic accuracy processes.

3.2.1 Imagination.

Two key features of empathic accuracy processes are imagination and the connection between observer and target. As discussed in chapter one, empathic processes rely heavily on an imagination component (see section 1.7.1), whereby the viewer takes on the perspective of the target in order to ascertain the emotions and thoughts they are experiencing. This imagination element can be seen in Davis’ (1996) multidimensional measure of dispositional empathy, where he includes both a perspective-taking and a fantasy subscales. The perspective-taking subscale is designed to ascertain the propensity for a person to take the perspective of another, thereby enhancing understanding of the target’s experiences. Likewise the fantasy subscale assesses how often or likely a person is to become involved with fictional characters, experiencing their feelings, sharing the character’s triumphs and failures (Davis, 1996).

3.2.2 Connectivity.

Connectivity between the observer and target is also important as it has been found that we are better able to empathise with whom we feel familiar (Krebs, 1975; Davis, 1996; Zagefka, et al., 2013) or similar to (Batson, et al., 2007). As a general rule, we are better able to identify the emotions of people who are known or close to us, compared to those that are not (Ickes, 1990;
Stinson & Ickes, 1992), therefore increasing the amount of interaction between the observer and target increases empathic accuracy (Marangoni, Garcia, Ickes, & Teng, 1995; Simpson, Ickes & Blackstone, 1995). Simply ‘imagining’ the other person’s situation, however, can also increase not only empathy processes, but also how much we value the other person’s welfare (Archer, Diaz-Loving, Gollwitzer, Davis, & Foushee, 1981; Batson et al., 2007; Batson, et al., 1995), in turn increasing our connection to the target. Increasing imagination processes, therefore, should have the dual effect of increasing empathy processes as well as increasing the connectivity between observer and target.

3.2.3 Facilitating both imagination and connectivity in photographic stimuli.

Although photographs are often used in studies of AR-EA abilities, there has been little research investigating how these imagination and connectivity processes can be facilitated or enhanced through the use of photographic stimuli. Some research has used instructions such as “imagine-self” / “imagine-other” conditions (see Batson, et al., 2003; Davis, et al., 1996; Davis et al., 2004; and Lamm, et al., 2007 for some examples). Imagining how another feels, as opposed to imagining how you would feel in their place, has been linked to increased empathy (Batson, Early, & Salvarani, 1997; Batson et al., 2003; Davis et al., 2004; Ruby & Decety, 2004). Therefore whilst the evidence suggests that ‘imagine-other’ instructions may facilitate imagination processes, with regards to photographic stimuli in particular, no other mechanism has been tested or used to achieve the same end.

In studies of empathic accuracy, connection to the target has been manipulated through varying the amount of interaction between participants (see Marangoni, et al., 1995; Simpson et al., 1995), or through the use of vignettes or descriptors that increase perceived similarity between the participant and target (e.g.: Krebs, 1975). Direct interaction is not possible with the use of photographs, whilst manipulation of similarity has returned mixed results. Batson et al. (1997) found that empathy/helping behaviours did not differ significantly between participants who were led to believe they were part of the same group and those that were led to believe otherwise. It
could be, however, that the manipulation of whether the person in need belonged to the same university or not was not an adequate manipulation of similarity. Again, no mechanism for increasing connectivity between observer and target has been found in relation to photographic stimuli specifically.

Photographs are not in and of themselves vehicles to automatically increase either imagination or connectivity, unless the photograph is of a person known to the participant, or involves a familiar memory. Without specific mechanisms to trigger imagination processes and to increase the connection between the observer and the target, it is questionable whether photographic instruments that do not employ other mechanisms to enhance these processes are actually measuring empathic accuracy abilities. Without something to facilitate these processes, participants may find it difficult to connect to the target, which may inhibit imaginary processes, in turn decreasing empathic accuracy. Therefore this current study will investigate if a photographic element, in the form of a male/female blacked-out silhouette, inserted into the photograph, can help facilitate empathic accuracy processes.

3.2.4 Presentation modes: full body versus head and shoulders only.

In chapters one and two it was argued that during the interpretation of another’s emotions, people do not just rely on the face for emotion cues, but instead consider cues from the entire body of the target (Aviezer, et al., 2011; Aviezer, et al., 2012; Nelson & Russell, 2011; Tracy & Robins, 2004). Traditional photographic stimuli, however, have often only included the head and shoulders of the target, on the assumption that the face is primary source of emotion interpretation (Eastwood, et al., 2001; Ekman, 2004; 2007; Ekman & Friesen, 1967).

Part of the aim of this thesis is the development of a new photographic measure to assess AR-EA abilities. In order to achieve this, different modes of presenting the target stimuli will be compared to each other. Three different photographic presentation modes will be compared: full body only (FB); head and shoulders only (H&S); and full body with silhouette (FBSil, the
imagination/connectivity element). In this way the results should reveal which presentation mode best assesses AR-EA abilities.

3.3 Study 1: Imagination and Connectivity

The aim of this first study is to investigate if inserting a specific stimulus element into the photographs designed to enhance imagination and facilitate connection between observer and target will impact AR-EA abilities. As per the literature presented in chapter one, both imagination and connectivity are important elements in empathic processes, and should facilitate empathy in participants. Using the photographs developed in chapter two, participants will be asked to identify the emotions of the targets in each of the photographs. The method and results of this study are presented below, along with a short discussion of the findings.

3.3.1 Exploratory questions, hypotheses, and testing conditions.

Although mostly exploratory in nature, some predictions can be made based on the findings from chapter two as well as previous research. It is expected that accuracy scores for emotion categories and intensity levels should follow similar patterns to those found in the previous chapter (see sections 2.4.1.2 and 2.4.2.2.1). Therefore neutral and happy should obtain the highest accuracy scores, whilst fear should attract the lowest scores. The lowest intensity level of emotional expression should attract the least accurate scores, and the most intense levels of expression, the highest accuracy scores.

In order to ascertain if different photographic presentation modes can impact emotion interpretation, three different types of photographic presentation were compared - full body only (with no silhouette; FB); head and shoulders only (H&S); and full body with silhouette (FBSil). Since the inclusion of a silhouette is a new component, with no research available, the analysis will be exploratory in nature; however, given that both imagination and connection seem to be facilitators of empathy processes, it is expected that those participants presented with the photographs containing the silhouettes (FBSil) will obtain higher accuracy scores compared to other participants.
In line with the research reviewed in Chapter two (see sections 2.4.1, 2.4.1.2, and 2.4.2.2.3), as well as the results obtained, it was hypothesised that female models would attract higher mean accuracy scores compared to male models; however, there should be no differences in emotion interpretation accuracy between male and female participants.

3.3.2 Method.

3.3.2.1 Participants.

A total of 187 participants were drawn from university students and members of the general public using snow-balling techniques, flyers, and canvassing in student common areas (such as Bush Court). Two participants (108 and 179) did not complete the questionnaires (with substantial sections left blank) and so were excluded, reducing the total from 187 to 185, leaving 116 females and 69 males. The mean age was 25.54, ranging from 17 – 63 years of age.

3.3.2.2 Materials.

Each booklet began with questions regarding demographic variables such as age, sex, education reached and SES (income). Education was measured on a sliding scale of increased achievement from 0 “Still in High School” to “Postgrad”. SES was indicated (i.e.: income brackets) ranged from “< 15k” (thousand) to “> 200k”. These were followed by 40 black and white photographs, 20 male and 20 female models. Photographs were randomised for each of the three different conditions - full body only (FB), head and shoulders only (H&S), or full body with silhouette (FBSil).

3.3.2.2.1 Creation of the ‘silhouette’

In order to enhance both imagination and connectivity a blacked-out figure, a silhouette, was inserted into each full body photographs. Two silhouettes were created, one male figure, and one female figure. These were created by morphing together features from several different models to create generic male and female shadowed figures (see Appendix H for a full description). Male participants were presented with a male silhouette, and female participants, a female silhouette.
Participants with the silhouette were instructed to ‘imagine that they are the silhouette in the photograph with the target’, and then instructed to interpret the target’s expressed emotion. It was hoped that this ‘imagination step’, would enhance both imagination processes and connectivity to the target, thereby facilitating AR-EA processes. The full body with silhouette presentation mode (FBSil) was tested against full body only (FB) and head and shoulders only (H&S) photographs. The exact procedure used for this testing has been outlined below.

3.3.2.3. Procedure.

All measures were presented in booklet form, with participants being assigned a condition randomly, although not all participants returned booklets (FB n = 66; H&S n = 56; FBSil n = 63). In both the FB and FBSil conditions each photograph was almost full page in size (20.77 x 13.84 cm), with only one image per page. In the H&S condition, the images were smaller (8.00 x 8.00 cm) with 3 photos displayed per page. Photographs were randomised for each of the different booklet conditions.

In the FB and H&S conditions participants were simply instructed to look at each image and circle the emotion word that most closely matched the photo. The emotion labels provided were: nothing, surprised, disgusted, fearful, angry, sad, and happy. Participants were also instructed not to overthink their responses, but to give their first impression. This is because in a social setting these judgements are generally made automatically (see Ekman & Friesen, 1974a; Knyazev, et al., 2010). In the FBSil condition an additional instruction was given below each image. “Imagine yourself IN this picture (you are the shadow). You are witnessing first-hand what the other person is feeling. Circle one of the words below that matches what that person is feeling.”

Delivery of the booklets varied slightly. Some booklets were filled out under supervision, some were posted to participants and returned via mail, and others were handed out to students in groups (Supervised n = 26; Postal n = 29; Other n = 130). Return rate for mailed booklets was 83.87%. A One Way ANOVA showed no significant differences in AR-EA scores for participants.
between the different testing environments \( F (2,162) = 2.48, p = .087 \), demonstrating that testing environment had no effect.

### 3.3.3 Results.

Participant responses were scored on a 0/1 basis where 0 indicated an incorrect response and 1 a correct response. This gave a maximum possible total AR-EA score of 40. Scores were also calculated in this fashion for each emotion category (Neutral, Surprise, Disgust, Fear, Angry, Sad, and Happy) and each intensity level (low, medium, and high). Initial descriptive statistics included the ‘neutral’ category, however, with only a possible score range of 0-2 was not included for the statistical analyses of emotion categories, intensity levels, model sex or participant sex in the results sections below.

Across all conditions participants scored an average of 24.67 out of a possible 40 (61.68%; SD = 3.66), with a range of 9-33. Individual emotion categories varied with the highest accuracies recorded for happiness and neutral categories, and the lowest for fear (see figure 3.1 below).
Mean percentages for the three different levels of intensity showed that the more subtly expressed emotions attracted lower accuracy scores than the more expressive versions (see figure 3.2 below).

![Graph showing mean percentages for intensity levels across all emotion categories except neutral](image)

*Figure 3.2. Mean percentages for intensity levels across all emotion categories except neutral*

Repeated measures ANOVA ($F(2,332) = 9.79$, $p < .001$, $\eta^2 = .056$) revealed no significant differences in mean accuracy percentages between low and medium intensity levels ($p = .344$), however, differences were significant between both the low and high ($p < .001$) and medium and high ($p = .016$) intensity levels.

**3.3.3.1. Photographic conditions.**

Figure 3.3 below shows that the H&S condition attracted less accurate mean scores ($m = 58.10\%$; $SD = 6.60$) than either the FB ($m = 64.25\%$; $SD = 7.70$) or FBSil ($m = 62.24\%$; $SD = 11.28$) conditions. A one-way ANOVA revealed that the differences were significant ($F(2,162) = 6.60$, $p = .002$, $\eta^2 = .082$) and large. Post hoc analyses showed that there was no significant difference between FB and FBSil conditions ($p = .444$), whilst the H&S condition attracted significantly lower accuracy scores than both FB ($p = .001$) and FBSil conditions ($p = .043$). It should be noted, however,
that Levene’s was violated in this test and therefore these results should be interpreted with some caution.

3.3.3.2. Model sex.

In most of the emotion categories female models attracted higher average accuracy percentage scores ($m = 76.71\%$; $SD = 13.25$) than male models ($m = 60.40\%$; $SD = 11.81$). Paired samples t-tests was used as participants responded to both male and female models. Results revealed that the difference overall was significant ($t(164) = 14.07, p<.001$; estimated cohen’s $d = 0.81$). This is a large effect, and as can be seen in figure 3.4 below, female models were more accurately rated compared to male models in all emotion categories except disgust, where male models attracted slightly higher averages.
3.3.3. **Participant sex.**

Female participants scored higher on average ($m = 25.57; SD = 3.04$) than males ($m = 23.10; SD = 4.13$). Independent Samples t-test confirmed that this difference was statistically significant overall ($t(163) = 4.40, p < .001; d = 0.68$), a moderate effect. Closer inspection revealed that this difference held only for the silhouette condition ($t(27.61) = 4.25, p < .001; d = 1.22$), however the effect in this case was very large. There were no differences in accuracy scores for the sexes in either the full body (FB) or the head and shoulders only (H&S) conditions.

Both male and female participants were better at reading female models than male models, as was indicated above. Male participants were more accurate in interpreting the female models overall, and at both low and high intensity levels ($p < .001$), but were better at interpreting male models at medium intensity levels of expression ($p <.001$), although this only held for the head and shoulders (H&S) condition. This pattern was also seen for female participants (all at $p < .001$ except for medium intensity level which was at $p = .004$). Therefore, with the exception of the head and
shoulders condition, at medium intensity, the emotions of females were more accurately identified by participants than the emotions of males.

3.3.4 Discussion.

The pattern of results achieved for the different emotion categories and intensity levels closely parallel results from the previous studies presented in chapter two. Similar to these findings, both happy and neutral categories attracted the highest accuracy ratings, and fear the lowest. Previous research has also found that the emotion happy is the most accurately recognised emotion (Hall & Matsumoto, 2004; Hoffman, et al., 2010; Montagne, et al., 2005; Slepian, et al., 2011; Thompson & Meltzer, 1964) and fear one of the most difficult to identify (Hall & Matsumoto, 2004; Hoffman et al., 2010; Montagne et al., 2005; Slepian et al., 2011).

Intensity levels also followed similar trends from those results obtained in chapter two (see sections 2.4.1.2 and 2.4.2.2.2), with similar effect sizes indicating a robust effect. As expected, those emotions expressed at the lowest level of intensity attracted the least accurate scores, whilst those at the highest level of expression intensity attracted the highest accuracy scores. This same performance pattern for different intensities of emotional expression has also been found in previous studies (Hoffman, et al., 2010; Palermo & Choltheart, 2004; Zuckerman, et al., 1976). The consistency in these patterns for both the emotion categories and emotion intensity levels between this study and the previous studies in chapter two, suggest good consistency of the AR-EA photographs.

Overall, AR-EA scores in this current study were lower compared to the second study conducted in chapter two. This is possibly due to the increased difficulty of the task. In the sorting-task in chapter two, participants were asked to not only categorise each photograph into an appropriate emotion category, but then to also sort those photographs into order of intensity for each model. This meant that participants had repeated exposure to individual models, allowing them to become familiar with their expressions, and compare expressions side by side. Also, participants were exposed to only 4 models (2 males and 2 females) compared to the current study.
in which six different models were presented. To make the task even more difficult, no one model was used for an entire emotion category or intensity level, meaning that participants were faced with the task of interpreting a different face for nearly every level of intensity for each emotion category. This increased task difficulty may explain the lower accuracy scores in the current study.

3.3.4.1. Photographic conditions.

The fully body only photographs (FB) attracted the highest mean scores, followed by the full body with silhouette (FBSil) condition, whilst the head and shoulders (H&S) photographs attracted significantly lower AR-EA scores than both. The large effect size obtained with this result indicates a very large difference between the different photographic presentations. This is interesting since much of the research conducted into emotion interpretations has relied upon only head and shoulders stimuli of targets in the photographs (Anitha, et al., 2010; Biehl et al., 1997; Ekman, et al., 1971; Hwang, et al., 2003; Langner et al., 2010; Matsumoto, 1992; Matsumoto & Ekman, 1989; Matsumoto et al., 2000).

Research has demonstrated that the face is not the only source of information when it comes to interpreting the emotions of others. Findings indicate that emotions can be identified from body cues alone (Atkinson, et al., 2004; Coulson, 2004; Dael, et al., 2012). Interpretation biases too can be evident from body cues only (Munoz, 2009). Other research suggests that both facial and body cues are necessary for the accurate identification of emotions (Nelson & Russell, 2011; Tracy & Robins, 2004). More importantly, Aviezer, et al. (2012) contest that even when told to, participants cannot separate the face from the body when identifying emotions in others. The current results seem to support this latest research, encouraging a rethink on the type of photographic stimuli that should be used in studies involving emotion identification tasks.

3.3.4.2. Model sex.

Participants found female model expressions easier to read compared to male models, in all photographic conditions and all levels of intensity, except for the H&S condition at medium
intensity. It is difficult to explain why this anomaly occurred in the H&S for medium intensity photographs as there is no research to draw upon that has compared the different presentation modes using different intensity levels of emotion. It should be noted that the effect size was again large, indicating that there is a real difference in participant’s abilities to read female versus male model expressions. More study comparing head and shoulders stimuli with full body stimuli, using emotions of varying intensity will be needed to confirm if this is specific to model sex influences or not.

Female models attracted higher AR-EA scores for all emotions except disgust, where male models had higher scores, although this difference was not significant. This is in line with previous research that has found that females are more expressive than men (e.g.: Buck, Miller & Caul, 1974; Buck, Savin, Miller, & Caul, 1972; Gross & John, 1995; Palermo & Coltheart, 2004; Wagner, et al., 1993; Zuckerman et al., 1975). One explanation for this may the social expectations and display rules that apply to the different gender roles of males and females. Research into gender roles and expression have found that women are expected, by others and themselves, to both experience emotions more intensely, and therefore express them more intensely (Hess et al., 2000; Kring & Gordon, 1998; La France et al., 2003), therefore explaining these results.

