The perception of emotion from the body movements displayed during interpersonal communication

Tanya J. Clarke

Thesis submitted for the degree of Doctor of Philosophy
Department of Psychology
School of Human Sciences
University of Surrey
2006
Abstract

This thesis addresses a range of interdependent themes centred on how emotions are picked up from the body movements of actors engaged in natural communication. There were 4 main aims. The first considered the importance of the social context. The second was to further our understanding of the mechanisms responsible for the perception of affect from biological motion. The third considered how the skill to read affect from body movements develops in children. The final aim was to further elucidate what is known about affect. A corpus of point-light displays were created from the body movements of pairs of female and male actors engaged in interlocution whilst emoting. A MacReflex motion analysis system was used to record their movements. Observers viewed either the original and/or various permutations of the original displays and made judgements about the portrayed emotions. The results of the first set of experiments, which tested whether subjects could recognise the emotions and also explored the effects of inversion, were in line with previous findings which showed that emotions can be recognised from biological motion and that inversion impairs recognition. However, this latter effect varied, depending on the emotion shown. Of central importance, the following two experiments considered the social context of the displays. For most of the emotions tested, it was found that seeing a natural interaction made a statistically significant enhancement to how well many emotions were recognised. This finding highlights the need for natural stimuli and the need to consider emotions at a social, as well as an intrapersonal, level. The mechanisms were further investigated in the next set of experiments. It was shown that biological motion of affect is processed differently to biological motion of locomotion. The final investigation focused on how children develop this ability: Children as young as five are able to identify some of the emotions shown and by age nine development is almost at adult levels. In summary, the findings reported in this thesis show that the social context is important in emotion and future studies should incorporate the social context in the stimuli. They also lend support for a) categorical theories of emotion and b) the existence of distinct modules in the mind.
Dedication

This work is dedicated to Mark Bradshaw; the person who inspired me to become interested in the question ‘how do we see?’ This ability is so elegantly installed in us that we never give the question a thought unless we are provoked into doing so. It is an easy question to ask but not an easy one to answer: Mark enjoyed trying to.

Mark could be an annoying supervisor at times but he was an inspirational mentor with a passion for his work. He was also a kindred spirit.

Mark died in October 2004. He was 43.
Acknowledgements

A huge thank you indeed to Dave Rose, for his invaluable support and advice and for being patient, kind and encouraging. Also, I would like to express my gratitude to Sarah Hampson and David Field for their encouragement and enthusiasm. Thank you to all of my friends including Sarah Livemore, Lucy Francis, Tanika Kelay, Ruth Justice, Deborah Kerr, Jane Hall, Tushna Vandravela, Mariana DaSilva and Leslie Notman and Ivan Chandler. Last but not least I am grateful to my parents for being supportive — always.

This thesis is for Michael and Alex
Contents

List of Figures ..................................................................................................................... ix
List of Tables ....................................................................................................................... xii

1. Introduction
   1.1 General Introduction .......................................................................................... 1
   1.2 Emotions ............................................................................................................. 6
   1.3 The importance of body movements and the evidence that body movements specify affect ................................................................. 19
   1.4 A naturalistic approach to event perception ...................................................... 30
   1.5 Summary ............................................................................................................ 34

2. Methods
   2.1 Overview ........................................................................................................... 36
   2.2 Actors ................................................................................................................ 36
   2.3 Script .................................................................................................................. 37
   2.4 Movement recording equipment ........................................................................ 38
   2.5 Marker and actor positions .............................................................................. 38
   2.6 Recordings ........................................................................................................ 39
   2.7 Dependent measures ......................................................................................... 41
   2.8 Subjects ............................................................................................................. 45

3. Effects of upright and rotated displays
   3.1 Introduction ....................................................................................................... 46
   3.2 Experiment 1 ..................................................................................................... 46
   3.3 Experiment 1: Method ....................................................................................... 51
      3.3.1 Subjects .................................................................................................. 51
      3.3.2 Apparatus .............................................................................................. 51
3.17 Experiment 2: Method ............................................................... 80
3.17.1 Subjects ..........................................................................................80
3.17.2 Apparatus, design, stimuli and procedure ....................................80
3.18 Experiment 2: Results .....................................................................................82
3.19 Experiment 2: Discussion .............................................................................84
3.20 Sex differences ...............................................................................................86
3.21 General discussion .........................................................................................89

4. Effects of social interaction
4.1 Introduction ...................................................................................................95
4.2 Experiment 3 ..................................................................................................96
4.3 Experiment 3: Method ..................................................................................99
4.3.1 Subjects ...........................................................................................99
4.3.2 Design and procedure.....................................................................99
4.4 Experiment 3: Results and discussion ...........................................................100
4.5 Experiment 4 ...................................................................................................103
4.6 Experiment 4: Method .......................................................................106
4.6.1 Subjects ...........................................................................................106
4.6.2 Design and procedure.....................................................................106
4.7 Results and discussion ....................................................................................108
4.8 General discussion ..........................................................................................110

5. Further analysis of the mechanisms
5.1 Introduction ....................................................................................................114
5.2 Experiment 5 ..................................................................................................115
5.3 Experiment 5: Method ...................................................................................118
5.3.1 Subjects ...........................................................................................118
5.3.2 Design and procedure.....................................................................118
5.4 Experiment 5: Results and discussion ..........................................................118
5.5 Experiment 6 ..................................................................................................121
6. Children's perception of emotion from biological motion
   6.1 Introduction .............................................................................................. 138
   6.2 Method ...................................................................................................... 142
       6.2.1 Subjects .......................................................................................... 143
       6.2.2 Stimulus materials and apparatus ..................................................... 143
       6.2.3 Design and procedure ...................................................................... 144
   6.3 Results ....................................................................................................... 146
   6.4 Discussion ................................................................................................ 152

7. Discussion
   7.1 Overview .................................................................................................. 156
   7.2 Summary of main findings ....................................................................... 156
   7.3 Implications for the study of emotion ...................................................... 159
   7.4 Validity, limitations and future research ................................................ 162
   7.5 Conclusion ............................................................................................... 167

8. References ................................................................................................... 168
List of Figures

Chapter 2

Figure 2.1. Schematic diagram (not to scale) of the recording set up. The thick, solid black lines show connections for the recording system, the dark dashed lines show the confinement area for the actors and the lighter dashed lines show the distance between the actors and the camera. ................................................................. 40