3.3.4.3. Participant sex.

Male and female participants had no significant differences in their accuracy scores in either the FB or the H&S photographic conditions. Female participants were, however, significantly more accurate than males in the FBSil condition, with an effect size indicating the difference to be larger than one standard deviation. Studies that have investigated differences in emotion accuracy identification between males and females reveal mixed findings. Some research states that there are no differences in the accuracy between the sexes (Hampson et al., 2006; Jorgensen & Howell, 1969), which reflects these current findings. Other research, however, has found that females are more accurate than men at reading emotions (Hall, 1978; Hall & Matsumoto, 2004), especially subtle
emotions (Hoffman et al., 2010). In no study is there evidence, overall, of men being better than females in this ability, except for specific emotions such as anger (e.g.: Wagner, et al., 1986).

When looking at male and female accuracy scores for the separate photographic conditions, however, it was found that males scored significantly lower than females in the silhouette condition. To refresh, the FB and H&S conditions required that participants simply look at the person in the photograph, and circle one of 7 available responses, to identify what emotion the target person was feeling/expressing. In the FBSil condition, there was the added step in the instruction set that participants first imagine that they were the silhouette in the actual photograph with the target. They were then asked to identify the emotion expressed by the target. The results seem to indicate that this added ‘imagination’ step impacted negatively, and significantly, on male participants.

This is counter to the expectation that the inclusion of the silhouette would enhance empathic accuracy processes, as recognition of another’s emotions theoretically, has a strong imagination component (Davis, 1980; Eisenberg & Miller, 1987; Hoffman, 1984). Previous research indicates that increasing the attention toward the target should increase empathy processes (Cialdini, Baumann & Kenrick, 1981). Whilst this held true for female participants, the same cannot be said for male participants who struggled to accurately identify the emotions of targets when imagining themselves as the silhouette in the photograph.

One explanation might be that the inclusion of the female silhouette for the female participants triggered female-gendered social role expectations, resulting in an increase in empathy and therefore empathic accuracy. Graham and Ickes (1997) posited the idea that if females were reminded of their stereotyped gender-roles, this increased both their empathy and their emotion interpretation abilities. Therefore, in this current study, the presence of another female in the form of a female silhouette may have triggered these gender-related expectations, influencing the results. Further and specific testing will be needed before this hypothesis can be confirmed.

Alternatively, it could be the case that females are better than males at imagining themselves into the photograph. Or it could be that females were simply more diligent at following
the instructions than the males. Whilst this might explain the slight increase in accuracy for female participants, it does not explain the decrease for males. One explanation may be that the inclusion of this ‘imagination’ step increased cognitive load, which then decreased accuracy. Davis, et al. (1996) found that increased cognitive load, via a memory task, negatively impacted perspective-taking abilities in participants. There is little research into the effects of cognitive load on empathy and upon empathic accuracy specifically, but the research generally indicates three different sets of findings: a) that high cognitive load can impact automatic empathic processing; b) that high cognitive load can reduce experienced emotion; and c) that high cognitive load can reduce emotion processing mechanisms. Hodges and Wegner (1997) argue that increased cognitive load can interfere with our ability to empathise with others. Rameson, Morelli, and Lieberman (2012) found in their neuroimaging study that cognitive load reduced empathy-related activity areas of the brain for participants low in dispositional empathy, however, individuals high in dispositional empathy, when asked to empathise, had increased empathic responding, even under high cognitive load. Van Dillen, Heslenfeld, and Koole, (2009) found that emotion recognition and processing areas of the brain reduced in activity when participants were engaged in cognitively demanding (arithmetic) tasks. Participants also reported feeling less negative emotions in response to emotionally negative stimuli the more demanding the arithmetic task (van Dillen et al., 2009). Kron, Schul, Cohen, and Hassin (2010) also found that under high cognitive load, participants reported experiencing less positive and negative emotions. A study by Knyazev, Slobodskoj-Plusnin, and Bocharov (2010) found that men processed facial emotion information in an ‘early processing stage’, implying that this processing is automatic, which may be impacted by cognitive load.

Mitchell, Nakic, Fridberg, Kamel, Pine, and Blair (2009) found that the amygdala region of the brain in participants, responsible for emotion processing, was less active in the high cognitive demand task compared to easier tasks. Another study found that increased working memory load inhibited accurate identification of emotion in a single face (Phillips, Channon, Tunstall, Hedenstrom,
& Lyons, 2008). In judgements of whether two faces displayed the same emotion or not, however, manipulations of working memory load had no impact (Phillips et al., 2008).

The conclusion, according to this body of research, is that increased cognitive load can inhibit or enhance empathic-based processing and responding. Overall, it seems that increased cognitive load has the potential to interfere with the accurate identification of emotions in others; however, there is also research to suggest that processing of some specific emotions, such as fearful and angry faces, may be facilitated in instances of high cognitive load (Fox, Yates, & Ashwin, 2012; Tracy & Robins, 2008; Yates, Ashwin, & Fox, 2010). Tracy and Robins (2008) argue that if emotion recognition is innate, based on evolutionary concepts, then emotion identification processes must be automatic and therefore not susceptible to cognitive load influences. In their study, participants were able to quickly and accurately recognise the stimulus emotions of happy, sad, surprised, disgusted, and angry even under conditions of cognitive load. Fear, however, required more deliberation and elicited reduced accuracy under cognitive load (Tracy & Robins, 2008).

The picture then is confusing and unclear. A recent study by Morelli and Lieberman (2013) has clarified the matter somewhat. They found that under high cognitive load areas of the brain associated with empathy and social cognition showed reduced activity, whilst other areas of the brain, the septal area and ventral AI (Anterior Insula) showed heightened activity during empathising tasks for all emotions and all cognitive load conditions (Morelli & Leiberman, 2013). This would indicate that there are both specific processing areas that are negatively impacted by cognitive load, and other areas that are not impacted at all.

In relation to the current study one explanation might be that in men, empathic accuracy processes are more effortful and therefore prone to more interference from cognitive load effects; however there is no research that has investigated the effects of cognitive load on emotion recognition in men and women separately. This is an interesting and intriguing result which could shed light on behaviours such as violence in highly social situations (e.g.: clubs and pubs). If men are susceptible to inaccuracies under conditions of high cognitive load, highly social situations may tip
the cognitive resource balance, resulting in more misidentification of emotions in others. This is, however, mere conjecture. Further study and replication of this result will be necessary to further this hypothesis.

3.3.4.4. Limitations.

Although the presentation order of the photographs were randomised for each photographic condition (FB, H&S, and FBSil), they were not randomised within each condition. Whilst this might not be a problem for the full body versions of the photographs, the head and shoulders presentation may have been biased by the presentation of 3 photographs per page, instead of one per page as in the other conditions. This may have set up biased responding for some of the photographs and may go some way to explain the lower accuracy ratings achieved.

3.4 Study 2: Psychometric Evaluation

In chapter two it was established that participants could successfully interpret the emotional expressions of the targets in the photographic stimuli, and do so across varying levels of expression intensity. There was some consistency in responding between the first and second study in chapter two indicating that the photographs could be interpreted with a measure of reliability. In the development of any new measure, however, validity and reliability need to be established with more formal processes. The aim of this section, therefore, is to conduct validity and reliability testing on the photographs selected from chapter two to forward the development of the AR-EA photographic measure.

It should be established with any new measure that it is measuring what it is purporting to, which is referred to as construct validity (De Vellis, 2012; O’Sullivan, 1982). As O’Sullivan (1982) points out, simply getting other people to judge the photographs is not enough. It needs to be established that the intended emotions match those that were actually expressed. In this case, construct validity was established in chapter two through the use of an accuracy criterion, where
models rated themselves on how well they believed they had captured each emotion (see Section 2.3.1). De Vellis (2012) also recommends that construct validity, specifically convergent and discriminant validity, be established through the testing of the measure against other theoretically related constructs.

Reliability is a measure of how consistently an instrument measures the construct, in this case AR-EA abilities. This means that the score on the measure should not change, unless another factor has caused a change in the construct being measured (De Vellis, 2012). Cronbach’s alpha is a common, practical and reliable method of establishing reliability (De Vellis, 2012), and establishes the internal consistency of the measure. In other words, that the items within the measure are highly intercorrelated (De Vellis, 2012).

The aim of this current study, therefore, was to establish which of the different presentation formats of the photographs attracted the highest validity and reliability scores. Specifically construct validity, (convergent validity and discriminant validity), and internal consistency (reliability) were tested. The validity and reliability testing will confirm which photographic presentation mode will be adopted as the new AR-EA photographic measure. The following study was conducted concurrently with Study 1: imagination and connectivity, above, therefore some of the descriptive statistics for participants, materials and procedure are similar to the above study. The results will be reported and a short discussion presented highlighting the outcomes.

3.4.1 Method.

3.4.1.1 Participants.

A total of 185 university students and members of the general public participated in the study. They were recruited using snow-balling techniques, flyers, and canvassing in student common areas (such as Bush Court). The mean age was 25.54 years, ranging from 17 – 63 years of age.
3.4.1.2. Materials.

3.4.1.2.1 Scale Selection.

The following scales were selected for testing convergent and discriminant validity. Research is presented establishing the expected relationship of each scale with AR-EA scores. Psychometrics for each scale are presented in table 3.1 below.

Convergent validity is based on the concept that theoretically similar constructs should be positively related to each other (O’Sullivan, 1982). As was established in chapter one, affective empathic accuracy is considered a part of the emotional intelligence construct (Mayer & Geher, 1996), which Salovey and Mayer (1989) argue is also a part of social cognition and social intelligence. Empathic accuracy for the interpretation of other’s emotions has strong links to both social information processing concepts (Ickes, et al., 1990), and social interactions (Davis, 1996; de Vignemont & Singer, 2006; Kohut, 1980; Riggio, 1986; Riggio, Tucker, & Coffaro, 1989; Schwartz, 1994). It was expected, therefore, that AR-EA scores would be positively correlated with social intelligence. In order to establish convergent validity the Tromsø Social Intelligence Scale (TSIS; Silvera, Martinussen, & Dahl, 2001) was selected. This instrument was developed to be a simple, but comprehensive, measure of social intelligence capturing three different dimensions of the construct. Social Information Processing (SP) includes cognitive-based processes such as the ability to understand others and predict their behaviour (Gini, 2006). Social Skills (SS) involves behavioural aspects such as fitting into new social situations and social adaptation (Gini, 2006). Finally, Social Awareness (SA) measures the ability to understand one’s own impact on others as well as understanding other’s choices and motives. The TSIS consists of 21 statements such as “I can predict other people’s behaviour”, and “I have a hard time getting along with other people” (reversed), which are assessed on a seven point Likert scale from 1 “Describes me extremely poorly” to 7 “Describes me extremely well”.

Discriminant validity is established by ensuring that the scores obtained on the instrument are not related to scores attained from other unrelated constructs (O’Sullivan, 1982). To do this,
two different scales were chosen. Empathic disposition should be unrelated to AR-EA abilities, whilst Narcissism should be negatively related to AR-EA scores.

The shortened version of Davis’ (1996) Interpersonal Reactivity Index (IRI) was developed to measure dispositional empathy by Zander (2011 unpublished, also see Lauterbach & Hosser, 2007, for a similar instrument). This is a 14 item scale assessed by a four point Likert scale (1-4) from 1 “Never” to 4 “Always”. It contains questions such as “I feel concerned for people having a hard time” and “I find it hard to see the other guy’s point of view” (reverse question). The shortened version retains two of the original subscales – Empathic Concern (EC) and Perspective-Taking (PT). Davis (1996) acknowledges that the IRI measures only tendency to empathise which may have no bearing on the actual ability to do so. As has already been demonstrated (see Chapter 1, section 1.4.1.4) AR-EA abilities are not synonymous with dispositional empathy, with research demonstrating little or no relationship between the two (Eisenberg-Berg & Lennon, 1980; Hall 1984; March, Kozak, & Ambady, 2007; Zaki et al., 2009). Therefore it is expected that IRI scores should share no statistical relationship with AR-EA scores.

The next scale was the Narcissistic Personality Inventory (NPI-16). The original NPI developed by Raskin and Terry (1988) consisted of 40 items. The NPI-16 consists of 16 opposing statements (narcissistic and non-narcissistic responses) where participants are asked to select one of the statements in the pair that best describes them. One example is “I hope I am going to be successful” (non-narcissistic response) versus “I am going to be a great person” (narcissistic response). The NPI is considered to be a good standard measure of subclinical Narcissism (Williams, Nathanson, & Paulhus, 2010). According to Pincus, Ansell, Pimentel, Cain, Wright, and Levy (2009), and Besser and Zeigler-Hill, (2011), Narcissism is associated with self-enhancement strategies that are maladaptive, including overly positive self-images and exploitation of others. Narcissism can also involve interpersonal and relational problems (Pincus et al., 2009) as well as self and emotional dysregulation (Besser & Zeigler-Hill, 2011). Narcissism has been demonstrated to have a negative relationship with empathy (Bushman, et al., 2003; Ehrenber, Hunter, & Elterman, 1996; Watson,
Grimsham, Trotter, & Biderman, 1984; Watson & Morris, 1991) and in particular with empathic accuracy tasks such as identifying the emotions of other (Ames & Kammrath, 2004; Baskin-Sommers, Krusemark, & Ronningstam, 2014). Wai and Tiliopoulos, (2012) also found deficits in facial identification with narcissistic individuals, but only for Angry expressions. Given the findings in the literature we would expect that the NPI-16 will be negatively correlated with AR-EA scores.

The final scale to be included was a social desirability scale. A socially desirable response is one in which participants present themselves in an overly-favourable light and is a constant problem when using self-report measures such as the ones described above (Fischer & Fick, 1993). Therefore it is important to have a mechanism that can detect the presence of any social desirability in participant responses, so that results can then be analysed with this in mind. In order to assess social desirability the short-form of the Marlowe-Crowne Social Desirability Scale (MCSD) was used. This was developed by Rudmin (1999) to reduce the Crowne and Marlowe’s (1960) original instrument from 33 to just 10 items. Participants simply respond either “true” or “false” to each of the 10 statements, which include “I sometimes try to get even, rather than forgive and forget” and “I have almost never felt the urge to tell someone off”. The results of this scale are then correlated with the other scales. If no social desirability is present, there should be no relationship, and the results obtained can be reported with confidence.

It is expected that the construct validity testing will return results as outlined earlier in the chapter (section 3.4): that the shortened IRI (dispositional empathy) will have no relationship with AR-EA score; that NPI (16; Narcissism) will have a negative relationship with AR-EA scores; and that TSIS will have a positive relationship with AR-EA scores.
Table 3.1

Cronbach's alpha levels for the 3 scales chosen for validity testing, as reported in their respective publications: TSIS, Silvera, et al. (2001); Shortened IRI, Zander (2011); NPI-16, Ames, et al. (2006).

<table>
<thead>
<tr>
<th>SCALE:</th>
<th>Subscales</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TSIS</td>
<td>SP</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SS</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SA</td>
<td>0.79</td>
<td></td>
</tr>
<tr>
<td>Shortened IRI</td>
<td></td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>NPI-16</td>
<td></td>
<td>0.72</td>
<td></td>
</tr>
</tbody>
</table>

3.4.1.2.2 Testing booklets.

Each booklet contained 40 black and white photographs portraying 20 male and 20 female models expressing six different emotions at three different intensity levels. The photographs included 2 neutral expressions (no emotion, one male and one female), and 2 ambiguous expressions (one male and one female). Photographs were randomised for each of the three different conditions - full body (FB), head and shoulders only (H&S), or full body with silhouette (FBSil). The booklets also contained questions regarding demographic variables such as age, sex, education reached and SES (income).

The booklets ended with four standardized scales as discussed above: the shortened IRI (Zander, 2011); the NPI-16 Narcissism scale (Ames et al., 2006); the Tromsø Social Intelligence Scale (TSIS; Silvera et al., 2001); and the Marlowe-Crowne Social Desirability Scale (MCSD; Rudmin, 1999). Psychometric properties for these measures can be found in table 3.1 above. (See Appendix I. for an example of a booklet and Appendix J. for the scales used)

3.4.1.3 Procedure.

Participants were randomly assigned a photographic presentation condition (FB, H&S, or FBSil). Full body photographs were presented one per page, whilst head and shoulders photographs
were presented three per page. Photographs were randomised for each of the different booklet conditions. (For more details on this procedure please refer to section 3.3.2.3 above.)

In the FB and H&S conditions participants were simply instructed to look at each image and circle the emotion word that most closely matched the photo. The emotion labels provided were: nothing, surprised, disgusted, fearful, angry, sad, and happy. In the FBSil condition an additional instruction was given below each image. “Imagine yourself IN this picture (you are the shadow). You are witnessing first-hand what the other person is feeling. Circle one of the words below that matches what that person is feeling.”

The four psychometric scales directly followed the photographs: Davis’ (1996) Interpersonal Reactivity Index (IRI); Narcissistic Personality Inventory (NPI-16; Ames et al., 2006); Tromsø Social Intelligence Scale (TSIS; Silvera et al., 2001); and Marlowe-Crowne Social Desirability Scale (MCSD; Rudmin, 1999).

3.4.2 Results.

Reliability was established using Cronbach’s alpha for each photographic condition. It was found that the ambiguous stimuli (1 male and 1 female photograph) detracted from alpha levels and so were left out of the reliability calculations. The FB condition attained an alpha level of only 0.343, which could not be radically improved with the removal of items. An alpha score of .021 was obtained for the H&S condition with the removal of one item (highest intensity of Disgust, male model). The FBSil condition returned an alpha level of 0.686. With the removal of one or two items this could be improved as high as 0.7, however, given that the items in the test vary in difficulty, and that Cronbach’s alpha will be reflective of only the most difficult items (Novick & Lewis, 1967; Lord & Novick, 1968) a level of 0.686 is considered very acceptable. Therefore, of the three different presentation modes, the silhouette condition was the most reliable in measuring AR-EA abilities.