Figure 2.2. Static frame of a point-light display showing the two actors where each actor is represented by 13 points of light. In this frame the actor on the left has his left arm occluded by the rest of his body. ........................................... 42

Figure 2.3. An image of the dependent measure used for most of the experiments reported in this thesis (a), and (b), an example of how the image might have looked after 5 out of 6 of the scales had been rated (for correctly recognising an anger portrayal) ........................................................................... 44

Chapter 3

Figure 3.1. Observers' mean ratings for the a) upright and b) inverted presentations of the emotion response categories of the portrayed and non-portrayed emotions. The mean ratings for the non-portrayed emotions were derived from the mean of the combined scores from the five non-portrayals in each case. Error bars show standard errors ................................................................. 53

Figure 3.2. Observers' mean ratings as a function of viewing block order for the a) upright and b) inverted presentations of the emotion response categories of the portrayed and non-portrayed emotions. The mean ratings for the non-portrayed emotions were derived from the mean of the combined scores from the five non-portrayals in each case. Portrayed UF and portrayed IF denote whether upright (UF) or inverted (IF) stimuli were shown first to observers. Error bars show standard errors .............................. 59

Figure 3.3. Observers' mean ratings as a function of experiment for a) upright and b) inverted presentations of the emotion response categories of the portrayed and non-portrayed emotions. The mean ratings for the non-portrayed emotions were derived from the mean of the combined scores from the five non-portrayals in each case ........................................... 63

Figure 3.4. a) Observers' mean ratings for each of the upright stimuli used in Experiment 1 as a function of total showing time; b) shows the subset of portrayals used for Experiment 1b ................................................................. 67

Figure 3.5. Observers' mean ratings for the upright presentations of the emotion response categories of the portrayed and non-portrayed emotions shown in Experiments 1 and 1b. The mean ratings for the non-portrayed emotions were derived from the mean of the combined scores from the five non-portrayals in each case. Error bars show standard errors ................................................................. 69
Figure 3.6.
Observers' mean ratings for the 90° rotated presentations of the emotion response categories of the portrayed and non-portrayed emotions. The mean ratings for the non-portrayed emotions were derived from the mean of the combined scores from the five non-portrayals in each case. Error bars show standard errors .......................................................... 82

Figure 3.7.
Observers' mean ratings for the upright presentations of the 'correct' response category for the portrayed emotions as a function of sex. Error bars show standard errors .......................................................... 88

Chapter 4
Figure 4.1.
Observers' mean ratings for the emotion response category portrayed by the actors when displayed as natural dyads, monads and reflections. Error bars show standard errors .......................................................... 102

Figure 4.2.
Accuracy of first speaker identification as a function of emotion. Error bars show standard errors. Solid line denotes chance score .......................................................... 109

Chapter 5
Figure 5.1.
Observers' mean ratings for the backwards presentations of the emotion response categories of the portrayed and non-portrayed emotions. The mean ratings for the non-portrayed emotions were derived from the mean of the combined scores from the four non-portrayals in each case. Error bars show standard errors. .......................................................... 120

Figure 5.2.
Illustrations of the presentations shown where a single frame of an anger monad is displayed in each of the 5 subconfiguration formats: a) The whole body was removed except for the head and shoulders, b) The torso was eliminated, c) The interlimb joints were removed so that internal joint angles were invisible, d) The arms (wrists and elbows) were omitted, e) The whole body was left out except for the wrists and arms .......................................................... 126

Figure 5.3.
Summary of results for each of the (correct) portrayed emotions for every stimulus configuration. Error bars show standard errors .......................................................... 127

Figure 5.4.
Observers mean ratings for presentations of the emotion response categories (a) anger, b) joy, c) sadness and d) fear) of the portrayed and non-portrayed emotions organised by stimulus configurations. Error bars show standard errors .......................................................... 128-130
Chapter 6

Figure 6.1.
Static frame of a point-light display showing the two actors where each actor is represented by 13 points of light. In this frame the actor on the left of the screen is expressing anger and the actor on the right is expressing joy. ........................................................................................................................... 144

Figure 6.2.
Accuracy of emotion identification as a function of age and emotion category. Error bars show standard errors. The results above each bar show results of one-sample t tests (df = 13 for the adults and 17 for the children) that compared actual performance with that expected by chance ...................................................147

Figure 6.3.
Mean accuracy of anger recognition as a function of age broken down by foil type. Error bars represent SEM. The results above each bar show results of one-sample t tests that compared actual performance with that expected by chance. ........................................................................................................................................149

Figure 6.4.
Mean accuracy of joy recognition as a function of age broken down by foil type. Error bars represent SEM. The results above each bar show results of one-sample t tests that compared actual performance with that expected by chance. ........................................................................................................................................150

Figure 6.5.
Mean accuracy of sadness recognition as a function of age broken down by foil type. The results above each bar show results of one-sample t tests that compared actual performance with that expected by chance. ........................................................................................................................................150

Figure 6.6.
Mean accuracy of fear recognition as a function of age broken down by foil type. The results above each bar show results of one-sample t tests that compared actual performance with that expected by chance... ........................................................................................................................................151
List of Tables

Chapter 1

Table 1.
Summary of the major studies where the perception of affect from point-light displays has been investigated ................................................. 26

Chapter 3

Table 3.1.
Observers' mean ratings for each emotion response category when (a) upright and (b) inverted point-light stimuli were presented. Figures in bold type denote the mean ratings for the portrayed emotions. Standard deviations are shown in parentheses .................................................. 54

Table 3.2.
Observers' mean ratings for each emotion response category when stimuli were blocked and counterbalanced by orientation. The ratings have been collapsed by presentation order and show (a) uprights and (b) inversions. Figures in bold type denote the mean ratings for the portrayed emotions. Standard deviations are shown in parentheses ........................................................................ 60

Table 3.3.
Table of observers' free responses to the point-light displays. Words are typed in capital letters when the portrayed emotion is expressed directly .............................................................................................. 72

Table 3.4.
Observers' mean ratings for each emotion response category when stimuli were upright and rotated 90°. Figures in bold type denote the mean ratings for the portrayed emotions. Standard deviations are shown in parentheses ........................................................................ 83

Table 3.5.
Observers' mean ratings for the upright portrayals shown in Experiments 1, 1a and 2. The ratings for males and females are shown separately. Figures in bold type denote the mean ratings for the portrayed emotions. Standard deviations are shown in parentheses ........................................................................ 88