Given that the FBSil condition was the only one to reach acceptable reliability levels, the validity testing was conducted only on the FBSil condition. Calculations were done for each scale
resulting in a total scale score and subscale scores for the shortened IRI, TSIS, and NPI(16) (see section 4.3 on scale selection for more detail). Table 3.2 below shows the average scores obtained by participants on each scale in the FBSil condition. For context, possible minimum and maximum possible score for each scale were included.

Table 3.2

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Min. Possible</th>
<th>Max. Possible</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shortened IRI</td>
<td>61</td>
<td>14</td>
<td>56</td>
<td>41.62</td>
<td>5.20</td>
</tr>
<tr>
<td>TSIS</td>
<td>53</td>
<td>21</td>
<td>126</td>
<td>63.15</td>
<td>13.95</td>
</tr>
<tr>
<td>NPI (16)</td>
<td>57</td>
<td>0</td>
<td>16</td>
<td>4.17</td>
<td>2.70</td>
</tr>
<tr>
<td>MCDS</td>
<td>61</td>
<td>0</td>
<td>10</td>
<td>5.21</td>
<td>2.01</td>
</tr>
</tbody>
</table>

As expected there was no significant relationship between AR-EA scores and the shortened-IRI \((r(55) = .11, p = .423)\). None of the Shortened IRI subscales were significantly correlated with AR-EA scores either. There was a significant, moderate, negative relationship between AR-EA scores and NPI (16) \((r(52) = -.453, p = .001)\), meaning that participants reporting high levels of narcissism obtained low AR-EA scores. Unexpectedly, the TSIS was not significantly related to the AR-EA scores \((r(48) = -.164, p = .255)\). There was one significant, weak correlation between AR-EA scores and the TSIS subscale of Social Awareness (SA; \(r(51) = -.30, p = .030)\). Surprisingly this was a negative relationship meaning the higher a participant reported having Social Awareness, the lower their accuracy scores for interpreting emotions.

Social desirability, as measured by the MCSD \((r(55) = -.08, p = .541)\) was not correlated with total AR-EA scores. There were also no significant relationships between social desirability and the shortened-IRI \((r(57) = .18, p = .174)\), or the NPI(16) \((r(55) = -.12, p = .354)\). There were significant, moderate positive relationships with total TSIS scores \((r(57) = .42, p = .001)\) and the perspective-
taking subscale of the shortened-IRI ($r(59) = .46$, $p < .001$), indicating that participants may have been presenting a more acceptable self-image. This may explain the finding of no relationship between AR-EA scores and TSIS, as participants may have been over-estimating their own social abilities, this lack then being reflected in the emotion identification task. The non-significant relationship with the shortened-IRI may also suffer from the same conclusion meaning that this result may have to be interpreted with some caution.

### 3.4.3 Psychometric Evaluation – Discussion.

Construct validity was tested using three different scales, intended to test both discriminant and convergent validity. It was hypothesised that AR-EA scores would not be related to dispositional empathy as proposed by previous research (Eisenberg-Berg & Lennon, 1980; Hall, 1984; March, et al., 2007; Zaki et al., 2009). This hypothesis was supported. This confirms that empathy is comprised of two individual constructs - disposition and accuracy. This strengthens the arguments made in the introduction of this thesis that both aspects of empathy need to be measured in order to gain full and proper insight into empathy processes (see Chapter 1 section 1.4.1.4). There are, however, few studies that have set out to directly compare dispositional tendencies with accuracy abilities, and more research is needed to understand how these two aspects of empathy work together.

Narcissism, as measured by the NPI(16), was hypothesised to have a negative relationship with empathic accuracy. This was based on the premise that narcissistic traits include self-focus and self-enhancements (Pincus et al., 2009; Besser & Zeigler-Hill, 2011), which would interfere with both empathy (Bushman, et al., 2003; Ehrenber, et al., 1996; Watson, et al., 1984; Watson & Morris, 1991) and emotion identification processes (Ames & Kammrath, 2004; Baskin-Sommers, et al., 2014; Wai & Tiliopoulos, 2012). The result was a significant negative relationship, with participants who reported possessing more narcissistic personality traits obtaining lower AR-EA scores than those who reported having less narcissistic tendencies. Given the results for both dispositional empathy
and narcissism, the discriminant validity of those photographs containing the silhouettes, as a measure of AR-EA abilities is confirmed.

Convergent validity was tested against the social information processing TSIS scale. Due to the close association of empathic accuracy processes with social information processing concepts (Davis, 1996; de Vignemont & Singer, 2006; Kohut, 1980; Riggio, 1986; Riggio et al., 1989; Schwartz, 1994), it was predicted that a positive relationship between the two measures should be revealed. The result unexpectedly, was a slightly negative non-significant relationship. Looking at the subscales of the TSIS no relationship was found between AR-EA scores and Social Information Processing nor Social Skills subscales, but there was a significant negative relationship with Social Awareness (SA). The SA subscale is purported to measure how aware one is of their impact on others, as well as how well others are understood, including their choices and motives (Silvera, 2001). The current results indicate that participants in this study may have over-estimated their abilities to interpret other’s emotions.

One explanation is that the participants reporting on this measure may not be as socially aware as they believe themselves to be. Another explanation may be found with the significant positive correlation between social desirability and TSIS. Participants, who actively promoted a more positive self-image, rated themselves higher in terms of social intelligence; however, this was not supported in the emotion identification task, which resulted in those participants who rated themselves as high on levels of social awareness, obtaining the lower AR-EA scores. The final explanation may come from research that indicates that participants, across a range of abilities, are not able to accurately predict their own performance. This negative relationship has been found for verbal and numeracy abilities (DeNisi & Shaw, 1977), intelligence and scholastic achievement (Mabe & West, 1982), and more specifically for self-reports of emotional intelligence and performance measures of the same (Brackett, Rivers, Shiffman, Lerner, & Salovey, 2006; Swann & Gill, 1997). Therefore it seems that participants, as a general rule, are not as enlightened regarding their own abilities as they might think. This could explain the unexpected discrepancy between the TSIS and
AR-EA scores. So whilst social intelligence theoretically should be positively correlated with emotion identification abilities, self-report measures to find convergent validity with a performance measure of the same are problematic.

Another consideration might be the complexity of the constructs under investigation. Social intelligence encompasses many different cognitive, behavioural and social processes, which might be difficult to adequately capture. Affect Recognition-Empathic Accuracy (AR-EA) abilities are complex, involving different empathetic and social processing mechanisms and constructs that interact on many different levels. In a study looking at the difficulties with replicating consistent findings for correlates of empathy, Ickes and colleagues (2000) found there were few individual difference correlates with empathic accuracy, concluding that the ‘best candidate’ for prediction was verbal intelligence, with a caveat that more research was needed. Therefore it could be that the two concepts are too complex to compare using such simple measures. Further testing of the AR-EA test will need to be done in order to confirm its convergent validity.

### 3.4.3.1 Limitations

A limitation of this current study was the fact that convergent validity was difficult to establish. This may be due to the variation in emotion intensities, which also had a negative impact on reliability measures, although the level reached was considered acceptable. The self-report measure for social intelligence did not accurately reflect participant performance in identifying the emotions of others, therefore further testing against other instruments or tasks will be necessary to establish convergent validity.

Another limitation may be that the instructions in the silhouette condition, being slightly different to the full body version. It is feasible that the instructions with the added imagination directive may have resulted in higher accuracy scores, and not the silhouettes themselves. This is something that will require specific study to clarify.
The results of the different photographic presentations lend some support to the validity of the photographs assessing AR-EA abilities. As per the literature and theory behind empathic accuracy processes, full body presentations were superior to head and shoulders only presentations. Also the inclusion of a silhouette in the photographs to enhance imagination and connectivity resulted in the best reliability scores of the three different presentation modes. This suggests that the photographs containing the silhouettes are indeed assessing affect recognition-empathic accuracy (AR-EA) abilities, and are therefore the best stimulus set for assessing this type of empathic accuracy.

3.5 Conclusion

The AR-EA scores obtained by participants in this current study were similar to those obtained in chapter two, for both the emotion categories and emotion intensity levels. Emotional expressions of the female models were more easily discerned by participants compared to the male models, and there were no differences in AR-EA scores between male and female participants (except in the silhouette condition). These results are not only in line with the literature, but are also consistent with the findings in chapter two attesting to the consistency of the AR-EA photographs.

The three different presentation modes of the photographs revealed some significant differences in AR-EA scores with both of the full body modes (FB and FBSil) attracting higher AR-EA scores than the head and shoulders only condition. This is an important finding as many of the photographic stimuli used in the research include only the head and shoulders of targets and not the full body. In much of the research there has been no direct comparison of head and shoulders and full body modes. Generally these have been tested separately – either head and shoulders, or body, but not full body. This current result does question the validity of research that has used only head and shoulders modes to assess empathic accuracy abilities, and points to the necessity of further testing using full body stimuli.
Of the three different modes tested, both the full body conditions obtained the highest mean AR-EA scores; however, Cronbach's alpha levels revealed that the silhouette condition was more reliable in this measure than the full body condition. This supports the theoretical basis of the inclusion of the silhouette to facilitate empathic accuracy processes for participants. Interestingly male participants did not perform as well as female participants when the silhouettes were included in the photographs. Whilst a range of explanations have been posited the literature remains unclear on the issue. Therefore only replication of this result can shed more light on this particular phenomenon. If replicated it has the potential to explain some of the antisocial behaviours perpetrated by males in highly social environments.

Construct validity for the AR-EA test was only partially established. Divergent validity was supported with no relationship being found with a dispositional measure of empathy, and a negative relationship being found with narcissism. Convergent validity was not established, as no significant relationship was found between a measure of social intelligence and AR-EA abilities. Several explanations were offered as to why this might be the case, the prominent one being that self-report measures do not accurately reflect actual performance in a similar domain. This implies that only another performance measure could be used to establish convergent validity, however, as discussed above (section 3.4.3), this may not be a practical solution. Further testing should help establish both reliability and validity.

Overall, the conclusion drawn from the psychometric evaluation is that the FBSil mode, those photographs that included a silhouette, are superior in both validity and reliability to both the head and shoulders only (H&S) mode, and the full body only (FB) mode. If one accepts that rationale that the silhouettes in the photographs can facilitate empathic processes for participants, then the results reported here would support that premise. For this reason it is proposed that the FBSil mode be adopted as the AR-EA photographic measure, designed to assess empathic accuracy abilities in participants.
The aims of the next set of studies will be to investigate the effect of social context on AR-EA abilities. Both the full body (FB) and full body with silhouette (FBSil) conditions will be used in the next series of tests in an effort to replicate some of the current findings. Social context will be embedded into the photographs in the form of different social settings as backgrounds: a kitchen scene; a bar scene; and a neutral background. The purpose of this is to see if manipulation of social context has any impact on emotion identification abilities, which would eliminate the need for heavy verbal loads in testing with the traditional use of vignettes.
Chapter 4

4.1 Overview

The purpose of this chapter was to continue with the exploration of how different elements within photographs might be manipulated to impact AR-EA abilities. Results from chapter three indicated that the insertion of a silhouette to enhance imagination and connectivity in participants successfully impacted AR-EA abilities, resulting in a more valid and reliable measure of AR-EA. The current chapter, therefore, looked at the potential influence of social environment. To achieve this, different social settings were inserted into the full body versions of the photographs developed in chapter two, allowing an assessment of which social settings would impact AR-EA scores, and to what extent this occurred.

As in chapter three, the current chapter also has a second aim. Psychometric evaluation of the different photographic presentation modes revealed that the full body photographs that included the silhouettes were the most valid and reliable. The conclusion was that this particular set of photographs represented the best measure of empathic accuracy abilities, and for this reason was chosen as the AR-EA photographic measure. The reliability results from chapter three, however, revealed that the Cronbach’s alpha level was borderline. It was therefore decided to conduct a second round of reliability testing using a test-retest methodology. This chapter then, like chapter three, is split into two sections. The first section will present Study 1: Social Context, an investigation into the effects of social setting on AR-EA abilities. This will include the method and results, followed by a discussion of the outcomes and implications of the findings. The second section, Study 2: Reliability – Test-rest, will present the method and results of the test-retest reliability of the AR-EA photographic measure. A short discussion will follow regarding the outcomes of this psychometric evaluation.

To begin, this chapter briefly reviews the research regarding the impact social context can have on empathy. The research demonstrated that our interpretations of others can be influenced by a number of factors, (see chapter 3 section 3.2 for an overview of some of these factors) including
the environment in which the emotion display has occurred. Emotional expressions are governed by display rules or norms (see Ekman & Friesen, 1969, as cited in Kupperbusch, 1999, also see Ekman, 1999b), which will influence the types of emotions expressed (Ekman, 1999b; Fridlund, 1991) as well as the intensity with which they are expressed (e.g.: Jakobs, et al., 2001). These influences on expression should therefore directly impact the accuracy with which these emotions are detected in others (see chapter 3 section 3.3.4 for a discussion on the influence of intensity levels on emotion identification). There has been little research investigating the direct impact social setting may have on AR-EA abilities. Due to this lack of literature in the area, it should be noted that this study was largely exploratory in nature.

4.2 Empathic Accuracy and Display Rules

Social context can govern which emotions we express and how intensely we will express them. The rules which govern the expression of emotions in different social situations are called display norms or display rules (See chapter 1, section 1.7.2)(Ekman & Friesen, 1969, as cited in Kupperbusch, 1999, also see Ekman, 1999b). These rules can be governed by culture (e.g.: Glikson & Erez, 2013) or gender roles (e.g.: Santiago-Menendez & Campbell, 2013).

Display rules can impact how and what emotions we express. For instance in the presence of others not from our own culture, we will suppress negative emotions and more freely express positive emotions (Glikson & Erez, 2013). Gender roles too have an impact, governing which emotions are more suitable for each of the sexes to display. Santiago-Menendez and Campbell (2013) found that girls admitted to crying more often that boys when sad, and also that only girls, and not boys, admitted to crying when angry. Emotional expressions will also be influenced by the presence of others, and if they are familiar (Fridlund, 1991) or strangers (e.g.: Jakobs, et al., 2001). Therefore in the social world, we deal with facial expressions of emotions that are not always intense, but are often subtle or masked (see Chapter 1, section 1.7.2 for a more detailed discussion). As demonstrated in Chapter 3 (section 3.3.4) varying the level of expression intensity of the emotions impacts the ability for others to accurately interpret that emotion. The introduction of
social context in the form of different photographic backgrounds, against which the models are presented, should interact with both the type of emotion expressed, and the intensity of that expression, impacting the accuracy with which participants are able to interpret those emotions.

4.3 Empathic Accuracy involves the Interpretation of Multiple Cues

Although much of the research into empathic accuracy involves only the interpretation of other’s facial emotions, in reality, people will take advantage of other cues if they are made available to them. In chapter three it was found that although the face and body may provide different types of emotion information (see section 3.2.4), people will take cues from both in order to interpret another’s emotions (see chapter 3, section 3.2.4 for review and section 3.3.3 for results). Research indicates, however, that this interpretation is not made in a social vacuum rather the environmental and social cues available will also have an influence on these abilities.

Research has demonstrated that background cues influence our interpretation of others. The colour of the background against which interpretations are made has been found to facilitate or inhibit the correct identification of facial emotions (Silverthorne, Gibson, Micklewright, & O’Connell, 1975; Young, Elliot, Feltman, & Ambady, 2013), whilst the physical placement of figures in a picture can also influence our interpretations of the emotions the characters may be experiencing (Marian & Shimamura, 2011). Some research has also suggested that contextual cues provided by the environment/social situation may be more important for interpretation processes than the expressed emotions themselves (Walbott, 1988; Carroll & Russell, 1996).

4.4 Empathic Accuracy and Social Context – Cognitive Processes

The ability to accurately interpret the emotions of others is also influenced by the observer’s own judgment(s) of the target and their circumstances. The research indicates that our memories and experiences (Hoffman, 1984; 1987; Wyer & Srull, 1986), prior and cumulative knowledge about the social context (Gesn & Ickes, 1999; Sze, et al., 2012; Wyer & Srull, 1986) and judgements about the target and their motives (De Melo, et al., 2013; Hoffman, 1984; 1987) can all influence AR-EA
abilities, facilitating or inhibiting them (see Chapter 1.7.2). Despite the fact that social context seems to have such a strong influence on empathic accuracy abilities, there seems to be a lack of research in the area. The study in this chapter is designed to look at the potential impacts of social context on AR-EA abilities, with the aim of shedding more light on the mechanisms at play and the way in which different social contexts can either facilitate or inhibit interpretations of emotions.

4.5 Study 1: Social Context

The aim of this study was to investigate if, in the presence of different social settings presented in the backgrounds of the photographs, AR-EA abilities in participants would be impacted. Three different backgrounds were inserted into both the full body (FB) and full body with silhouette (FBSil) photographs – a kitchen; a bar (in a tavern/pub); and neutral (no background). These backgrounds were chosen for their anticipated familiarity with a university student cohort and the amount of social experience participants may have had in similar social environments, making them instantly recognisable to participants. As was demonstrated in the literature reviewed in chapter 1 (section 1.7.2) it is predicted that AR-EA abilities will be influenced by the different social contexts in which the photographs will be presented in, although the direct nature of this influence is unknown due to a lack of research in the area, with no research of this specific nature having been conducted previously.

4.5.1 Method.

4.5.1.1 Participants.

A total of 213 university students ranging in age from 17 – 70 years of age (m = 25.19), with 150 females and 63 males participated in the study. Students were recruited via posters and approached in lunch and common areas. Most had indicated that they had undertaken some university of TAFE based education, earning below $15,000 per year.
4.5.1.2. Materials.

Booklets consisted of 38 black and white photographs of male and female models posing six different emotions (happy, sad, fear, angry, disgust, and surprise) and one neutral pose. There were two photographic conditions – full body (FB) and full body with silhouette (FBSil). The silhouette photographs were designed to enhance imagination processes in participants, as well as increase the connectivity between the participant and the target in the photograph. This technique was described in more detail in chapter three (see section 3.3.2.2.1, and also see Appendix H). Each photograph was displayed on one A4 page of the booklet with instructions and response categories below each photograph (nothing; surprised; disgusted; fearful; angry; sad; happy). The booklets also contained demographic information including age, sex, education and income levels. In the previous study (chapter three) two additional ‘ambiguous’ photographs were included for testing, however it was found that these added nothing to the overall instrument, and in fact were a potential source of bias, and were therefore left out for the current testing.

Three different backgrounds were included in both the FB and FBSil photographs: a kitchen; a bar (in a tavern); and neutral. The neutral background was from the original photographs were taken against a grey background. The tavern and kitchen backgrounds (see Appendix K.) were converted to black and white and inserted behind the photographs using the software Photoshop. This resulted in nine different booklets overall: FB with kitchen, bar, and neutral; FBSil (male silhouette) with kitchen, bar and neutral; and FBSil (female silhouette) with kitchen, bar, and neutral. In each different booklet photographs were randomised to avoid ordering effects. The range of possible scores for each booklet condition was from 0 to 38.

4.5.1.3. Procedure.

The method used was similar to that presented in chapter three (see section 3.3.2.3). Participants were approached during lunch hours over the weeks of second semester (2013), in Bush Court, an informal lunch space frequented by students. Other students were approached in a lecture theatre. Participants were informed about the study and asked if they wished to participate.
Participants were also offered the chance to win one of 3 $50.00 gift cards for their participation (drawn by random ballot). Those who volunteered to fill in a booklet were asked for their consent and then given a booklet to fill in (randomly assigned). Booklets were collected when participants had finished.

In the full body (FB) condition, participants were instructed to simply view the photograph and identify the emotion the target was expressing by circling one of the seven available responses below the photograph (see Appendix L. for booklet examples). In the silhouette (FBSil) condition there was an additional instruction. Participants were asked to first look at the silhouette in the photograph and imagine that the silhouette represented them, that they were in the photograph with the target. Having done this, they were then to identify the emotion expressed by the target by circling one of the available responses. Participants were instructed to not over-think their responses, but to use their first instincts (see chapter three, section 3.3.2.3).

NB: It was later discovered that in the FBSil Kitchen background condition, one photograph was erroneously presented twice, whilst another photograph (low intensity fear – H8) was left out. Only the first appearance of this photograph was scored, and the second ignored. All responses to the photograph ‘H8’ in the fear category, in all photographic and background conditions, were ignored in the following analysis. All scores were converted to percentages to reflect the missing variable in the ‘fear’ emotion category.

4.5.2 Results.

The aim of this study was to see if different social settings, inserted into the backgrounds of the photographs, might impact AR-EA abilities. Due to a lack of literature in this specific area the following analysis was exploratory in nature. Significance levels, therefore, were adjusted to $p <= .01$; however, any results that were below .05 but above .01 have also been reported in order to avoid Type II errors (De Vellis, 2012).
4.5.2.1. Social context.

In both the full body (FB) and the silhouette (FBSil) conditions, the neutral background elicited higher accuracy percentages ($m = 70.95$, $SD = 8.36$ fb; $m = 69.28$, $SD = 9.72$ sil) than either the kitchen ($m = 67.03$, $SD = 10.98$ fb; $m = 66.17$, $SD = 7.25$ sil) or bar ($m = 69.19$, $SD = 9.106$ fb; $m = 69.21$, $SD = 8.62$ sil)(see figure 4.1 below). This indicates that participants were most accurate in identifying emotions when no background was present. A one-way ANOVA, however, revealed that this difference was not significant for either the full body (FB; $F(2,93) = 1.40, p = .252, \eta^2 = 0.029$) or the silhouette (FBSil) conditions ($F(2,100) = 1.29, p = .279, \eta^2 = 0.025$). Participants’ overall emotion accuracy scores did not seem to be influenced by the presence of the different backgrounds.

Looking at the individual emotion categories for the full body (FB) condition, it was revealed that participants did differ significantly in their accuracy scores for the ‘anger’ ($F(2,101) = 7.39, p = .001, \eta^2 = 0.128$). Post hoc analyses showed that the significant difference lay between the scores for the neutral background ($m = 69.74$, $SD = 17.27$) and the bar background with participants viewing the bar demonstrating significantly less accuracy ($m = 51.56$, $SD = 20.02$). There were no significant
difference in accuracy scores between the kitchen ($m = 58.82, SD = 22.56$) compared to either neutral or bar. No other significant differences were evident for any other emotion category. Therefore pictures with the bar background elicited less accuracy for anger compared to the other backgrounds.

In the silhouette (FBSil) condition differences were found for both surprise and sad emotion categories. These differences were not significant to the $p = .01$ level, however, they were below the $p = .05$ level and therefore should be reported. In the surprise category, accuracy percentages were highest for the bar background ($m = 70.83, SD = 18.85$), and lowest for the kitchen ($m = 57.22, SD = 18.92$), with the neutral background percentages falling between the two ($m = 60.32, SD = 19.81$). The difference was partially significant ($F(2,105) = 4.74, p = 0.11, \eta^2 = 0.083$), with post hoc analyses showing that this difference was below the $p < 0.05$ level between the neutral and bar backgrounds, and the kitchen and bar backgrounds, but not different between the neutral and kitchen backgrounds. Surprise, therefore, was best identified by participants when the bar background was present.

Sadness percentage scores were also partially significant between the different backgrounds ($F(2,103) = 3.97, p = 0.22, \eta^2 = 0.072$), where the lowest percentages were present with the bar background ($m = 53.81, SD = 18.12$), followed by the kitchen ($m = 62.07, SD = 14.70$) and then neutral backgrounds ($m = 62.08, SD = 19.41$). Only the difference between the neutral and bar backgrounds were significant ($p = .018$) according to post hoc analyses. Where sadness is concerned, the bar background inhibited accurate identification.

4.5.2.2. **Photographic condition by social context.**

Comparison of the photographic conditions for each background revealed no significant differences in accuracy percentages between FB and FBSil photographs. This means that regardless of whether the photographs contained only full body images, or had silhouettes inserted into them, participants’ performances were the same for the emotion identification task.
Analysis of individual emotion categories, however, did reveal one difference. For the emotion ‘anger’ in the neutral background, accuracy percentages were significantly higher in the full body condition ($m = 69.74, SD=17.27$) compared to the silhouette condition ($m = 55.16, SD=19.65$) ($t(78) = 3.51, p = .001, d=0.79$). Therefore participants had more trouble accurately identifying ‘anger’ when the silhouettes were present compared to only the full body images. According to Cohen’s $d$ this effect was quite large. This effect was not replicated with either the kitchen or bar backgrounds.

4.5.2.3. **Participant sex and social context.**

4.5.2.3.1 Full body (FB) photographic condition.

For each of the backgrounds female participants outperformed male participants in overall accuracy percentages for interpreting emotions (see figure 4.2 below). This overall difference was only partially significant for the bar background ($t (28) = 2.43, p = .022; d = 1.18, females m = 71.06, SD=8.67; males m = 61.71, SD=7.13$). Cohen’s $d$ shows that this difference is greater than one standard deviation indicating two distinct groups. Neutral backgrounds elicited the most accurate responses, followed by the bar and then kitchen backgrounds.

![Figure 4.2. Comparison of male and female participants' overall emotion category accuracy percentages by background](image)

Figure 4.2. Comparison of male and female participants' overall emotion category accuracy percentages by background
Analysis revealed that female participants were significantly more accurate at interpreting sad emotions \((m = 67.24, SD=15.74)\) compared to males \((m = 47.92, SD=20.77)\) in the neutral background condition \((t (35) = 2.87, p= 0.007, d = 1.05)\), and marginally better at interpreting the emotions of the male models (female participants \(m = 66.367, SD=10.80\)) compared to male participants \((m = 56.25, SD=8.10; t (34) = 2.52, p = .016, d = 1.09)\). In both cases the difference, according to effect sizes, seems to be quite large. There was no significant difference in performance between male and female participants in interpreting female model emotional expressions, indicating that when no backgrounds were present females were better at interpreting both sad emotions and those emotion expressed by the male models compared to the male participants.

There was little difference in accuracy scores for the kitchen backgrounds. Differences in the interpretation of Surprise did come close to traditional significance \((t (30) = 2.11, p = .043, d=.91)\) with females \((m = 66.67, SD=14.14)\) achieving higher accuracy percentages than male participants \((m = 52.78, SD=16.39)\). Again this difference is large, even though significance is marginal. Overall, the kitchen background offered no consistent advantage or disadvantage for either sex.

Bar backgrounds revealed several differences beyond the overall emotion category accuracy percentages reported above. Female participants \((m = 80.67,SD=14.97)\) were superior in accurately discerning disgust expressions \((t (29) = 2.53, p = .017, d = 1.21)\) compared to male participants \((m = 63.89, \ SD=12.55);\) low intensity expressions (females \(m = 68.36, SD=10.86;\) males \(m = 53.03, SD=3.71)\)(\(t (29) 3.37, p = .002, d = 1.90)\); and also better at identifying expressions from female models (female participants \(m = 78.10,SD=12.31\)) better than male participants \((m = 65.79, SD=7.25)\)(\(t (29) = 2.34, p = .027, d=1.22\)). The bar backgrounds revealed the most differences between the sexes, with all effect sizes greater than one standard deviation, indicating a strong influence on AR-EA abilities.
4.5.2.3.2 Silhouette (FBSil) photographic condition.

The performance between male and female participants seemed to be more stable in the silhouette photographic conditions. For both the neutral and kitchen backgrounds, no significant differences were found between accuracy percentages for either male or female participants, overall or for any specific emotion category. Bar backgrounds elicited one significant difference for the emotion ‘anger’ \( (t(35) = 2.78, p = .009, d=0.99) \) with females achieving higher accuracy percentages \( (m = 61.59, SD=23.80) \) compared to male participants \( (m = 41.67, SD=15.68) \), a difference of almost one full standard deviation. When the silhouette was present along with a bar background, males had more trouble accurately identifying anger than did females.

4.5.3 Discussion.

4.5.3.1 Social context.

Total emotion accuracy percentages indicated that the neutral background elicited the most accurate responses compared to both the kitchen and bar backgrounds. Although this difference was not significant, there were some significant differences in individual emotion categories. In the FB condition, anger responses were less accurate with the bar background present compared to the neutral background. The large effect size here indicating that the bar background had a strong impact on participant’s interpretations of anger. In the FBSil condition both surprise and sad emotions showed variations in accuracy percentages between backgrounds, however the effect size was quite small indicating that there may be no real difference. For surprise the scores were highest with the bar background, but for sadness, they were lowest for the bar background. The one common denominator in these results is that these differences seem to focus around the bar background. In some cases the bar background seemed to facilitate the interpretation of the emotion (surprise) and for others it inhibited it (anger and sadness). This result reflects the findings of Gesn and Ickes (1999) who stated that social context could act to enhance or undermine emotion identification accuracy.
The fact that participants seemed to demonstrate more difficulty for the negative emotions (anger, sadness) in the bar background compared to a more positive emotions such as surprise, is also interesting and could go some way to explain incidences of violence in these venues. The reason for this inaccuracy, however, is unclear. Participants were not questioned as to their experiences, negative or positive, in any venues of this type, and so the individual experiences and knowledge that each person brought to each particular stimulus is unknown. Research that links participant memory and experience with different social venues (backgrounds) and their performance on emotion identification tasks is the next step to unravelling this particular phenomenon.

As per the previous studies presented in this thesis, female participants continued to display better accuracy for interpreting emotional expressions than their male counterparts, although not all differences were significant. In the FB condition females showed superiority for several different emotions including sadness in the neutral backgrounds, surprise in the kitchen background, and disgust in the bar background. The differences between males and females, according to the effect sizes, were more than one standard deviation, which is considerable. Some previous research indicates that these differences may be indicative of sex differences. A general advantage has been found for women in accurately identifying the facial emotions of others (e.g.: Montagne, et al., 2005). Evolutionary theorists Hampson, et al. (2006) found that women processed negative emotional cues faster than men, and that women were significantly more accurate than men in identifying disgust and anger. Hall and Matsumoto (2004; study 1) found that female participants outperformed male participants in accurately decoding disgust, happiness, sadness, and surprise.

The current finding that female participants demonstrated better accuracy for more subtly portrayed emotions does have precedence (Hoffman, et al., 2010). There are few studies, however, that have dealt with variations in intensities of expressions and the potential influence this may have on AR-EA abilities. As already reported (see Chapter 1 section 1.6.2, and chapter 2, section 2.2.3.3), women do seem to have an overall advantage for the interpretation of emotions (e.g.: Hall, 1978;
Montagne et al., 2005), although other research has found that no sex differences exist when motivations other than sympathy are in play (Klein & Hodges, 2001). Knyazev, et al., (2010) suggest that variations between men and women may be due to different superiorities at the different emotion recognition processing stages (early processing stage for men, and late processing stage for women). Although the research to date may be somewhat unclear, the current study does suggest that women are more accomplished at reading more subtle expressions of emotion as suggested by Hoffman et al. (2010). It may be that only through the investigation of more nuanced levels of emotional expression will consistent sex differences be revealed.

4.5.3.2. Participant sex / model sex interaction

One of the most interesting results from this study comes from the variation in men’s abilities to interpret the emotions of male or female models. With no backgrounds male participants were worse than female participants at reading the emotions of the male models, however, with the bar as a backgrounds, male participants had most trouble interpreting female models. With the difference being more than one standard deviation between male and female participant performance, this seems to be a strong effect. There are several possibilities that might explain this difference.

Research has demonstrated that the sex of the target may impact emotion identification accuracy in observers. According to Hampson, et al. (2006) emotional interpretations may be enhanced by motives of reproduction (attractiveness). They found that women expressing negative emotions were rated as less attractive compared to when they expressed positive emotions. Tracy and Beall (2011) reported that, overall, women’s faces were rated as more attractive than men’s faces. Research has also found that women tend to be rated as being more emotionally expressive than men (Hampson, et al., 2006; Kring & Gordon, 1998). Given these findings it would be expected that males should demonstrate enhanced recognition of female model’s faces compared to male models, because a) female faces are considered more attractive, and b) female faces are more expressive. This, however, was not the case in the bar background.
If we expand on the hypotheses of Hampson et al. (2006) one explanation may be that with the bar as a social setting stimulus, males are more cognizant of male emotional expressions because they represent competition for female attentions. Rahman and Anchassi (2012) argue that men’s ability to interpret other male facial emotions may be an adaptive mechanism to enhance monitoring of threat from other males. If reproduction motives are an influence, then it would make logical sense that males might pay more attention to other male competitors in a social environment that has been traditionally linked to the activity of finding a mate. The influence of social situations on emotion recognition, however, has little literature behind it, and therefore requires further investigation in order to address this particular finding more definitively.

4.5.3.3. Participant sex / silhouette interaction

In the previous chapter (chapter 3 sections 3.3.3.2 and 3.3.4.3) it was reported that male participants showed significantly lower accuracy scores overall for identifying emotions when the photographic silhouette was present in the pictures. This inhibiting effect was thought to be produced by the extra cognitive load when participants imagined themselves firstly as the silhouette and then ‘in’ the photograph with the target. Although this overall effect was not seen in this present study a difference still became evident in the FBSil condition with the bar backgrounds. Male participants were significantly less accurate in interpreting angry emotions compared to female participants, and this difference was strong according to the large effect size gained. This is particularly interesting as this finding indicates another specific social environmentally driven deficit for men, which may help to explain the aggressive behaviours often seen at pubs, clubs, bars and taverns. Again there is little research in this specific area, however, we do know that social context, motivations and social goals do have the potential to influence our interpretations of others (Wyer & Srull, 1986; De Melo et al., 2013). Hoffman (1984; 1987) proposed that advanced empathy mechanisms would include the memories and previous knowledge of the observer. When witnessing an event affecting another, these higher order cognitions would be triggered, drawing on the experiences of the observer. The result of this could be to inhibit or facilitate the emotion
interpretation abilities of the observer, depending upon these individual experiences (Hoffman, 1984; 1987; Gesn & Ickes, 1999; Sze, et al., 2012; Wyer & Srull, 1986). The experiences we have in different social settings can influence the way we interpret both events and others within that specific setting, a finding which has found support in the current study.

4.5.3.4. Limitations

One limitation of this study is the lack of questioning of participants in terms of their experiences, thoughts, feelings or memories for the different types of social venues portrayed in the photographic backgrounds. This information may have shed more light on the mechanisms at work during AR-EA processes. The mistake in one of the booklets that necessitated the elimination of one of the fear photographs may also have impacted results, especially since this brought down the number of fear photographs being analysed from 6 to only 5. None of the conditions were found to have any influence on interpretations of fear, although this could have been because of the reduced number of stimuli for that particular emotion category.

The fact that participants were in a social environment when filling out the booklets may have impacted results as well, although it is unclear in what way. Some of the booklets containing the neutral backgrounds were filled in at the beginning of a lecture, with students seated facing toward the front of the lecture theatre, and prepared for class. This meant that less social interaction was taking place in this testing environment compared to those who were tested in the Bush Court setting. This could explain the higher accuracy scores acquired, although specific testing to compare social and non-social testing environments would need to be done to confirm this hypothesis (also see chapter 3 section 3.3.4.3 for a discussion on the possible effects of social interactions with cognitive load and AR-EA abilities).