Chapter 4

Table 4.1.
Observers' mean ratings for each emotion response category when dyads, monads and reflected stimuli were presented. Figures in bold type denote the mean ratings for the portrayed emotions. Standard deviations are shown in parentheses ........................................................................ 101
Chapter 5

Table 5.1.
Observers' mean ratings for each emotion response category when backwards point-light stimuli were presented. Figures in bold type denote the mean ratings for the portrayed emotions. Standard deviations are shown in parentheses ................................................................. 119

Chapter 6

Table 6.1.
Observers' mean scores for each correct emotion category broken down by emotion category of the accompanying foil. Standard deviations are shown in parentheses. ........................................................................... 148
Chapter 1
Introduction

1.1. General introduction

Humans are rarely completely still and, “throughout our evolutionary history, being an animal has been reliably — if imperfectly — correlated with self-generated motion” (Cosmides and Tooby, 1994, p 100). Therefore, animal movement or ‘biological motion’ is a major property of the world. Humans are inherently social and one of the supreme achievements of evolution is that vision allows us to pick up exquisitely the movements made by other conspecifics in order to make sense of our surrounding social environment. Sure and immediate understanding about the meaning of others’ movements has an impact on our survival. Expressed via movement of the face and body and by speech and vocalisation, emotions are also a fundamental part of being human. Emotions shape our own existence as how we communicate them has implications for how others behave, which in turn affects our own behaviour. When communicated via movement, it is clear that emotions can be expressed and perceived via facial (e.g. Ekman, 2003) and bodily channels (e.g. Dittrich, Troscianko, Lea, and Morgan, 1996; Wallbot, 1998). Four important areas of enquiry with respect to the latter context, which are the focus of this thesis, are:

i) To see whether, and if so how, the social context is important for the perception of affect\(^1\) from biological motion.

ii) To further our understanding of the mechanisms responsible for the perception of biological motion.

iii) To investigate how the skill to read emotion from body movements develops in childhood.

iv) To further elucidate what is known about emotions.

Johansson (1973), in his seminal paper on biological motion perception described a method by which pictorial cues such as colour, texture, boundaries and edges could

---

\(^1\) The terms ‘emotion’ and ‘affect’ are used interchangeably in this thesis.
be omitted from any display of an event. Therefore, the only information left was about the motion of the structure. Any moving structure could be subjected to this treatment but naturally human movement has been the main focus of attention. Initially, this effect was achieved by attaching bright spots to the major joints of an actor and recording in darkness. Since then, Johansson’s method has been replicated many times and more technologically advanced procedures have been employed to create such displays. The term ‘point-light display’ is coined to describe this type of presentation as typically all that can be seen in a display is a collection of light points on an otherwise blank background. For example, between 11 and 13 points of moving light are typically used to show a point-light human. Subjects can easily identify correctly activities such as running, walking and jumping in these displays (e.g. Johansson, 1973). When such a display is shown every viewer can almost instantaneously pick out a human performing the action and, importantly, movement is critical for these types of displays to be identified as when observers view a static display of a point-light figure they do not recognise the figure as being human (Johansson, 1976). This technique has been invaluable for grasping an understanding of biological motion as it strips away almost all of the other pictorial cues in the optic array that would still exist in normal ‘full-light’ video displays. Thus, the information remaining is from movement.

This thesis addresses the perception of point-light displays and in particular the use of this technique to explore how emotions are perceived when only cues from body movement are available. However, a different approach to this question is taken as a shift is made to consider emotions at the social rather than the more traditionally examined intrapersonal level. Therefore, the point-light displays used as stimuli in this thesis are novel because they show two people (a ‘dyad’) conversing naturally to explore how emotions are perceived when they are portrayed in a communicative context.

---

2 The term ‘event’ is used throughout this thesis as a term for occurrences that involve dynamic movement as the key feature.
3 Some spatial structure from the relationships between the point-lights also remains.
The four empirical chapters in this thesis address a range of interdependent themes concerned with how affect is recognised from body movement. There is now considerable empirical evidence that shows emotions can be readily perceived from full-light bodily expressions (e.g. Dittrich et al., 1996; Wallbot 1998) and, more impressively, reduced cue situations (point-light displays) of the same types of movements (e.g. Dittrich et al., 1996; Atkinson, Dittrich, Gemmell, and Young, 2004). As reviewed in detail below, point-light displays that have been used in the past to convey the expression of affect have been useful; however, it is important that there should be consideration for viewing the subject from a social angle.

The first of the empirical work is discussed in Chapter 3. The research here examines the nature of perception when emotions are expressed via biological motion in more natural communication behaviours than have been previously explored. Therefore, the work in this chapter first assesses the utility of the behavioural samples used as stimuli for the rest of the research reported in this thesis and also allows comparisons to be made between the results of this work and that from other research laboratories.

The experiments reported in Chapter 3 address the question of how readily emotions can be identified from interpersonal expressions of them and how manipulations of display orientation affect perception. To alter the orientation in biological motion displays is a common procedure that elucidates the processing mechanisms responsible for recognition of such displays in that it reveals how the images might be processed by the perceptual system. For example, when inverting point-light stimuli on the image plane, intrastimulus kinematics such as velocity of the dots and relative physical positioning remain constant. Therefore, the percept cannot be accounted for by low-level processing — where the image is thought to be constructed by the visual system computing the relative motions of the dots (e.g. Webb and Aggarwal, 1982).

Both qualitative and quantitative measures of analysis are discussed in Chapter 3; the qualitative work attempts to validate the quantitative results. The chapter also investigates the issue of whether males and females perceived the stimuli differently from each other.
Importantly, the work in Chapter 4 (Experiments 3 and 4) investigates how the perception of emotion is modified by the social context; an area that has so far been largely neglected in this research field. The aim of the work described here was to take a first step at examining certain aspects of communication with regard to recognition of affect from biological motion. Emotions should be considered at a social level as emotions are ‘intentional’, i.e. they are always ‘about’ something and these things are usually people or events (Manstead, 2005). Furthermore, during communication, emotions are frequently shared between people. Two important and novel issues are explored in the chapter. The first pertains to the importance of the interaction between the dyads in the display. In this experiment the displays are manipulated so that the natural dyadic exchange is removed to see if perceptability is an emergent property of the social interaction. The second investigates the communicative interaction independently of affect. It is asked whether the third person observer (i.e. the subject in the experiment) is able to identify the interlocutor from the minimal information available.