4.6 Study 2: Reliability – Test-Retest

The results regarding the reliability of the AR-EA photographic measure were unclear in chapter three. A Cronbach’s alpha level of 0.686 was reported. Whilst below the ideal 0.7 required
for a measure to be considered reliable (De Vellis, 2012), it was proposed that given that the measure contained items varying in difficulty, this alpha level may have been considered acceptable (De Vellis, 2012). The aim in this current study, then, was to confirm the reliability of the AR-EA photographic stimuli by conducting a second reliability test in the form of a test-retest methodology. The AR-EA photographic measure (FBSil mode) was compared to the full body (FB) photographs to confirm if the silhouette mode remained more reliable. Both sets of photographs (with neutral backgrounds) were tested twice over a period of 2-4 weeks and the results analysed. It was predicted that, as in chapter three, the full body with silhouette photographic (FBSil) condition would attract higher reliability than the full body only (FB) photographs.

4.6.1 Method.

4.6.1.1. Participants.

The test-rest study involved 65 university students, 45 females and 20 males. The mean age was 27.55 years, with an age range of 18 to 70 years old. Students were recruited through lectures where students volunteered to fill out the questionnaires before the lecture. Students were then approached again in the same lecture 2-4 weeks later (depending upon availability). This ensured standardisation of the testing environment across both sessions.


Two different sets of booklets were used for each testing session (T1 and T2). The booklets contained 38 black and white photographs of male and female models expressing six different emotions, each at three different levels of intensity (see Appendix M. for an examples of images in the booklets). There were two presentation conditions, full body (FB) and full body with silhouette (FBSil). (See Chapter 3 section 3.3.2.2.1 and appendix H. for more details on the silhouettes, their development and use). In the full body condition participants were instructed to look at the photograph and determine the emotion being expressed by the target by circling one of seven responses presented below the photograph “nothing; surprised; disgusted; fearful; angry; sad;
happy”. In the silhouette condition participants were asked to imagine that they were the shadowed silhouette and imagine that they were in the photograph with the target. They were then asked to identify the emotion expressed by the target by circling one of the seven responses. Participants were instructed not to over think their responses and to give only one response per photograph (see chapter 3, section 3.3.2.3).

Booklets were randomised between presentation conditions (FB and FBSil) and also between T1 and T2 so that participants were not presented with the same photographs in the same order, reducing order and practice effects. As well as photographs, questions regarding participant demographics were also included such as age and sex.

4.6.1.3. Procedure.

For both testing sessions the testing environment was held constant. Participants were approached to volunteer their time filling out a booklet at the beginning of a lecture. They were informed that they would be asked to fill in the booklets again in 2 – 4 week’s time, in the same lecture theatre. Participants were offered the chance to win one of three $50.00 Coles/Myers vouchers for their participation. At the end of 10 minutes of testing, participants handed back their booklets. Their booklet number was recorded against their name to ensure that they received the same testing condition at T2. These records were destroyed at the end of the second round of testing to maintain anonymity. The same procedure was carried out 2 – 4 weeks later depending upon the availability of the different lectures.

4.6.2 Results – Test-retest.

The average time between testing was 18.68 days. The shortest period was 12 days (only one participant) and the longest was 28 days (18 participants). Pearson’s correlations revealed no significant relationship between number of testing days and total emotion category score percentages, therefore amount of time did not influence participant performance.
The average emotion category score percentage was 69.93%, which is above chance. Mean raw score (percentages) at T1 was $m = 68.75\%$ (SD = 9.18) and at T2 was $m = 68.64\%$ (SD = 9.08). In the full body (FB) photographic condition, the mean accuracy percentages for the total emotion score at T1 was $m = 70.09\%$ (SD = 8.56) and at T2 was $m = 69.64\%$ (SD = 8.50). In this photographic condition (see figure 4.3 below), happy elicited the highest accuracy percentage followed by neutral, disgust, anger, sad, fear and surprise. There was little variation in the mean percentage scores for happy indicating the presence of a ceiling effect. A series of Repeated Measures t-tests revealed no significant differences between T1 and T2 for any of the emotion categories.

The mean accuracy score percentage for the total emotion scores for the Silhouette condition at T1 was $m = 67.57\%$ (SD = 9.67) and at T2 was $m = 67.78\%$ (SD = 9.60). Again, there seemed to be a ceiling effect operating for ‘happy’, which attracted the highest accuracy percentage. The next highest, as in the full body condition, was neutral followed by disgust, sad, surprise, fear and anger. A series of Repeated Measures t-tests revealed no significant differences between T1 and T2 percentage scores for any of the emotion categories (see figure 4.4 below).
In order to establish the levels of reliability between the two testing times a series of bivariate correlations were done, for the total emotion score and for each emotion category. Each photographic condition was analysed separately in order to determine which one was more reliable.

The full body (FB) photographic condition total emotion score percentages were significantly, strongly and positively correlated between T1 and T2 ($r(25) = .81, p < .001$). Looking at the individual emotion categories (see table 4.1 below) only 4 of the total 7 emotion categories returned significant correlations between the two testing times. This suggests that some emotions, fear, anger and happiness, in the full body condition were more reliable than others.

Table 4.1

<table>
<thead>
<tr>
<th></th>
<th>Neutral</th>
<th>Surprise</th>
<th>Disgust</th>
<th>Fear</th>
<th>Anger</th>
<th>Sadness</th>
<th>Happiness</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r$</td>
<td>0.69</td>
<td>0.43</td>
<td>0.48</td>
<td>0.23</td>
<td>0.23</td>
<td>0.46</td>
<td>0.27</td>
</tr>
<tr>
<td>$df$</td>
<td>29</td>
<td>29</td>
<td>26</td>
<td>27</td>
<td>27</td>
<td>26</td>
<td>29</td>
</tr>
<tr>
<td>sig</td>
<td>0.000**</td>
<td>0.015*</td>
<td>0.010*</td>
<td>0.223</td>
<td>0.219</td>
<td>0.013*</td>
<td>0.135</td>
</tr>
</tbody>
</table>

$NB$: ** => $p < .001$; * => $p < .05$

Figure 4.4. Mean accuracy score percentages for each emotion category comparing T1 and T2 for the silhouette (FBSil) photographic condition.
The silhouette (FBSil) photographic condition total emotion score percentages were significantly, moderately and positively correlated between the two testing times ($r(31) = -0.66, p < .001$). Table 4.2 below shows the correlations between T1 and T2 for each emotion category. Only one emotion did not reach significance, surprise, with a probability level just above the .05 threshold. All other emotions were significantly and positively correlated between the two testing times.

Table 4.2

<table>
<thead>
<tr>
<th></th>
<th>Neutral</th>
<th>Surprise</th>
<th>Disgust</th>
<th>Fear</th>
<th>Anger</th>
<th>Sadness</th>
<th>Happiness</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r$</td>
<td>0.40</td>
<td>0.31</td>
<td>0.58</td>
<td>0.61</td>
<td>0.59</td>
<td>0.74</td>
<td>0.50</td>
</tr>
<tr>
<td>$df$</td>
<td>32</td>
<td>31</td>
<td>32</td>
<td>32</td>
<td>31</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>sig</td>
<td>0.018*</td>
<td>0.075</td>
<td>0.000**</td>
<td>0.000**</td>
<td>0.000**</td>
<td>0.000**</td>
<td>0.003*</td>
</tr>
</tbody>
</table>

NB: **=> $p < .001$; * => $p < .05$

The correlation for the overall emotion scores between T1 and T2 was higher for the FB condition compared to the FBSil condition. It was in the FBSil condition where 6 of the 7 emotion categories were positively and significantly correlated between T1 and T2, whereas this was only the case in 4 of the 7 emotions in the FB condition. The FBSil condition, then, demonstrated more reliability across all emotion categories compared to the FB condition.

4.6.3 Discussion.

The reliability for the test-retest demonstrated, once again, that the silhouette (FBSil) condition was a more reliable measure of affect recognition-empathic accuracy (AR-EA) than the full body (FB) condition. Correlations showed that the length of time between testing did not influence accuracy scores, however this may have been due to the shortness of the time period between 2-4 weeks only. Further testing, over a longer time period, would be necessary to verify the current findings.
Correlations between the two testing times revealed that the total emotion score percentages were significant for both the full body (FB) and the silhouette (FBSil) photographic conditions, although the correlation was stronger in the FB condition. When the individual emotion categories were investigated, however, the silhouette (FBSil) photographic condition demonstrated far more stability and reliability compared to the FB condition. In the FB photographs three emotions did not attain significant correlations (fear, anger and happiness), whereas the FBSil had only one emotion which did not reach significance (surprise). This result indicates that the FBSil photographic condition was the most reliable measure of AR-EA abilities across all emotion categories.

This is in line with the findings in chapter three where the FBSil condition returned a higher Cronbach’s alpha level compared to the FB condition. Although the correlation in this present study is not considered high, it should still be considered adequate given that the measure includes items of varying difficulty. The different difficulty levels meant that the most difficult items would have attracted the lowest correlations and the least difficult the highest, negatively impacting the overall correlation results. In conclusion, the FBSil photographic stimuli have been demonstrated to be more reliable across all seven emotion categories compared to the FB photographs. The results of the psychometric evaluations conducted in the previous chapter and the current chapter confirm that the photographs that include the silhouette are a better measure of AR-EA abilities, and therefore should be considered as the AR-EA photographic measure.

4.7 Conclusion

Study 1: Social Context investigated the impact of social setting on participant AR-EA abilities found that the different social contexts can influence empathic processes. Although the findings were subtle, being confined to specific emotions, model sex, and levels of expression intensity, the results do indicate that social context inserted into the photographs via backgrounds did influence accuracy for emotion identification. It is particularly interesting that most of these results focused on the bar background, a social environment known for violence and anti-social behaviours. The
inhibiting effect for the interpretation of anger, disgust, and sadness, could go some way to explaining some of the more antisocial behaviours often found in venues of this type. If, as the results suggest, the social environment of a bar has an inhibitory effect on the interpretation of emotions, especially for males, then this misinterpretation could have a negative impact on social interactions, resulting in unwelcome behavioural outcomes. The explanation for this effect was hinged on previous findings that social and cognitive schemas may impact emotion interpretations (see chapter 3 section 3.3.4.3). These cognitive schemas are built up from expectations, experience and memory, which then act as a bias or filter for information regarding the emotions of others (Gesn & Ickes, 1999; Wyer & Srull, 1986). More specific research is needed to confirm this effect and the processes behind it, however, this current study, and its findings, have opened the door to a new area of inquiry that promises much in explaining some anti-social behaviours in others, as well as shedding new light on AR-EA abilities.

A final evaluation of reliability also confirmed the findings from chapter three that those photographs including the silhouette were superior to other photographic modes for assessing AR-EA abilities in participants. This particular set of photographs, therefore, has been dubbed the ‘AR-EA photographic measure’. The measure offers new ways of investigating empathic accuracy abilities, promising new insights into sex differences as well as the ability to manipulate elements such as social context to investigate potential factors in empathic accuracy processes.
Chapter 5 – General Discussion and Conclusions

5.1 Overview

Empathic accuracy, in the context of this thesis, was defined as: A response to another’s experiences through the ongoing interpretation processes of the other’s thoughts, feelings, motives, situation, as well as verbal, visual, social, and historical cues, that may result in an emotional reaction in the observer, and which may subsequently influence behaviour. This definition is important as it highlights the evolving nature of empathic processes. Empathy is not a static, one-off event. It is an ongoing, interpretative and information gathering process that can facilitate and enhance our social interactions with others. By interpreting other’s emotions, we come to understand their situation from their point of view, adjusting our own behaviours accordingly. For this reason it is important that we understand the role that empathy has to play in our social lives, as well as the factors that may influence it.

One of the aims of this thesis was to investigate the possible influence of inserting different empathy-related elements into photographs on affect recognition-empathic accuracy (AR-EA) abilities. In order to achieve this, an original set of photographic stimuli were developed for testing with the final goal being the development of a photographic measure to assess AR-EA abilities. This represented the second major goal of this thesis.

The opening chapter provided an overview of the empathy construct distinguishing between dispositional empathy and empathic accuracy. It was determined in this review that, in order to fully understand how empathy operates, study of both dispositional empathy and empathic accuracy were needed. Various empathic accuracy methodologies were reviewed. It was concluded that photographs depicting models displaying various emotions represented the most cost effective and adaptive means for assessing AR-EA abilities. Photographs can be easily manipulated to include specific elements designed to increase imagination and the amount of connection between the observer and the target. This could be achieved through the insertion of a blacked-out silhouette figure via which participants could imagine themselves into the photograph with the target. Social
context could also be easily manipulated in photographs through the inclusion of different social settings as backgrounds. Thus the first major aim of the thesis was established.

Chapter two focused on the initial development of a set of photographs to be used for testing and the development of an AR-EA photographic measure, the second major aim of the thesis. The photographs were full body black and white images of male and female models expressing six basic emotions (happy, sad, fearful, disgusted, surprised, and angry; as per Ekman (1992a; 1992b; 1999a), plus one neutral expression (see chapter 2 section 2.3, and appendix F). With the exception of neutral, each emotion was expressed at three different levels of intensity: low (slightly happy); medium (very happy); and high (extremely happy). Each emotion category and each intensity level was represented by each sex, giving a total of 38 photographs. The photographs were initially tested across two studies (see chapter 2) to establish base-line response patterns, which were used as a basis of comparison for the studies to follow.

In the first section of chapter three, the influence different empathy-related constructs (imagination and connectivity) might have on AR-EA abilities were investigated. Three different photographic presentation modes were used: full body only (FB); head and shoulders only (H&S); and full body with silhouette (FBSil). The silhouette consisted of a blacked-out gender-oriented figure (male and female), via which participants could imagine themselves into the photograph with the target. It was hoped that this mechanism would enhance imagination processes and also promote connectivity between the participant and the photographic target: two essential elements in empathic accuracy processes. Results indicated that in both the full body presentation modes (FB and FBSil), participants obtained higher AR-EA scores compared to the head and shoulders only (H&S) presentation. It was also found that the insertion of the silhouette into the photographs impacted AR-EA abilities, especially for male participants, whose ability to accurate interpret other’s emotions was inhibited.

The second section of chapter three focused on the psychometric evaluation of the three different sets of photographic stimuli: FB, H&S, and FBSil. Through validity and reliability testing it
was discovered that the photographs containing the silhouettes had superior psychometric qualities compared to the other two presentation modes. It was proposed, then, that the silhouette photographs should form the AR-EA photographic measure.

Chapter four continued with the investigation into photographic elements that might influence AR-EA abilities in participants, namely, social context. Using different social settings, (a kitchen, a bar, and neutral) as backgrounds in the full body photographs, the potential impact on AR-EA scores was investigated. It was found that the ‘bar’ background had the most profound effect on empathic accuracy, facilitating the accurate interpretation of some emotions, whilst inhibiting others.

The second study in chapter four extended the investigation into the reliability of the AR-EA photographic measure. The full body silhouette (FBSil) photographs were compared with the full body only (FB) photographs in a test-retest methodology. It was found that scores correlated significantly across the two testing times for both modes of photographic presentation, however, those photographs that contained the silhouettes were more consistent across the seven different emotion categories. This result confirmed that the photographs with the silhouettes included should be used as the AR-EA photographic measure for assessing empathic accuracy abilities.

The findings for each study have already been discussed (see chapters 2, 3 and 4), therefore what follows is an attempt to bring these findings together in order to look at the influences on empathic-accuracy more generally. It will be concluded that the elements introduced into the photographs were successful at influencing AR-EA abilities, which raises new questions and possibilities for future research. The superior accuracy scores obtained with the FB versions of the photographs (both FB and FBSil), over that of the more traditional mode of H&S, whilst not overthrowing previous research findings, does suggest that more research is necessary to find the best method, using photographs, for assessing AR-EA abilities. More than this, was the conclusion that the full body photographs that contained the silhouettes were the best for assessing AR-EA abilities and so was named the ‘AR-EA photographic measure’. Finally the implications of these
findings will be discussed, introducing, again, new directions for research and possible explanations for the interaction of social context, behaviour, and empathic accuracy.

5.2 Enhancing Imagination and Connectivity via Photographic Stimuli

One of the aims in chapter three was to test the viability of inserting an imagination facilitation element into the photographs and to investigate if this would have any influence on participant AR-EA scores. The silhouette (one male and one female) was designed to ‘place’ the participant in the photograph with the target, thereby enhancing not only imagination processes in the participant but to also increase the level of connection between participant and target. As established in both chapter one (section 1.7) and chapter three (section 3.2) both imagination processes and connection between the observer and target are important elements in empathic accuracy processing.

The results indicated that participant AR-EA scores did increase with the presence of the silhouette, compared to head and shoulders stimuli only, although there was no significant differences between the full body presentation modes (FB and FBSil). What was revealed, however, was a difference in performance between male and female participants.

5.2.1 The impact of imagination on sex differences in AR-EA abilities.

The inclusion of the silhouette revealed a difference in performance between males and females. In chapter three it was found that whilst there was no difference between male and female total AR-EA scores, there was a significant difference in overall scores in the FBSil condition (see sections 3.3.3.3 and 3.3.4.3). The inclusion of the silhouette coincided with a slight increase in female scores, and a decrease in male scores, resulting in a significant difference between the sexes. Whilst there is no previous research that has investigated this particular technique, it was suggested that the difference may have been the result of increased cognitive load for the task. The inclusion of the silhouette and the instruction to imagine oneself into the photograph with the target may
have increased cognitive load in participants, which had a negative impact on AR-EA abilities in males (see chapter 3 section 3.3.4.3 for a fuller discussion of this).

This result, however, was not replicated in the neutral background condition in chapter four, with no overall differences found between males and females regardless of photographic condition (FB or FBSil) (see section 4.5.3.3). It was suggested that this may have been due to the testing environments used in chapter four. In chapter three, students were approached in a social setting during lunch hour (see section 3.3.2.3). In chapter four, some students were tested in a lecture situation where, arguably, there would be less free socialising (see sections 4.5.1.3 and 4.6.1.3). It was theorised, therefore, that the combination of a social testing environment used in chapter three, with the extra imagination steps required in the silhouette condition resulted in additional cognitive load that negatively impacted male AR-EA scores (see chapter 3 section 3.3.4.3 and chapter 4 section 4.5.3.3). This theory will need to be directly tested providing new directions for research. If it is found that social interaction plus additional imagination steps does negatively impact AR-EA abilities in males, this may offer some interesting insights and explanations regarding male social behaviours.

It was also posited that the instructions for the silhouette condition, rather than the silhouettes themselves, may have been responsible for the higher reliability and accuracy scores (see section 3.4.3.1). The only way to test this is to specifically study the impact of instructions upon participant accuracy ratings. This could be done with the full body and full body with silhouette photographs separately in order to distinguish which, the instructions or the silhouettes, are having this impact on AR-EA abilities.

5.2.2 Expression intensity and sex differences in AR-EA abilities.

The research reviewed in both chapters 1 (section 1.6.2) and 3 (section 3.3.4.3) revealed inconsistent findings in relation to gender differences in empathic accuracy tasks. Most of the research indicated that females had an advantage in this ability, outperforming their male counterparts (Hall & Matsumoto, 2000), however other research found no differences at all.
Hoffman, et al. (2010) varied the intensity of the different emotional expressions, from subtle to more intense expressions. They found that the female participants were better than male participants but only for more subtle expressions. The findings from both chapters 3 and 4 indicate that, females are not only better than males at accurately identifying more subtle emotional expressions, but also more intense expressions. Men seemed to identify medium level expressions best. Like Hoffman et al.’s (2010) study, it was only with the inclusion of varying levels of emotional expression that differences between the sexes were revealed. This suggests that there may be more to empathic accuracy processes than originally thought, and that these processes may differ between males and females depending upon the level of emotional intensity displayed. There are various theories that attempt to explain sex differences in empathy abilities such as: evolutionary explanations (Hampson, et al., 2006); display rules (Kring & Gordon, 1998); differing recognition processes (Knyazev, et al, 2010), and even different gender-related motivations (Graham & Ickes, 1997). The current findings suggest, however, that men and women may vary in their ability to interpret emotions depending upon the intensity with which they are expressed.

5.3 Social Context

Chapter four focused on the potential influence social context might have on AR-EA abilities. Whilst there is some research regarding the impact of social context on empathy and empathic accuracy tasks, the research is problematic, with poor research designs and the confounding of constructs (such as emotional priming)(see chapter 1, section 1.7.2.1). Given that photographs offer a simple way of introducing social context into the AR-EA task, it is surprising that there is no other research that has attempted to do or test this method and its influence.

Three different backgrounds were compared: neutral (no background); a kitchen; and a bar (as in a tavern or pub). Both the full body (FB) and full body with silhouette (FBSil) modes were used resulting in nine different conditions. Photographs were randomised for each condition, with instructions for each presentation mode remaining the same as in chapter three.
It was found that the different photographic backgrounds did influence AR-EA abilities in participants. Rather than influencing overall scores, however, differences were revealed at the more nuanced levels of emotion category and model sex. The neutral and kitchen backgrounds revealed little differences, except in the FB mode where surprise was more accurately interpreted by females compared to males for the kitchen background. The bar background produced the most differences.

5.3.1 Social context and photographic mode interactions.

In the FB condition, the emotion surprised attained significantly higher accuracy scores with the bar background compared to both the neutral and kitchen backgrounds. Sadness, however, attracted the lowest scores with the bar background. Females scored significantly higher than male participants with the bar background and were also better at interpreting the emotion disgust. Compared to male participants, female participants were also better at interpreting the emotions of female models. At the lowest level of expression intensity, females were also superior. In the FBSil condition, the emotion surprise was most accurately identified in the presence of the bar background, whilst accuracy scores for sadness were lowest with this background. Compared to female participants, males were poor at accurately identifying the emotion anger, at all levels of intensity, and with both male and female models.

These results suggest that empathic accuracy processes can be influenced by the social context in which an emotion is being interpreted. It is particularly interesting that most of the statistically significant results centred upon the bar background. It was suggested in chapter four that the schemas participants may draw upon in the ‘bar’ situation may influence their AR-EA abilities. Gesn and Ickes (1999) made the argument that the knowledge and memory that come with specific social contexts could enhance or inhibit emotion identification processes (see chapter 1 section 1.7.2 for a fuller discussion of this). As it stands, the current findings are isolated and more research will be needed to ascertain how these cognitive schemas and social contexts influence AR-EA abilities, however, these findings do provide some interesting avenues for further research.
5.4 Comparing and Testing Photographic Presentation Modes

Traditionally photographic stimuli for empathic accuracy studies rely on the head and shoulders only of the target. This was due to the understanding that the face offered the most information when it came to emotion interpretation (Ekman, 1965; 2004; Ekman & Friesen, 1967; Ekman & Friesen, 1974a; Ekman et al., 1980). Comparisons of the head/face of a target and the full body of the target were rare, with most of the research focusing on the differences between the face and body (Ekman & Friesen, 1967), rather than looking at head and body together. A review of the literature, however, revealed studies that found that when presented together, participants would use both head and body to interpret the emotions of others, and in fact could not separate the two (Aviezer, et al., 2011; Aviezer, et al., 2012; see chapter 1 section 1.6.1 for a fuller discussion of this issue).

A series of tests were conducted to ascertain which type of photographic presentation would best measure AR-EA abilities. In chapter three, the study compared three different presentation modes: full body only (FB); head and shoulders only (H&S); and full body with silhouette (FBSil; which will be discussed in the following section 5.3 below). Reliability and validity tests were done on each presentation type (see section 3.4).

Results revealed that the head and shoulders presentation obtained the lowest accuracy scores, as well as the lowest reliability scores. This mode was therefore excluded from further testing. Both FB and FBSil modes obtained similar accuracy scores. Cronbach’s alpha, however, was highest for the FBSil condition, reaching .686. Due to the varying levels of difficulty amongst the testing items, this level was considered acceptable (see chapter 3 section 3.4.3).

Validity was tested using various scales measuring constructs that were expected to be related in specific ways to AR-EA abilities (see section 3.4). As expected, AR-EA scores and dispositional empathy scores, as measured via the Shortened IRI, were not related, confirming discriminant validity. There was a significant negative relationship between AR-EA and a measure of narcissism, also establishing discriminant validity. It was expected that AR-EA scores would be
positively related to a measure of social intelligence, given the close relationships between empathy and theory of mind, and emotional intelligence constructs (see chapter 1 section 1.2.3.1 and chapter 3 section 3.4). Convergent validity, however, was not established. This was thought to be the result of comparing a social intelligence disposition with an actual task measuring ability (see chapter 3, section 3.4.3).

Chapter four compared only the FB and FBSil conditions in a test-retest methodology in order to try and confirm the reliability results obtained in chapter three. Both the FB and FBSil conditions attracted significant positive correlations between time1 and time2 testing, although the FB correlations were stronger. Looking at the individual emotion categories, however, it was revealed that the FBSil condition was much more consistent across all emotion categories, with all but one emotion (surprise) failing to reach significance. In the FB condition three emotions (fear, anger and happiness) did not reach significance.

Looking at the results for both reliability and validity across the studies in chapters three and four, it was concluded that the FBSil condition offered a more reliable, stable and valid measure of AR-EA abilities, and therefore was considered to be the best measure of AR-EA. For this reason the photographs that contained the silhouettes formed the new AR-EA photographic measure.

5.5 Overall Findings.

Taken together the findings from these studies point to some interesting new possibilities for future research, and question some of the techniques used in the research to date. The versatility of photographs as a medium through which various factors can be manipulated has been born out. The findings from this thesis demonstrate that photographic elements can be utilised to improve the nature of empathic accuracy assessment and research. A mechanism which does not appear to have been used in previous empathy research.
5.5.1 Photographic presentation mode.

One of the major findings of these studies was the superiority of both the full body photograph types when compared to head and shoulders only presentations. There have been few studies that have actually compared these different presentation modes, and the finding that participants were far more accurate at interpreting the emotions of others when both the face and body were present questions the validity of those studies still relying on head and shoulders only stimuli. Whilst Ekman and colleagues (Ekman & Friesen, 1967; Ekman & Friesen, 1969) may be correct in stating that the face is information rich when it comes to emotional expressions, when participants were given extra information, in the form of the target’s body, they used it. This can be seen in the open ended responses given by participants in the very first study in chapter 2 (sections 2.4.1.2 and 2.4.1.3) when asked how they went about interpreting the target’s emotions. Most said that if they were unsure about the face, then they looked at the body. This is particularly important given that the intensity of target expressions varied. When the facial expression was subtle, participants could and did use the target’s body to help in their interpretations.

5.5.2 Imagination, connectivity, and social context.

The photographs were easily manipulated to enhance empathy processes, as well as change the social context within which emotional expressions were presented. The simple inclusion of a silhouette as a mechanism via which participants could imagine themselves into the photographs primed imagination processes increased the connection between participant and target. Beyond the more traditional directions given to participants to either ‘imagine self’ or ‘imagine other’, imagination processes have neither been manipulated nor facilitated in this way by other studies. The fact that more subtle differences between male and female participants was revealed through the use of this technique tells us that we need to do more to uncover and understand the differences in empathic accuracy between the sexes. If, as posited, cognitive load can have a negative impact on male empathic accuracy processes, then this presents opportunities to explain antisocial behaviour in socially taxing situations, such as a pub, or restaurant.
The results concerning the impact of social context on AR-EA abilities gives this explanation some merit. The simple insertion of different backgrounds into the photographs influenced the interpretation of specific emotions, the interpretation of subtle emotions, and the interpretation of emotions expressed by different sexes. Overall male participants seemed most impacted, with most of the results coming from the bar background. Whilst the interpretation of some emotions seemed to be facilitated by the background, others were inhibited, meaning that social schemas may be operating, triggered by the specific social context. These schemas, in turn, influence emotion identification. Not only this, but it was proposed in chapter 3 that there may also be evolutionary processes at work (see section 3.3.4.3). The fact that male participants, in the bar context, had difficulty interpreting the emotion ‘anger’ and also had difficulty with female model expressions, may be a result of primal sexual drives based in evolutionary theory. Whilst these suggestions are conjecture until tested, the current findings give new direction and purpose to research in these areas.

5.5.3 The AR-EA photographic measure.

The second major aim of this thesis was to develop a set of photographic-based stimuli to measure AR-EA abilities. Original photographs were taken for this task, and doctored to produce three different photographic presentation modes: full body only (FB); head and shoulders only (H&S); and full body with silhouette (FBSil). Head and shoulders presentations have been used in the majority of the research, however, they ignore other emotion cues that might be used in the process of interpreting another’s emotions (see chapter 1 section 1.6.1). By comparing the three different photographic modes it was established that the full body photographs with the silhouette attracted higher accuracy ratings of the target’s emotions compared to H&S only photographs, as well as revealing superior validity and reliability scores compared to either H&S or FB presentation modes.

Psychometric evaluation of the silhouette photographs occurred over chapters three and four. Each time this set of stimuli outperformed the other types of photographs (FB and H&S). Due
to the results it was suggested that the inclusion of the silhouette aided imagination and connectivity processes in participants, which, in turn, facilitated empathy processes. The result, therefore, was a more reliable, valid and stable measure of AR-EA abilities that was dubbed the AR-EA photographic measure. Future testing of the AR-EA photographic measure will need to be done in order to establish norms and scoring procedures.

5.6 Limitations

As already stated within each chapter there were limitations in each of the studies undertaken. In chapter 2 very small sample sizes made some statistical comparisons unadvisable, although the findings here did give a baseline for the types of responses that could be expected from this particular set of photographic stimuli. These response patterns were closely matched throughout the other studies attesting to the reliability of the photographs.

A problem with the initial testing design in chapter two, open-ended responses by participants, resulted in the testing having to be conducted again using a slightly different testing methodology. The low accuracy scores obtained in this first test was thought to be the result of participants being able to give open-ended responses when identifying the emotions expressed emotions by targets. This lead to more than 270 unique words and/or phrases to describe just the six basic emotions and one neutral expression. It is fascinating that the human vocabulary for emotion is so large, highlighting the importance of emotions in our everyday lives. The other unintended consequence of this research design was the failure of participants to identify the ‘neutral’ expression. It was suggested that this arose from a lack of information. Participants were asked to identify each emotion, but were not told that ‘no emotion’ was a possible response, nor were participants instructed that each emotion was to be portrayed more than once at varying levels of intensity. This may have set up a demand characteristic for participants, leading them to come up with a new emotion label for each new facial expression, hence the large number of different responses. In previous literature accuracy ratings have always improved when the emotion categories have been made available for participants to respond against (see chapter 2 section
2.4.2), therefore a change in method using a forced-choice design resulted. The new methodology resulted in improved accuracy ratings for every emotion category.

In chapter three, although imagination and connectivity were the empathic factors manipulated, participants were not asked if they felt closer to the photographic target, nor asked if they found it difficult or easy to imagine themselves into the photographs as instructed. The exact nature of the changes in participant processing that occurred as a result of the inclusion of the silhouettes was, therefore, unclear. Further testing is necessary to ascertain the exact effect the silhouettes in the photographs may be having on participants.

A similar problem occurred in the studies of social context in chapter four. Participants were not asked what thoughts or memories they associated with each background, and therefore results could not be compared to the schemas they potentially brought to the testing. Again a study to specifically investigate this would be necessary in order to understand how the interaction between social context and cognitive schemas influence AR-EA abilities.

5.7 Conclusions

The aim of this thesis was to investigate through the use of photographic stimuli, what factors may influence AR-EA abilities. It was found, as per the theoretical models proposed (see chapter 1, sections 1.3, 1.4 and 1.7), that imagination, connectivity, and social context can all enhance or inhibit empathic accuracy processes. Imagination and connectivity enhanced empathic processes for females, whilst seeming to inhibit the same processes for male participants, under specific testing environments. Social context, in particular settings of a bar (as in a tavern or pub), facilitated the interpretation of surprise, but inhibited interpretations of anger and sadness. For male participants the bar setting also inhibited their ability to accurately interpret the emotions of the female models.

All of these findings point to the nuanced workings of empathic accuracy processes that are only revealed when factors such as emotional expression intensity are varied, and other empathy-related factors (imagination, connectivity, and social context) are manipulated. Elements which can
be easily manipulated using a photographic medium, revealing a new way of using photographic stimuli in empathy research. The mechanisms used in the studies above, a blacked-out silhouette, and backgrounds depicting different social contexts, are simple and easy to use. The ability to enhance or facilitate specific aspects of the empathy construct, such as imagination and connectivity, can give important insights into empathic accuracy processes, such as the differences between males and females when an extra imagination step is introduced, or when more subtle emotional expressions need to be interpreted. The influence of social context was also an important finding in these studies, as being able to understand the deficits in empathic accuracy that may be operating in different social settings will bring new insights to human behaviour. Through these insights we can begin to look at other influences that may then act to change or modify behaviour.

The other aim of this thesis was to develop a photographic-based measure for assessing AR-EA abilities. The results from the testing conducted throughout this thesis suggested that full body photographs, and not head and shoulders only stimuli, were not only a more accurate, but also more reliable measure of AR-EA abilities. In particular, those full body photographs that contained the silhouettes revealed superior psychometric properties compared to the other photographic modes. Whilst current findings do not necessarily refute previous literature, it does call into question past studies that have used head and shoulders stimuli only, and highlights the fact that researchers should not necessarily rely upon previously ‘tested’ methods, but rather endeavour to find new and improved ways to assess, measure, and capture, different psychological concepts and processes. The AR-EA photographic measure will require further testing and development, but it does point the way to how testing stimuli, such as photographs, can be ‘reinvented’ to further the research into empathy processes.

The AR-EA photographic measure represents a new way of assessing AR-EA abilities. The inclusion of the silhouette in the photographs enhances empathy process (imagination and connectivity) whilst the variation in expression intensity allows for a more nuanced investigation of AR-EA abilities. The measure contains 38 black and white full body photographs of male and female
models portraying six basic emotions (surprise, disgust, fear, angry, sad, and happy) and three different levels of expression intensity (low, medium, and high), plus two neutral expressions (one male, one female). The measure has several advantages, the first being the low language load, with participants simply choosing between one of seven emotions labels. It is a paper and pencil test, meaning that, when formulated, administration and scoring procedures will be standardised and simple to carry out. These factors make for an adaptable and cost effective method of assessing AR-EA abilities in a wide range of participants.
References


174


177


Doi: 10.1037/0022-3514.32.6.1068
APPENDIX A.

Procedures used with Photographic Models during the Photographic Session

Arrival:
- Upon arrival, photographic models were introduced to the experimenter (Tracey Woolrych) and the photographer (Liv Stockley).
- The model was then asked to read and fill in the consent/media release forms
  - One form was for the model to keep, and one to be kept for the study
- The model was also given a detailed information sheet to keep

Briefing – Photographic session procedure

- The photographic model was then informed of the procedure to be followed during the photographic session.
- The following was read out to each model:

  "Today we will be taking full body black and white photographs of you while you express different emotions. These emotions are anger, fear, disgust, surprise, sadness and happiness. We will ask for three different levels of intensity for each emotion – low, medium, and high. For instance we might begin by asking you to imagine how it would feel to be only a little angry with someone or something. Then ask you to be very angry, and finally, extremely angry.

  There will only be myself and Liv in the room with you when you act out these emotions. We want you to be as natural as possible – so you can just straight out act the emotion, or we have some eliciting exercises that can help you as well.

  For each emotion, we want you to get as close to natural as possible. This will mean that some of these emotions may feel uncomfortable. It is ok if this happens, but if it gets too much, then you can stop the session at any time.