Chapter 5 follows on from the work in Chapter 3 that investigates the rotated displays as it contains further experiments where the visual information specified in each display is manipulated in order to investigate the mechanisms that might be responsible for the percept. The first experiment (Experiment 5) concentrates on whether playing the displays in reverse affects recognition. Recent evidence shows that point-light displays of walking dogs (Pavlova, Krägeloh-Mann, Birbaumer, and Sokolov, 2002) and walking humans (e.g. Vanrie, Dekeyser, and Verfaillie, 2004) can still be recognised when this manipulation is carried out although some perturbations are documented in both cases. These observations, in combination with the general findings regarding effects of display orientation, have led Pavlova et al. (2002) to make important claims about the organisation of the mechanisms responsible for biological motion. Experiment 5 will therefore explore whether such claims are transferable to abstract entities, i.e. emotions rather than concrete activities.
It has been demonstrated that there is considerable redundancy in a full-body point-light walker and findings are suggestive of a system that categorises components of the body (Pinto and Shiffrar, 1999). However, whether this is true for the specific case of affect is a question that is so far unanswered. Therefore, the second experiment reported in Chapter 5 (Experiment 6) is concerned with the minimal stimulus cues from which information is available for affect within a point-light display.

Despite the fact that event perception is a thriving research area there have been very few studies that consider how such skills develop in childhood. There are some studies to show that the ability to perceive the information contained within point-light displays develops early in infancy (e.g. Arterberry and Bornstein, 2002; Bertenthal, Proffitt & Kramer, 1987; Fox and McDaniel, 1982). Further, there is no shortage of studies conducted to assess typically developing children’s ability to recognise emotion; for example to identify emotion from faces (e.g. Izard, 1971; Walden & Field, 1982), voice prosody (Stifter & Fox, 1987), music (Cunningham & Sterling, 1988; Boone & Cunningham, 2001) and static body postures (Tracy, Robins, and Lagattuta, 2005). However, these two research areas have rarely been combined in a way that sets out to study how the skill to recognise affect from body movement develops. This imbalance is addressed and examined in the final empirical chapter (Chapter 6) where the developmental trajectory for acquisition of this skill is assessed by studying how well 5-, 7- and 9-year-old children can identify emotion from the displays. Finally, in Chapter 7 it is discussed how the work contributes to the topic of emotion. Research methods and future experiments are also discussed in this chapter.

The remainder of this chapter discusses the theoretical and empirical backgrounds to each of these issues beginning with a discussion on emotions.
1.2. Emotions

Emotions are complex but here follows a restricted review which is largely related to the studies reported in this thesis. Emotions are clearly a fundamental aspect of being human and have been discussed at least since the times of Aristotle. They play a significant part in our everyday lives and therefore, the topic of emotion cannot be ignored.

Damasio (2000) has defined emotions as “a collection of physiological responses triggered by certain brain systems when the organism represents certain objects or situations” (p-15). This definition is a biologically based ‘hardwired’ view and it stands next to the Darwinian position that emotions are adaptive. Clearly, there are other viewpoints but most psychologists recognise the fact that emotions serve a functional purpose and would not dispute the fact that there are biological mechanisms responsible for their production.

There has been vast attention devoted to the topic of emotion in recent years and yet there are only a few fundamental characteristics of affect that many theorists agree about. For example, most theorists agree that there are some ‘basic’ emotion categories which should receive higher status than the rest, although there is little consensus about which emotions should have a higher status and how many of them should be considered in this way (Ortony and Turner, 1990). For example, Watson (1930) considered that only the emotions of love, fear and rage were fundamental. In contrast, Frijda (1986) has considered that as many as 18 emotions should be considered as being basic. Yet a few theorists advocate the position that there are no basic emotions. For example, Averill (1994) has argued that it is not profitable to consider a small subset of basic emotions as it means that nonbasic emotions will receive less attention. However, nearly all theorists include anger, joy, sadness and fear as basic emotions (Ortony and Turner, 1990). Also, there is much support for disgust to be included in this list of basic emotions (e.g. Izard, 1977). Although emotions tend to be discussed in terms of discrete categories more than one emotion
can be experienced at the same time, for example one can be angry about a friend’s behaviour but feel sad for the friend at the same time (Ekman, 2003).

Emotions have multiple components which include physiological, expressive and experiential (feeling) components which serve inter- and intra-personal functions (Ekman, 1994). There is substantial evidence to demonstrate that emotions have a physiological basis. For instance, Damasio (2000) has shown that damage to the prefrontal cortex is inversely associated with patients’ abilities to process emotion normally.

Most emotions have clear behavioural characteristics for that emotion; for instance, hot anger has very obvious expressive signals that are expressed via the voice, the face and the body. These external signals are also observable in other primates (Ekman, 1994) and particular facial characteristics for specific emotions can be observed cross-culturally (e.g. Ekman and Friesen, 1974; Ekman, 2003). However, several emotions such as contempt have few external signs (Ekman, 2003).

There is some support for the view that emotions are involuntary and short-lived patterns of thought and behaviour which have a quick onset (e.g. Ekman, 1994). That they have a quick onset pertains to the adaptive value of emotions as they serve to enable us to act quickly in response to important events (Ekman, 1994). As such, emotions have developed over evolutionary time to prepare an animal to act in response to environmental stimuli. However, Damasio (2000) has argued that emotions have varied time profiles, some have a quick onset but others have a gradual onset and a slow decay, which can best be considered as “background emotions” (p-16). Ekman (1994) has argued that any state that does not have a quick onset should be considered a mood. The distinction between emotions and moods, at the present time, appears to be unclear.

Although most authors agree that emotions serve as functional mechanisms, there is also the view that emotions serve no purpose. Emotions do not appear to be functional all of the time; for example, learned helplessness could be equated to apathetic sorrow (Frijda, 1994a). This author has argued that emotions may simply
be remnants of past adaptive responses but serve no present functional purpose, or alternatively, that they were never functional — they may be quite simply an expression of a person’s awareness of an event rather than an adaptive mechanism.