  To help you express these emotions, we have some props for you, and some scenarios or situations that you can bring to mind or imagine. Not all of these might work for you – and that is fine. If one scenario isn’t working for you, we will choose another.

  After each emotion, we will ask you how you feel and how well you think you captured the emotion – we will be recording these responses. This is not to test you, but to check to make sure you are ok, and to get your own opinion on how well you think you are expressing these emotions – do not be overly hard on yourself – but respond honestly.

  Each emotion will be dealt with separately, and we will have short rest periods between each one, during which we may also ask you do some activities, such as puzzles, or some exercises. This is to make sure you feel better, and to make sure no one emotion is lingering for you. The activities and exercises are designed to help get rid of any excess emotion before we move onto the next section so we don’t get ‘mixed’ emotions.

  REMEMBER: If you become uncomfortable, you can stop at any time. Or, if you find that this is not working for you at all, you can end the session completely. We want to make sure you are comfortable in doing this, so if you aren’t, then please let us know.
Before we begin – do you have any questions?
To begin we will do some warm-up exercises to get you going.”

Relaxation and stretching exercises:
- The photographic model was taken through a series of short physical exercises designed to warm up the body and face, ready for the expressing of emotions.
- See Appendix B for more details of these exercises.

THE PHOTOGRAPHIC SESSION:
For each emotion set, the same basic procedure was followed.
- Firstly it was explained to the model what emotion was to be captured and at what level of intensity
  o Where different intensity levels were sought, the lowest intensity level was asked for first, followed by the medium intensity and finally the most intense version of the emotion.
  o Models were asked to draw on their own experiences where possible in expressing each emotion
  o Where models had trouble with this, suggestions in the form of scenarios, or props were used to help with the elicitation of the required emotion.
- Secondly the model was asked how they felt about the expressing of that particular emotion.
  o These responses were recorded (see Appendix D. for the participant emotion scale used for this)
  o Models were then asked how they felt generally
  o Cleansing exercises were then used if necessary to extinguish any remaining emotion (to prevent blending of emotions; see Appendix C. for some examples)
  o The model was then asked if they were ready to ‘move onto the next emotion’. If they replied “No”, then more cleansing exercises were completed until the model felt they were ready to continue.

The following is a brief outline of the procedure and elicitation techniques used for each emotion set.

NEUTRAL:
“First we will just get a picture of you with no emotions – so keeping your face NEUTRAL”
  o The photographic model was asked to adopt a relaxed stance –
     ▪ A photograph of this will be taken – as a baseline
     ▪ The model was then asked “How do you feel?” –
       • (anticipate responses such as silly, nervous, to calm.)
  ASK participant how they felt about that –
  o Did they feel as if they captured a NEUTRAL face?
  o How do they feel now?
     ▪ MOVE ONTO CLEANSING EXERCISES IF NECESSARY
  o “Ready to move on?” (yes / no)
  o If No – then do more cleansing exercises if necessary – then pose the question again.
  o “Ready to move on?” (yes / no)
SURPRISE:
“The next emotion will be SURPRISE”
- First start with **slightly surprised**:
  - *You just ran into a person accidently turning a corner*
- Now try a more intense version – **very surprised**:
  - Give the mode a chance to try on their own first – then,... OR
  - *Form an ‘O’ with your mouth and open your eyes wide*
- Now for the most intense level – **extremely surprised**:
  - *You just won the lottery!*
  - *You just opened an envelope with winning tickets for a trip for 2 all expenses paid holiday to Tahiti!*

**ASK participant how they felt about that** –
- Did they feel as if they captured SURPRISE? - RECORD
- How do they feel now? - RECORD
- MOVE ONTO CLEANSING EXERCISES IF NECESSARY
- How do you feel now?
- Ready to move on?

CONTEMPT / DISGUST:
“The next emotion will be CONTEMPT or DISGUST – we have a few props to help you with this one”.

- First start with **slightly disgusted/contempt**:
  - --- HANDLING the slime or icky spider
  - *Going into the toilet straight after someone else has been*
  - *Having to wipe the toilet seat after someone else left a mess*
  - *Catching a whiff of the rubbish bin*
- Now try a more intense version – **very disgusted/contempt**:
  - *Having to dispose of some rotting vegetables*
  - *Cleaning out a wound full of pus*
  - *Picking up a dead cricket/cockroach*
- Now for the most intense level – **extremely disgusted/contempt**:
  - *Encountering the putrid smell of a rotting body*
  - *Having to speak to a convicted serial child rapist*
  - *Sifting through rotting rubbish with your bare hands*

**ASK participant how they felt about that** –
- Did they feel as if they captured DISGUST/CONTEMPT? - RECORD
- How do they feel now? – RECORD
- MOVE ONTO CLEANSING EXERCISES IF NECESSARY
- How do you feel now?
- Ready to move on?

FEAR:
“The next emotion will be FEAR”

- First start with **slightly fearful**:
  - *Your being in an unfamiliar dark room*
  - *Watching a horror movie*
  - *Running into a spider’s web at night*
- Now try a more intense version – **very fearful**:
  - *Waiting to go into an exam*
Standing up to give a speech at an important family event (and everyone is looking at you)
Standing on the ledge/plane doorway to take a bungee-jump/skydive for the first time
Being in the house alone at night and hearing an unfamiliar sound – like an intruder?

Now for the most intense level – extremely fearful:
- Walking out on stage to deliver a speech to 20,000 people
- Handling a live spider/snake
- Being confronted with some holding a gun to you

ASK participant how they felt about that –
- Did they feel as if they captured FEAR? - RECORD
- How do they feel now? – RECORD
  - Do you want to take a break?
  - MOVE ONTO CLEANSING EXERCISES IF NECESSARY
- How do you feel now?
- Ready to move on?

ANGER:
“The next emotion will be ANGER”
- First start with slightly angry:
  - Someone just cut you off driving on the freeway
  - Someone just bumped into (deliberately) at the pub, spilling your drink
  - Someone just pushed in front of you in a cue
- Now try a more intense version – very angry:
  - You just found out a close friend lied to you
  - The company just stuffed up your pay (for the 3rd fortnight in a row) which means you won’t get paid on time and a repayment will be missed
  - It’s a hot day, and you have been waiting patiently for the parking spot for the last several minutes, only to have someone pull into it and steal your spot (despite the fact that you had your indicator on!)
  - You have told you child (PET?) four times to behave, without success – in fact they seem to be continuing despite you!
- Now for the most intense level – extremely angry:
  - After a day of continual frustrations, where nothing seems to go right, an idiot scrapes along the side of your car with a shopping trolley
  - A stranger provokes to the point where you think you will explode
  - You are having a fight with your partner – and it has escalated to the point of yelling and pot throwing!

ASK participant how they felt about that –
- Did they feel as if they captured ANGER? - RECORD
- How do they feel now? – RECORD
  - Do you want to take a break?
  - MOVE ONTO CLEANSING EXERCISES IF NECESSARY
- How do you feel now?
- Ready to move on?

“OK – YOU ARE DOING GREAT – ALMOST THERE – “
SADNESS:
“The next emotion will be SADNESS”
  o  First start with **slightly sad:**
    ▪ Watching a sad or romantic movie with a tragic ending
    ▪ Empathising with an ad that shows starving children in Africa
    ▪ Not getting that part in the play you auditioned for
  o  Now try a more intense version – **very sad:**
    ▪ Having your boyfriend/girlfriend of only 8 weeks break up with you
    ▪ Getting a really bad mark on an assignment
    ▪ Not getting that much needed job
    ▪ hearing of a devastating tragedy – such as the earthquakes in New Zealand
      or Japan, the tsunamis in Indonesia, or the tragedy of 09/11 - imagine back
      to how you felt about all those people
  o  Now for the most intense level – **extremely sad:**
    ▪ Failing an important exam
    ▪ Break up of a long term relationship
    ▪ The death of a loved one

**ASK participant how they felt about that –**
  o  Did they feel as if they captured SADNESS? - RECORD
  o  How do they feel now? – RECORD
  o  Do you want to take a break?
  o  MOVE ONTO CLEANSING EXERCISES IF NECESSARY
  o  How do you feel now?
  o  Ready to move on?

“NOW SOMETHING TO MAKE YOU FEEL BETTER…..”

HAPPINESS:
“The next emotion will be HAPPINESS”
  o  First start with **slightly happy:**
    ▪ Walking along you look down and find $2!
    ▪ You just saw a little bird outside your office window
    ▪ A stranger just smiled at you
  o  Now try a more intense version – **very happy:**
    ▪ You just had a wonderfully productive day – and the boss even
      complimented your work
    ▪ You just spent some quality time with your best friend – catching up about
      old times
    ▪ You just got a promotion / you got that role
    ▪ You just got a great mark on an assignment
  o  Now for the most intense level – **extremely happy:**
    ▪ You just had the news that you got that dream job
    ▪ You landed a part on Broadway
    ▪ You just won first division....
    ▪ It’s your birthday!
    ▪ You’ve just been proposed to
    ▪ Your partner just said “I love you” for the first time
    ▪ You just got asked out on a date by your dream guy/woman

**ASK participant how they felt about that –**
  o  Did they feel as if they captured SADNESS? - RECORD
  o  How do they feel now? – RECORD
- Do you want to take a break?
- MOVE ONTO CLEANSING EXERCISES IF NECESSARY
  o How do you feel now?
  o Ready to move on?

De-Brief:
- The models were then de-briefed in the following manner.
- “Well done – you have done very well. We just have to check the photos to make sure they all came out before you go... in the meantime...
  o :How do you feel now? Any questions? What did you think about the session? Any suggestions you might make? “
- After all photographs were checked, and the model confirmed that they felt fine, they were released.

  “AGAIN – THANK YOU SO MUCH FOR YOUR PARTICIPATION – I GREATLY APPRECIATE IT!”

________________________________________
APPENDIX B.

Relaxation and Stretching Exercises

**BODY SHAKE OUT:**
- Get participants to stand with feet slightly apart
  - Imagine that someone is holding a string that runs from the centre top of your head right down through the centre of your body
- Close your eyes and take a slow deep breath in, hold it for a bit, then slowly let it out again
  - Repeat this 3 times
- Open your eyes slowly
- Roll your shoulders gently – forwards, then backwards
- Roll your head gently & slowly – look up, look down, look left, look right. Put your left ear to your left shoulder, now your right ear to your right shoulder.
- Stretch your arms up over your head, now let them just fall back down to your side
- Shake your hands out, and now shake out your shoulders
- Shake your feet out – one at a time

**FACIAL WARM UP:**
- Open up your mouth wide and gently move your jaw back and forth – now close
- Stick out your tongue and wiggle it around a bit
- Scrunch up your nose and then relax
- Purse your lips, and suck in your cheeks, now relax
- Now – gently place your fingers on your face and massage your cheeks, forehead, temples and jaw.
APPENDIX C.

Cleansing Exercises (examples only)

The tasks were designed to distract the model from their emotions by engaging them cognitively through the use of puzzles and trivia.

Card sorting:
- Into numerical/value order – eg all the aces, all the twos etc.
- According to suit
- Numerical/value order according to suit
- As fast as you can!

Word puzzles:
See if you can solve these puzzles?

Example:  

<table>
<thead>
<tr>
<th>HEAD</th>
<th>SHOULDERS</th>
<th>ARMS</th>
<th>BODY</th>
<th>LEGS</th>
<th>ANKLES</th>
<th>FEET</th>
<th>TOES</th>
</tr>
</thead>
</table>
What is the next two letters in this sequence? O, T, T, F, F, ?, ?

Problem solving puzzles:
Taken from a Logic Problem puzzle book.

Word Find Book:
See if you can solve this word find? – see book

Simple Sudoku Puzzle Book:
See if you can solve one of these Sudoku puzzles? – see book

Trivia and general knowledge questions:
- General Knowledge questions:
  o “What is the capital of Queensland?”
  o “What are the three primary colours?”

Counting and math problems
- Count down slowly by 3s starting from 200 and working back to 0.
- What is 15 + 8?

Gentle relaxation:
- Breath in slowly through your nose for a count of 3
- Then breath out through your mouth for a count of 4
- Now we are going to extend this – concentrate on your breathing
- Breath in slowly through your nose for a count of 5
- Now hold your breath for a count of 3
- Breath out slowly through your mouth for a count of 5.
### APPENDIX D.

**Model Emotion Expression Capture Scale**

**HOW WELL DO YOU FEEL YOU CAPTURED THE EMOTION?**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Didn’t capture the emotion at all</td>
<td>Felt I captured the emotion, but only slightly</td>
<td>Felt I captured the emotion, okay – but not brilliantly</td>
<td>Felt I captured the emotion quite well</td>
<td>Completely captured the emotion</td>
</tr>
</tbody>
</table>
APPENDIX E.

Sample of Booklet used in Study 1: Open-ended Responding

(NB: photographs in these appendices are not representative of the quality of the end product used for testing)

Please look at each photograph in turn, and answer the questions below each photo. Make sure you answer each question before moving on to the next photograph. Don’t take too long considering your responses – there is no right or wrong answer – just give your first impression.

What do you think this person is feeling? ___________________________________

____________________________________________________________________

What do you think this person is thinking? __________________________________

____________________________________________________________________

What do you think this person is feeling? _________________________________

____________________________________________________________________

What do you think this person is thinking? _________________________________

____________________________________________________________________
What do you think this person is feeling? __________________________________
____________________________________________________________________

What do you think this person is thinking? _________________________________
____________________________________________________________________

Please note:
The first photograph in this booklet example represented Surprise, at low intensity.
The second photograph represented Fear, at high intensity.
APPENDIX F.

Procedure used for selection of photographs to be used for further testing.

Percentage accuracies were calculated for all photographs from the second round of testing. The first selection of photographs was based on an accuracy threshold of 70% or better (see Table 1. below). In the case where there was more than one choice between models for a specific level of an emotion, the highest scoring photograph was chosen. If the photographs obtained the same accuracy score, the photo with the clearest view of the face would be chosen.

Table 1 Appendix G.
Percentage accuracies by photograph, emotion, intensity and model.

| EMOTION | INTENSITY | MALE MODELS | | | FEMALE MODELS | | | | |
|---------|-----------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
|         |           | Model A | Model C | Model F | Model H | Model B | Model D | Model E | Model G |
| Neutral |            | 87.50 | 87.50 | 62.50 | 37.50 | 100.00 | 87.50 | 62.50 | 87.50 |
| Surprise| slightly   | 62.50 | 25.00 | 25.00 | 12.50 | 25.00 | 0.00 | 37.50 | 100.00 |
|         | very       | 100.00 | 100.00 | 75.00 | 25.00 | 87.50 | 87.50 | 75.00 | 87.50 |
|         | extremely  | 75.00 | 87.50 | 75.00 | 75.00 | 87.50 | 75.00 | 62.50 | 100.00 |
| Disgust | slightly   | 75.00 | 62.50 | 25.00 | 87.50 | 12.50 | 50.00 | 62.50 | 100.00 |
|         | very       | 75.00 | 0.00 | 0.00 | 50.00 | 100.00 | 100.00 | 87.50 | 0.00 |
|         | extremely  | 100.00 | 12.50 | 50.00 | 25.00 | 100.00 | 100.00 | 100.00 | 12.50 |
| Fear    | slightly   | 12.50 | 12.50 | 0.00 | 50.00 | 37.50 | 75.00 | 100.00 | 0.00 |
|         | very       | 37.50 | 25.00 | 50.00 | 87.50 | 62.50 | 75.00 | 75.00 | 25.00 |
|         | extremely  | 62.50 | 25.00 | 62.50 | 50.00 | 100.00 | 100.00 | 100.00 | 75.00 |
| Anger   | slightly   | 75.00 | 0.00 | 25.00 | 75.00 | 50.00 | 50.00 | 62.50 | 87.50 |
|         | very       | 87.50 | 0.00 | 12.50 | 87.50 | 100.00 | 87.50 | 87.50 | 50.00 |
|         | extremely  | 100.00 | 50.00 | 50.00 | 75.00 | 100.00 | 62.50 | 100.00 | 62.50 |
| Sadness | slightly   | 0.00 | 25.00 | 12.50 | 25.00 | 87.50 | 37.50 | 87.50 | 0.00 |
|         | very       | 100.00 | 62.50 | 50.00 | 25.00 | 87.50 | 75.00 | 87.50 | 50.00 |
|         | extremely  | 87.50 | 87.50 | 62.50 | 25.00 | 87.50 | 100.00 | 75.00 | 100.00 |
| Happiness | slightly | 100.00 | 100.00 | 50.00 | 50.00 | 87.50 | 100.00 | 100.00 | 87.50 |
|         | very       | 75.00 | 87.50 | 0.00 | 87.50 | 87.50 | 87.50 | 100.00 | 87.50 |
|         | extremely  | 62.50 | 87.50 | 100.00 | 75.00 | 37.50 | 75.00 | 75.00 | 25.00 |
The 70% or better threshold resulted in 31 photos of male models and 50 photos of female models. For the female models, a full range of photographs could be selected using this criteria, that is, that each emotion and intensity level were represented. The selection for male models was more restricted. For the male models Neutral, Disgust, Anger, and Happy were the only emotions to achieve 70% or better accuracy across all levels of intensity. Other emotion categories yielded only one or two photographs reaching the 70% or better threshold. By lowering the threshold to 60% for male models a full range of photographs from across all emotion categories and intensity levels could be selected, with two exceptions: the lowest level of intensity for sad and fear.

To ensure that all emotion categories and intensity levels were represented by both sex, the best performing male models for low intensity sad and fear categories were chosen. For the lowest level of intensity for fear model H was selected with an accuracy of 50%. The lowest intensity level for the sad category is not as robust with both Model C and Model F being the highest with accuracy percentages of only 25%, therefore either photograph could be selected.