Many authors argue that emotions have valenced states, that is, they appear to be either positive or negative (pleasant or unpleasant). Ortony and Turner (1990) have argued that to be affectively valenced is a necessary state for an emotion, otherwise it is simply considered as a cognitive state. Surprise does not have any particular valenced state (see Ortony and Turner, 1990; Ekman, 2003) and so accordingly should not be considered as an emotion (Ortony and Turner, 1990). However, terms of ‘negativity’ and ‘positivity’ might be best avoided altogether as attributes of emotions. For example, Ekman (2003) points out that sometimes negative emotions can be enjoyed — such as having a good cry at a sad film. Further, negative emotions often serve us in a positive way. For example, the feeling of anger or jealousy may serve to protect an important interpersonal relationship if the emotion is directed towards an individual who is a potential threat to that relationship. Furthermore, Buck (1999) has argued that if positive and negative emotions are simply defined in terms of approach or avoidance then anger would be a positive emotion. However, it can be contended that sometimes anger leads to avoidance behaviour (Davidson, 1994).

That emotions have valenced states is a concept that pertains to the structure of affect. There are two main positions that relate to this issue — one of which (already discussed) is that there are discrete categories where each category is qualitatively different from the next. The other position, derived as a result of factor analyses and multidimensional scaling, is that affect can be described in terms of a set of (usually two) dimensions — of which valence is normally considered to be one and the second usually relates to the degree of activation (energy) of the affect (Russell, 1980, Russell and Barrett, 1999). These dimensions are often thought to be inter-related so that all affect can be accounted for in terms of particular combinations of pleasure and activation (Russell, 1980, Russell and Barrett, 1999). In this dimensional model, affect is generally represented in a two-dimensional space in which emotions fall in a circle (sometimes called a circumplex) where, for example, happiness falls at 0° and
distress at 135° (Russell, 1980 — see Russell and Barrett, 1999 for various schematic representations of dimensional accounts).

Clearly, it can be seen from the short discussion above that the topic of emotion is complicated and that authors often disagree. However, disputes regarding the topic can only be solved by a continuation of scientific exploration.

Here follows a brief description for each of the emotion categories that will be investigated in this thesis. Romantic love is included and discussed to a greater extent than the other categories as it is suggested that this state should be considered as an important emotion.

**Anger**

The word anger covers numerous experiences from an annoyance that someone has parked tightly next to your own vehicle to the rage that is felt if your teenage child is attacked as they walk home from a party. Ekman (1994) has postulated that the antecedents for anger include threats and insults and an incident that violates one’s own values. Thus offensive acts towards an individual, or the individual’s family and possessions, can be considered as the antecedent to this emotion. The action tendency of anger is to attack or aggress, or sometimes to withdraw (Davidson, 1994). Anger produces increases to normal heart rate which demonstrates the involvement of the sympathetic nervous system (which becomes active in response to aversive stimuli) and is responsible for ‘fight or flight’ behaviour. (Le Doux, 1994).

**Fear**

Fear has probably received the most attention of all the emotions. Fear is an aversive emotional state that performs the role of mediating escape in times of danger (Öhman, 2000) and similar expressions are observable in other animals which suggests they also experience fear. There is a detectable physiological basis for this emotion as there are strong skin conductance responses to fear-evoking stimuli such
as spiders and snakes (e.g. Flykt, 2005). Like anger, fear produces increases to normal heart rate which suggests the involvement of the sympathetic nervous system (Le Doux, 1994). Furthermore, there is overwhelming evidence that there is a specific region of the brain — the amygdala — that is heavily involved in the processing of fear (LeDoux, 1994; Öhman, 2005). The amygdala is served by a fast subcortical processing highway that does not involve the cortex (Le Doux, 2000). There are vast differences in the ways that fearful experiences are dealt with which depend on the severity of the threat but the general action tendency is to escape, become still or to attack (Öhman, 2005).

Sadness

Unlike the functional accounts for fear or anger, it is harder to understand the motivation for the feeling of sadness. Most theorists agree that sadness is the resultant state that is associated with a particular loss or failure to achieve important goals or a failure in important social relationships (Power, 1999). Sadness may be seen as a cry for help although it is not necessarily manifested by crying (Clore, 1994). Emotional disorders such as depression are related to sadness (Power, 1999). There is neuroscientific evidence to suggest that the amygdala is involved in the processing of sadness stimuli (e.g. Adolphs and Tranel, 2004).
Romantic Love

Like sadness, love is normally focused on a particular person or an event. Love in general is associated with the involvement of affection. This is the most contentious of all the categories included for investigation here as romantic love is not regarded as an emotion by all psychologists. In particular, Ekman (2003) considers that because romantic love is an enduring state it does not fit in with the criteria that emotions are briefly felt. However, Ekman (2003) does concede that one aspect of love, which he describes as extreme pleasure defined as ecstasy or bliss can be considered an emotion due to its fleeting presence.

Love was considered to be an important emotion by some scholars in the last century. For example, Watson (1930) only considered that there were three emotions, of which love was one. Further, Harlow (1958) wrote a classic paper on the nature of love and considered love to be a fundamental emotion. Love is also considered as a basic emotion in some modern taxonomies (e.g. Shaver, Wu and Schwartz, 1992). Indeed, the view by a layperson is that love is an important emotion: if participants are asked to list what they consider to be emotions, love is one of the most common responses (Fehr and Russell, 1991).

Hatfield and Rapson (1993) argue that scholars draw a distinction between two types of love — ‘passionate love’ and ‘companionate love’. These authors postulate that passionate love is an intense emotion that is a state when an individual longs for another. In contrast, Hatfield and Rapson (1993) argue that companionate love is a far less intense emotion which combines feelings of attachment and commitment as well as intimacy. Therefore, Hatfield and Rapson’s definition of passionate love may be similar to Ekman’s (2003) definition of love as a state of bliss.

That love is an emotion expressed when another person is desired shows that this emotion relates to approach behaviour (Gonzaga, Keltner, Londahl and Smith, 2001). This view is substantiated by evidence that the experience of love is correlated to
Duchenne smiles\(^4\), head nods, nonhostile arm and hand movements and leaning towards the partner (Gonzaga et al., 2001). Furthermore, from the analysis of videotapes of couple's facial expressions, interpersonal synchrony and proximity were found to be important correlates of romantic love (Gada, Bernieri, Grahe, Zuroff and Koestner, 1997). Clearly, it can be seen that love has a strong interpersonal element.

Love also has a physiological basis as the expression of this emotion is linked to the production of oxytocin which is a neuropeptide associated with bonding behaviour (Carter, 1998).

**Joy**

Frijda (1994a) has argued that it is hard to discern the function of joy. However, as discussed shortly, joy may be better understood if it is considered at an interpersonal rather than an intrapersonal level. Joy is an emotion that is contingent on goal achievement but it is most likely to be related to more than just the satisfaction in achieving the goal (Clore, 1994). For example, the level of feeling joyous will depend on a number of other factors which may be based on personality, the satisfaction placed on biological needs (such as sex and food) and/or relief after problematic or demanding activities have ended (see Frijda, 1994a; Clore, 1994). Ekman (1994) has argued that the antecedents for joy are pleasure, praise, relief and excitement. Unlike love and sadness, joy may not necessarily be focussed on a particular person or event (Averill and More, 1993). However, it is reasonable to suggest that the same point could be argued for sadness.

\(^4\) A Duchenne smile is created by using the zygomatic major muscle (which pulls the lips up into a smile) and the orbicularis oculi muscle (which creates wrinkling around the eyes).
Disgust

There is some dispute between theorists as to what exactly disgust relates to. Rozin and Fallon (1987) argued for a narrow definition of disgust in that it is a food related emotion: it is adaptive in that it serves to protect the body from noxious substances. Alternatively, Rozin, Haidt and McCauley (2000) postulated a two-factor model of disgust that consisted of 'core' disgust which is the threat of eating rotten food, for example. The second factor related to 'animal reminder disgust' which was postulated to be associated with death and poor hygiene. Ekman (2003) also concedes that disgust is related to food in terms of feelings of repulsion towards noxious sensory stimulation. However, Ekman also argues strongly that moral disgust is an important factor to consider when defining disgust. Ekman (2003) argues that the most intense feelings of disgust relate to physical repulsion (such as the Nazi atrocities that occurred in concentration camps). Anger is an emotion most confused with disgust and anger reactions can turn into disgust over time (Ekman, 2003). There is less involvement of the sympathetic nervous system for disgust than for some of the other emotions discussed here as the experiences of anger, fear or sadness produce a larger increase in heart rate than disgust (Levenson, 1994).
Emotions are social

The prevalent view of emotions is that they function at the private, intrapsychic level. This assumption is not contended here as emotions certainly have physiological and cognitive components and they function to serve us individually. However, this is not the only view. Humans are social in nature and we are biologically adapted to living in social groups (Dunbar, 1998). Furthermore, we depend on each other for survival. For example, collecting resources, dealing with threatening situations and raising families are behaviours shared between people (Ainsworth, 1989).

Emotions are a significant part of our everyday social life: people are important to one another and they make appraisals about each other's emotional states. For example, the emotions of love, hate, anger, jealousy, envy and contempt are centred on interpersonal appraisals (Parkinson, 1996) and Schaver, Wu and Schwartz (1992) found that out of 600 written descriptions of emotions by lay people around three quarters featured interpersonal relationships as a key theme in the description.

That emotions are communicative in nature was a position also advocated by Darwin (1872, 1999) — although Aristotle also discussed the social element of emotion. Darwin alluded to the close coupling of emotions and communication throughout his work, as the anecdotes described by him were often inherently social in nature. For example, when discussing anger he gave an example of a caught murderer who was surrounded by a mob of angry people “snarling with their teeth and making at him like wild beasts” (p-239). However, Darwin did not relate the communicative value of emotions to the adaptive principles that he advocated, which Ekman (1999) considers to be a fundamental omission: Ekman (1999) postulated that emotions are expressed in order to communicate to others, informing the observer that something important is happening to the expressor.

Emotions are adaptive in that they function to serve us at this social level, for example, to maintain the cohesiveness of a group or to maintain familial bonds in
attachment (or to avoid interaction altogether, Levenson, 1994). For example, an emotion such as joy not only has the function of maintaining the individual’s behaviour but signals to others to continue their interaction with the individual (Emde, 1988). Furthermore, momentary displays of romantic love serve to promote and maintain long-term commitment behaviours in relationships (Gozanga et al., 2001). The communicatory function of emotions occurs from an early age, as illustrated by the way that infants will look towards their caregiver for assurance that a novel object can be approached. If the caregiver smiles the child will approach the object but if the caregiver expresses fear, disgust or anger the child will not approach the object (Klinnert, Emde, Butterfield and Campos, 1987). The social exchange of affect is two-way, for example, toddlers have been shown to use particular displays of affect (sadness) in order to gain support from their caregivers (Buss and Kiel, 2004).

Thus, rather than thinking of emotions as states that reside only at the intrapersonal level, emotionality can be thought of as a dynamic relational process that occurs between the individual and the environment (Campos, Campos and Barrett, 1989; Blair, 2003). Therefore, the study of emotions at a social level must be of interest as emotions have interpersonal or intergroup regulatory outcomes.

Parkinson (1996) has argued that the study of emotions at a social level should replace cognitive and physiological approaches. Parkinson’s (1996) view is based on the argument that emotions should be seen as ways to communicate between people. For example, Parkinson (1996) advocates that fear signals ‘help me’ and sadness signals ‘comfort me’.

Whilst the interpersonal and communicative aspects of emotions are important, a balanced view is required. Therefore, it is argued here that the investigation of emotions should include a social approach to supplement rather than replace enquiries made at physiological and cognitive levels.
Mechanisms and behavioural evidence to suggest information is processed at a social level

This discussion now turns to the growing physiological and psychological evidence to support the view that emotions should be considered at a social level. First, the evidence to suggest that there are neurons that function to guide social behaviour is reviewed.

In the neuropsychological literature it has been proposed that there is a coding system for action orientated behaviour at an interpersonal level (see Gallese, Keysers and Rizzolatti, 2004). This is based on the finding of so called ‘mirror neurons’ in monkeys’ brains that are activated not only when a monkey performs a task but also when a conspecific performs the same task (Gallese et al., 2004). This theory has been extended to humans as there is accumulating evidence that similar neural mechanisms exist in humans; in functional magnetic resonance imaging (FMRI) studies the same brain areas are involved for observing an action in another person as when the observed action is executed by the subject (Gallese et al., 2004). Similar findings have been reported that extends this theory to other sensory modalities; participants have been found to show activation in pain related areas of the brain when they watch others be subjected to painful stimuli (Hutchinson, Davis, Lozano, Tasker and Dostrovsky, 1999). Further, there may also be other shared interpersonal representations in humans (Thomas, Press and Haggard, 2005). These authors found that subjects were faster to respond to tactile stimulation to their own body when viewing a visual cue in a congruent (rather than an incongruent) anatomical location on a model’s body.