To make this final decision, participant feedback regarding the photographs was consulted. Participants indicated that both of these male models seemed to be the ‘same’ in many of the photographs, making identification of emotions difficult. More specifically though model C was reported as looking embarrassed in many of the photographs, a category not available for selection. This blending of more than one emotion represented a confound and therefore the entire series for Model C was removed from further testing, resulting in Model F being selected.

Another issue arose for the female models in terms of their clothing. Model D was dressed in a loose black dress, which some participants said made reading of the body language difficult. It was believed that this could also result in a confound, and so the entire series of photographs for Model D were also removed. This left 3 male and 3 female models.
representing the six different emotions, across the 3 different intensity levels, plus neutral.

This resulted in one complete set of 19 male stimulus photographs and one set of 19 female stimulus photographs. Based on this, the initial instrument was made up of a minimum of 38 photographs. It was also decided to include one male and one female photo that had attracted 0% scores. These were to act as deliberately ambiguous material to compliment the low intensity levels of the emotion categories. Table 2 below shows the final selection of photographs for further testing.

Table 2 Appendix F.
Photograph selections for one male and one female model per emotion category and intensity level.

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Intensity</th>
<th>Male Model</th>
<th>Female Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
<td>----</td>
<td>A1</td>
<td>B1</td>
</tr>
<tr>
<td>Surprise</td>
<td>Slightly</td>
<td>A2</td>
<td>G2</td>
</tr>
<tr>
<td></td>
<td>Very</td>
<td>F3</td>
<td>B3</td>
</tr>
<tr>
<td></td>
<td>Extremely</td>
<td>H4</td>
<td>E4</td>
</tr>
<tr>
<td>Disgust</td>
<td>Slightly</td>
<td>H5</td>
<td>G5</td>
</tr>
<tr>
<td></td>
<td>Very</td>
<td>A6</td>
<td>B6</td>
</tr>
<tr>
<td></td>
<td>Extremely</td>
<td>A7</td>
<td>E7</td>
</tr>
<tr>
<td>Fear</td>
<td>Slightly</td>
<td>H8</td>
<td>E8</td>
</tr>
<tr>
<td></td>
<td>Very</td>
<td>H9</td>
<td>E9</td>
</tr>
<tr>
<td></td>
<td>Extremely</td>
<td>A10</td>
<td>B10</td>
</tr>
<tr>
<td>Anger</td>
<td>Slightly</td>
<td>H11</td>
<td>G11</td>
</tr>
<tr>
<td></td>
<td>Very</td>
<td>H12</td>
<td>E12</td>
</tr>
<tr>
<td></td>
<td>Extremely</td>
<td>F13</td>
<td>B13</td>
</tr>
<tr>
<td>Sad</td>
<td>Slightly</td>
<td>F14</td>
<td>E14</td>
</tr>
<tr>
<td></td>
<td>Very</td>
<td>A15</td>
<td>B15</td>
</tr>
<tr>
<td></td>
<td>Extremely</td>
<td>A16</td>
<td>B16</td>
</tr>
<tr>
<td>Happy</td>
<td>Slightly</td>
<td>A17</td>
<td>E17</td>
</tr>
<tr>
<td></td>
<td>Very</td>
<td>H18</td>
<td>G18</td>
</tr>
<tr>
<td></td>
<td>Extremely</td>
<td>F19</td>
<td>E19</td>
</tr>
<tr>
<td>Ambiguous</td>
<td>----</td>
<td>F6</td>
<td>G8</td>
</tr>
</tbody>
</table>

The selected full body photographs were then checked for the adequacy of the head and shoulders view. If the facial expression of the corresponding head and shoulders photograph was obscured in any way, or deemed unclear, the photograph (both full body and head and shoulders) was swapped out for another model, however this proved to be unnecessary. One photographs A15 showed the model’s hand in the head and shoulders view.
This was removed using Photoshop. The final result was a set of 40 full body photographs, and 40 head and shoulders only photographs.
APPENDIX G.

Final photographs selected across six models by emotion category and intensity level.

**NEUTRAL:**

![Neutral](image1)

**SURPRISE:**

<table>
<thead>
<tr>
<th>INTENSITY</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image2" alt="Surprise Low" /></td>
<td><img src="image3" alt="Surprise Medium" /></td>
<td><img src="image4" alt="Surprise High" /></td>
<td></td>
</tr>
</tbody>
</table>
DISGUST:
INTENSITY:  Low  Medium  High

FEAR:
INTENSITY:  Low  Medium  High
ANGER:
INTENSITY: Low | Medium | High

SAD:
INTENSITY: Low | Medium | High
HAPPY:

**INTENSITY:**
- Low
- Medium
- High

AMBIGUOUS:

- Disgust: Medium intensity
- Fear: Low intensity
APPENDIX H.

Creation of the Silhouette.

The insertion of silhouettes into the photographs is designed to enhance both the imaginativity for the participant and the connectivity between the participant and the photograph target. Due to the distinctive nature of many aspects of the existing photographs, silhouettes were created using composites from several different photographs in order to create one generic male silhouette, and one generic female silhouette.

The female silhouette used Model B as the base. The clothing was changed by blending the skirt and legs from Model E. The buckles that stuck out on Model E’s shoes were deleted and the shape slightly changed, as was the shape of the skirt. As well as this, the hair line of Model B was changed from straight here to softly kinked hair, and the shoulder line was changed.

The male silhouette used Model H as the base. The legs were lengthened by blending Model F’s lower half. Model H’s shirt cuffs were deleted and the arm shape changed and made to look a little further from the body. The neckline was slightly narrowed and lengthened, and the hair line was also changed.

The next problem was the orientation of the silhouettes. The silhouettes should give the impression of facing away from the participant, facing forward towards the target in the photograph. However, the feet of the silhouettes were pointing the wrong way for this. Changing the direction of the feet proved to be onerous, with outcomes looking less than natural. This problem was overcome via the solving of another problem.

The silhouettes, when placed in the photographs, were of an equal height to the photographic targets. This reduced the optical perspective of the participant standing in front of the target within the photograph. Therefore silhouettes were enlarged. This produced the perspective that they were then standing in front of the photographic model, and also solved the problem of the feet as, in most cases, they were cut off at the bottom of the photograph. In cases where this did not happen, the target in the photograph was repositioned.

Once the blended silhouettes had been created a reversed (horizontal plane) image of each was also saved. This meant that 4 silhouettes had been fully created, 2 male and 2 female. These were then ‘tested’ with the full body neutral images to be used in the photographic instrument, Model A (male) and Model B (female). Several viewings of these photographs revealed that the reversed images of both silhouettes seemed to work the best and produce the most generic male and female forms, as well as looking ‘natural’ within each photograph.

The silhouettes were then manually inserted into each black and white, fully body photograph. The photographs were then viewed several times to make sure that the correct optical perspective was achieved for each photograph. See examples below:
APPENDIX I.

Example of testing booklets: Study 1 and Study 2

Expressions of the Human Face and Body
- Study 1 (Ethics Approval 2012/165)

PARTICIPANT CONSENT:
I have read the Information letter about the nature of this study and what the testing involves. Any questions I have about the research process have been answered to my satisfaction. I agree to take part in this research. I am aware that this study is anonymous and no personal details are being collected or used by the research. I know that I may change my mind and withdraw my consent to participate at any time; and I acknowledge that once I have completed the testing it may not be possible to withdraw my data.
I understand that all information provided is treated as confidential by the researchers and will not be released to a third party unless required to do so by law. I understand that the findings of this study may be published and that no information which can specifically identify me will be published.

Please circle one:  

Agree  Disagree

PARTICIPANT INFORMATION:  
AGE: _________  SEX: M / F

EDUCATION LEVEL:
What is the highest level of education you have completed? (please tick one)

☐ Still in High School
☐ Less than High School (Year
☐ Graduated from High
☐ Some University, or other Tertiary study (eg TAFE)
☐ Bachelor’s Degree
☐ Postgraduate Degree

SOCIO-ECONOMIC STATUS:
Approximately what is your total annual income (before tax)? (Optional)

☐ Less than $15,000
☐ $15,000 - $24,999
☐ $25,000 - $39,999
☐ $40,000 - 69,999
☐ $70,000 - $99,999
☐ $100,000 - $150,000
☐ $150,000 - $200,000
☐ Over $200,000

On the following pages you will see some black and white photographs. Please look at each photograph in turn and then tick the box you think is the closest match to the photo. Make sure you answer each question before moving on to the next photograph. Don’t take too long considering your responses – there is no right or wrong answer – just give your first impression.
Full Body Only (FB) presentation

What do you think this person is feeling? (please circle one)

Nothing        Surprised        Disgusted        Fearful        Angry        Sad        Happy

238
What do you think this person is feeling? (please circle one)

Nothing    Surprised    Disgusted    Fearful    Angry    Sad    Happy
Head and Shoulders Only (H&S) presentation

What do you think this person is feeling?
(please circle one)

Nothing  Surprised  Angry  Fearful
Disgusted  Sad  Happy

What do you think this person is feeling?
(please circle one)

Nothing  Surprised  Angry  Fearful
Disgusted  Sad  Happy

What do you think this person is feeling?
(please circle one)

Nothing  Surprised  Angry  Fearful
Disgusted  Sad  Happy

240
Imagine yourself *IN* this picture (you are the shadow). You are witnessing first-hand what the other person is feeling. Circle *one* of the words below that matches what that person is feeling.

Nothing  Surprised  Disgusted  Fearful  Angry  Sad  Happy
Imagine yourself IN this picture (you are the shadow). You are witnessing first-hand what the other person is feeling. Circle one of the words below that matches what that person is feeling.

Nothing    Surprised    Disgusted    Fearful    Angry    Sad    Happy
APPENDIX J.

Scales used for validity testing: Study 2

Shortened IRI

**Instructions:** For each of the statements below, indicate how well it describes YOU by putting a circle around a letter;

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Never</td>
<td>S</td>
<td>Sometimes</td>
<td>O</td>
</tr>
</tbody>
</table>

1. I feel concerned for people having a hard time.       N   S   O   A
2. I understand people better by imagining how things look to them.       N   S   O   A
3. I don’t listen to other people when I know I’m right.       N   S   O   A
4. I don’t feel much pity for people being treated badly.       N   S   O   A
5. I try to understand what the other person is feeling when they upset me.       N   S   O   A
6. I get upset by things I see happen.       N   S   O   A
7. I feel protective towards someone who needs help.       N   S   O   A
8. I find it hard to see the other guy’s point of view.       N   S   O   A
9. I don’t feel sorry for people with problems.       N   S   O   A
10. I look at everyone’s side of an argument.       N   S   O   A
11. I’m a pretty soft-hearted person.       N   S   O   A
12. There are two sides to everything and I look at both of them.       N   S   O   A
13. Other people’s problems don’t upset me much.       N   S   O   A
14. I imagine how I would feel if I were somebody, before I judge them.       N   S   O   A
Tromsø Social Intelligence Scale

Using the 1 - 6 scale below, please indicate your agreement or disagreement with each question by circling the appropriate number next to it.

<table>
<thead>
<tr>
<th>Q.</th>
<th>Question</th>
<th>Describes me extremely well</th>
<th>Describes me extremely poorly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q.1</td>
<td>I can predict other peoples' behaviour</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Q.2</td>
<td>I often feel that it is difficult to understand others' choices</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Q.3</td>
<td>I know how my actions will make others feel</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Q.4</td>
<td>I often feel uncertain around new people who I don't know</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Q.5</td>
<td>People often surprise me with the things they do</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Q.6</td>
<td>I understand other peoples' feelings</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Q.7</td>
<td>I fit in easily in social situations</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Q.8</td>
<td>Other people become angry with me without me being able to</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Q.9</td>
<td>I understand others' wishes</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Q.10</td>
<td>I am good at entering new situations and meeting people for the first time</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Q.11</td>
<td>It seems as though people are often angry or irritated with me when I say what I think</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Q.12</td>
<td>I have a hard time getting along with other people</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Q.13</td>
<td>I find people unpredictable</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Q.14</td>
<td>I can often understand what others are trying to accomplish without the need for them to say anything</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Q.15</td>
<td>It takes a long time for me to get to know others well</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Q.16</td>
<td>I have often hurt others without realizing it</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Q.17</td>
<td>I can predict how others will react to my behaviour</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Q.18</td>
<td>I am good at getting on good terms with new people</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Q.19</td>
<td>I can often understand what others really mean through their expression, body language, etc</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Q.20</td>
<td>I frequently have problems finding good conversation topics</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Q.21</td>
<td>I am often surprised by others' reactions to what I do</td>
<td>1</td>
<td>5</td>
</tr>
</tbody>
</table>
**Narcissistic Personality Inventory (NPI-16)**

Read the pairs of statements below – for each pair choose the one that describes you best by circling it.

<table>
<thead>
<tr>
<th>I hope I am going to be successful</th>
<th><strong>OR</strong></th>
<th>I am going to be a great person</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sometimes I tell good stories</td>
<td><strong>OR</strong></td>
<td>Everybody likes to hear my stories</td>
</tr>
<tr>
<td>I insist upon getting the respect that is due me</td>
<td><strong>OR</strong></td>
<td>I usually get the respect that I deserve</td>
</tr>
<tr>
<td>People always seem to recognize my authority</td>
<td><strong>OR</strong></td>
<td>Being an authority doesn’t mean that much to me</td>
</tr>
<tr>
<td>I don’t like it when I find myself manipulating people</td>
<td><strong>OR</strong></td>
<td>I find it easy to manipulate people</td>
</tr>
<tr>
<td>I can make anybody believe anything I want them to</td>
<td><strong>OR</strong></td>
<td>People sometimes believe what I tell them</td>
</tr>
<tr>
<td>I always know what I am doing</td>
<td><strong>OR</strong></td>
<td>Sometimes I am not sure of what I am doing</td>
</tr>
<tr>
<td>I expect a great deal from other people</td>
<td><strong>OR</strong></td>
<td>I like to do things for other people</td>
</tr>
<tr>
<td>I like to be the centre of attention</td>
<td><strong>OR</strong></td>
<td>I prefer to blend in with the crowd</td>
</tr>
<tr>
<td>I am no better or nor worse than most people</td>
<td><strong>OR</strong></td>
<td>I think I am a special person</td>
</tr>
<tr>
<td>I am an extraordinary person</td>
<td>OR</td>
<td>I am much like everybody else</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>----</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>It makes me uncomfortable to be the centre of attention</td>
<td>OR</td>
<td>I really like to be the centre of attention</td>
</tr>
<tr>
<td>I like having authority over people</td>
<td>OR</td>
<td>I don’t mind following orders</td>
</tr>
<tr>
<td>There is a lot that I can learn from other people</td>
<td>OR</td>
<td>I am more capable than other people</td>
</tr>
<tr>
<td>I try not to be a show off</td>
<td>OR</td>
<td>I am apt to show off if I get the chance</td>
</tr>
<tr>
<td>When people compliment me I sometimes get embarrassed</td>
<td>OR</td>
<td>I know that I am good because everybody keeps telling me so</td>
</tr>
</tbody>
</table>
Marlowe-Crowne Social Desirability Scale (MCSD)

Listed below are a number of statements concerning personal attitudes and traits. Please read each item and decide whether the statement is true or false as it applies to you. For each item, please circle **TRUE** or **FALSE**.

| Q.1. No matter who I'm talking to, I'm always a good listener | TRUE or FALSE |
| Q.2. There have been a few occasions when I took advantage of someone | TRUE or FALSE |
| Q.3. I sometimes try to get even, rather than forgive and forget | TRUE or FALSE |
| Q.4. When I don't know something, I don't at all mind admitting it | TRUE or FALSE |
| Q.5. There have been occasions when I felt like smashing things | TRUE or FALSE |
| Q.6. I never resent being asked to return a favour | TRUE or FALSE |
| Q.7. I have almost never felt the urge to tell someone off | TRUE or FALSE |
| Q.8. I am sometimes irritated by people who ask favours of me | TRUE or FALSE |
| Q.9. I sometimes think when people have misfortune they only got what they deserved | TRUE or FALSE |
| Q.10. I have never deliberately said something that hurt someone's feelings | TRUE or FALSE |
APPENDIX K.

Social context – backgrounds

Kitchen

Bar
Full Body (FB) Presentation (examples only)
Silhouette (FBSil) presentation (examples only)
APPENDIX L.

Examples of testing booklets: Study 1 (Social context)

Silhouette (FBSil) – Neutral background

Imagine yourself IN this picture (you are the shadow). You are witnessing first-hand what the other person is feeling.

What do you think this person is feeling? (please circle one)

Nothing    Surprised    Disgusted    Fearful    Angry    Sad    Happy

251
Imagine yourself **IN** this picture (you are the shadow). You are witnessing first-hand what the other person is feeling.

What do you think this person is feeling? (please circle one)

Nothing  Surprised  Disgusted  Fearful  Angry  Sad  Happy
Imagine yourself IN this picture (you are the shadow). You are witnessing first-hand what the other person is feeling.

What do you think this person is feeling? (please circle one)

Nothing    Surprised    Disgusted    Fearful    Angry    Sad    Happy
Full body (FB) – Neutral background

What do you think this person is feeling? (please circle one)

Nothing  Surprised  Disgusted  Fearful  Angry  Sad  Happy

254
What do you think this person is feeling? (please circle one)

Nothing  Surprised  Disgusted  Fearful  Angry  Sad  Happy
What do you think this person is feeling? (please circle one)

Nothing    Surprised    Disgusted    Fearful    Angry    Sad    Happy
APPENDIX M.

Example of testing booklets: Study 2 (Test-retest)

Full body (FB) presentation

What do you think this person is feeling? (please circle one)

Nothing  Surprised  Disgusted  Fearful  Angry  Sad  Happy

257
Imagine yourself **IN** this picture (you are the shadow). You are witnessing first-hand what the other person is feeling.

What do you think this person is feeling? (please circle one)

- Nothing
- Surprised
- Disgusted
- Fearful
- Angry
- Sad
- Happy