For the special case of affect, the role of the insula has recently been found to be fundamental in the feeling and understanding of disgust (Gallese et al., 2004). The same area of the insula is activated in FMRI studies when a subject experiences an unpleasant odour and when they are exposed to facial expressions of disgust. This evidence has led Gallese et al. (2004) to suggest that the same neural mechanisms may be responsible for first and third person experiences of emotion.
However, Jacob and Jeannerod (2005) contend the view that there are neurons in the brain that exist to assist in social behaviour. These authors have argued that the existence of mirror neurons does not prove that these are the social mechanisms for intentional communication; rather they consider imitating another person’s movements is only likely to be sufficient for understanding his or her motor intention not their social aims. Furthermore, these authors argue that matching behaviour is not equal to social behaviour. Moreover, matching behaviour is not adaptively advantageous — for example, when exposed to threatening stimuli it is more useful to flee rather than imitate another person’s movements (Jacob and Jeannerod, 2005). Indeed it can also be argued that single cell recordings do not give the whole story (Rose, 1996).

Even if, as yet, there is no concrete physiological basis for how social knowledge is processed, behavioural evidence does show that individuals are influenced by the emotions of others. Empirical evidence shows that positive emotions are facilitated by interpersonal interactions such as smiling, which was found to be positively correlated with the presence of a friend when happy films were viewed (Jakobs, Manstead and Fischer, 1999). Smiling was also enhanced when friends watched sad films together (Jakobs, Manstead and Fischer, 2001). Surprisingly then, in the latter case, although the stimulus was a sad film the presence of another person promoted smiling behaviour and inhibited the production of sad expressions. Conversely, sad expressions were produced when the sad movies were watched alone. Other evidence shows that emotions can have harmful effects on the outcomes of an event as it has been shown that the demands made by one person of another (the actor) alters depending on the perceived emotional state of the actor (van Kleef, De Dreu and Manstead, 2004). These researchers demonstrated, in an experiment that involved two-way computer-mediated negotiation, that participants conceded more to an angry opponent than a happy one.
Conclusion

The topic of affect is much debated but many authors agree that there are basic categories of emotions which include anger, joy, sadness, fear and disgust. Love might also be considered a basic emotion as there are distinct (interpersonal) displays of love and love appears to have a physiological basis (the production of oxytocin).

The prevalent view of emotions is that they function at the private, intrapsychic level. In this discussion it has been argued that emotions are also social. So far, the physiological evidence in support of this view is weak, however psychological evidence shows that the investigation of affect from a social perspective is an important avenue of enquiry. It is argued that physiological, cognitive and social approaches to the investigation of emotions are equally valuable.
1.3. The importance of body movements and the evidence that body movements express affect

The human visual system has evolved to be able to extract important information from the environment. Movement is one such valuable property as humans and animals are rarely still (Cosmides and Tooby, 1994). Therefore dynamic, rather than static information, should be considered as a foundational feature of perception. The dynamics of behaviour add a substantial amount of information to that in a static display (McArthur and Baron, 1983). For example, when a ball is stationary its weight cannot be known simply by looking at it but can be instantly calculated when watching it roll down a hill (McArthur and Baron, 1983). The following discussion explains why the focus of attention in this thesis is about movement specifically expressed via the body.

Dynamic information conveyed by the body is undoubtedly important. For instance, body movement alone can be used as a form of communication, such as in gestures, in slumping or standing erect or moving quickly or slowly. Moreover, cross-cultural evidence confirms the importance of body movement as a mode of communication: actions directed towards objects by communication of gesture are evident in the cultural expressions of tribal rituals and these gestures act as substitutes for speech (Donald, 1991). Furthermore, body movement in combination with speech (and also with other bodily systems such as facial expression) is fundamental in natural human communication (Bachorowski, 1999; Wachsmuth, 2002; Planalp and Knie, 2002).

Body movement is likely to be a defining feature of our phylogenetic past as, in evolutionary terms, movement of the body is postulated to be a precursor to language acquisition (Rizolatti and Arbib, 1998). These authors suggest that early hominids could have described an object or an event by using referential arm movements which, over time became accompanied by vocalisations. Rizolatti and Arbib (1998) argued it was unlikely for facial movement to be a precursor to language acquisition for two reasons: firstly, oro-facial communication is limited to a few actors, secondly, actions directed towards objects, that often accompany speech are inherent
in body language (Rizolatti and Arbib, 1998). Therefore, speech is likely to have evolved from body movement due to the flexibility and potential expansiveness of the brachio-manual system (arm movements).

Thus, according to Rizolatti and Arbib (1998), the use of body gestures and movements was fundamental as a precondition for language evolution as it allowed for group communication. Proximity is of course still important within the domain of social exchange: it is clearly frustrating to be out of earshot from a speaker and at a distance so that their facial expressions cannot be clearly seen. For body movement to be a cue would mean that humans can pick up information that can be utilised when other communication channels are not easily available. Indeed, the perception of body movement is fundamental as it is likely to be the first information available when approaching or being approached by others.

The foregoing arguments suggest that body movement is a foundational characteristic of the communication system and yet the published literature shows that there has been a tendency to investigate expressions produced via the face rather than body. Almost without exception, all emotions are expressed (and can be perceived) via movement of the face and the body. Affect can be perceived from static facial expressions and from static body postures but it is far more natural to investigate emotion in the way in which it is naturally manifested (Atkinson et al. 2004).

This discussion now reviews what literature there is that has dealt with the specific case of the perception of affect from body movement.
Evidence to suggest that body movement specifies affect: Full-light observational and behavioural studies

Darwin's (1872/1999) classic work detailed specific body movements and postures for many emotions. For example, his description of an angry man was as follows: "He carries his head erect, with his chest expanded and the feet planted firmly on the ground. He holds his arms in various positions with one or both elbows squared, or with the arms rigidly suspended by the sides" (p.242). As well as body movements, Darwin also described the accompanying vocalisations for some emotions. For example, joy comprised "purposeless and extravagant movements of the body" coupled with "the utterance of various sounds" (p.80). After discussion in such eminent research it is surprising to find that the perception of emotion from bodily expression has been researched relatively infrequently in comparison to the amount of work conducted on the expression of emotion in faces. In fact, research into affect from body movement has taken low status. Ekman (2003) emphasises the importance of facial expressions of affect and for instance, when describing anger he states that, "like all emotions, anger has a powerful signal in both face and voice" (p.125). However, Ekman neglects to mention the powerful, high movement activity of the body that also signals this emotion.

Despite the dearth of apparent interest in research that investigates body movements as a medium for expression, there is some evidence to suggest that specific body movements are associated with emotional states.

de Meijer (1989) has demonstrated that emotion attributions can be made from specific types of body movements. In this study, subjects were required to attribute affect to specific movement categories produced by actors shown in full-light (i.e.,video) conditions. For instance, stretched, vertical movements of the whole body were found to correlate with joy and forceful movement was associated with anger. However, a problem with this study was the fact that actors were asked to produce body movements in spatial dimensions and with specific dynamics such as force and
velocity, i.e. they were not asked to express emotion. Therefore, whilst this study contributes to the debate, it does not tell us about the perception of emotion per se.

In a different approach, Wallbot (1998) investigated the specific link between the production and the perception of affect. This author employed twelve actors to produce expressions of affect, which were videotaped. That a large sample of actors were used was an important feature of the design as it has been argued that there are individual differences in actors’ movement styles and therefore this measure allowed for variation in performances (Wallbot, 1998). Fourteen emotions were produced including anger, joy, disgust, sadness and fear. The body movements were then coded from the played (silent) videotapes for the quality of movements produced such as ‘collapsed movement of the upper body’ and ‘arms stretched out to the front’ or ‘stretched out sideways’. It was found, using interrater reliability techniques, that 66% of the posture and movement categories distinguished different emotions. Therefore, these results showed that there appeared to be qualitative differences in movement between emotion categories. For example, a salient feature of anger was found to be lateralized and frontal movements with a lot of movement activity. In contrast, sadness was expressed with a collapsed upper body including forward facing shoulders and low movement dynamics (Wallbot, 1998).

In a similar study, younger and older adults were able to recognise successfully the emotions of anger, sadness and happiness from videotapes (where facial information was blurred) which showed either one actor or two actors (together in the scene) depicting affect (Montepare, Koff, Zaitchik and Albert, 1999). In this study, the actors were given a scene in which to enact the emotion, for example, clapping ones hands (to depict joy). In most scenes the actors worked together, which Montepare et al. (1999) argued would ensure ecological validity; however, the actors were silent during their enactments, the scenes were scripted and it is not made clear as to how the pairs of actors worked together. However, when the videotaped displays were rated for salient movement dimensions similar results were documented as those recorded in Wallbot’s (1998) study. For example, anger displays were rated highly for incorporating movements that were jerky, fast and expansive.
Taken together, the results of these behavioural studies suggest that there might be specific patterns of movement associated with particular affects, which is in line with Darwin’s (1872, 1999) observations. However, the use of the point-light technique in experimental research might allow for firmer conclusions to be made about the importance of information from kinematics for the specification of affect.

Evidence to suggest that body movement specifies affect: Point-light studies

So far it has been shown that emotion is likely to be expressed via movement of the body but these observations have all been made in full-light conditions. A key question, however, concerns the contribution of solely kinematic information from the body that specifies affect. In Wallbot’s (1998) study, information from the face may have been available and, in the de Meijer (1989) study, although actors were instructed to keep neutral facial expressions this may not have always been the case: for instance, micro-expressions of affect cannot be concealed (Ekman, 1993) and it is possible that these would have been discernable in such displays. Even if facial expression can be discounted as a potential source of information for recognition, there are many other cues in the optic array that can be used as cues for perception. In a full-light display there are fundamental properties of the visual world such as continuous surfaces, edges, shapes, background, depth cues, colour and texture, all of which contribute to perception (Gibson, 1950). These properties may provide cues to enable attributions about affect. For example, a red face may signify that an actor is angry. Even a simple static shape such as a triangle can convey emotional meaning (see Pavlova, Sokolov and Sokolov, 2005). Point-light displays remove almost all of these elementary properties of the visual scene. In a point-light display the only information is that of kinematics, which includes reduced spatial information; this cannot be avoided as spatial information is derived from the relations between the point-lights.

From certain experiments there is an increasing amount of evidence to suggest that the information contained within a point-light display of the body is sufficient to specify affect. The major studies to show that affect can be recognised from displays of biological motion are summarised in Table 1. Walk and Homan (1984) were the
first researchers to use the point-light method to investigate affect. By showing subjects point-light displays of different dances and emotions, they demonstrated (using a matching task as the dependent measure) that subjects could recognise fear, anger, contempt, surprise, sadness and happiness. Next, Dittrich, et al. (1996) examined the ability of observers to recognise affect in standardised dance movements. Specifically they investigated surprise, fear, anger, disgust, grief and joy. Observers were shown full-light and point-light displays of affect and ranked them in order as to which emotion each display was most likely to depict. Observers' rankings were recorded and converted into mean scores as a measure of emotion recognition accuracy. Dittrich et al. (1996) found that all of the emotions could be recognised above chance in the point-light condition with the exception of disgust. However, some emotions were recognised with more success than others. Portrayals of grief and joy were better identified than portrayals of fear and anger. Portrayed disgust was the least well identified and it was confused with other emotions (non-portrayals) as similar ratings as those received for disgust were given to fear and sadness non-portrayals. Dittrich et al. (1996) reported that the full-light displays were recognised more readily than the point-light displays. Disgust was also recognised above chance in this condition.

In another study where dance movements were used as stimuli, Brownlow, Dixon, Egbert and Radcliffe (1997) showed dance novices and dance experts point-light displays of happy and sad dances. These researchers demonstrated that perceivers with dance experience were able to make finer judgments about dance movements in which affect was portrayed over novices' judgments of the same. This result shows that experience makes a difference to perception.

Atkinson et al. (2004) replicated and extended Dittrich et al.'s (1996) study by not only showing observers full-light and point-light displays of the same emotions (except for surprise which was omitted) but also by adding a condition where actors exaggerated their body movements by increasing the vigour of the movements. In Atkinson et al.'s (2004) study, actors — rather than dancers — portrayed the emotions. Subjects viewed and rated the displays by using a forced choice method. In both the full-light and point-light conditions, Atkinson et al. (2004) found that
subjects could recognise affect. Similar results were found to the Dittrich et al. (1996) study in that affect could be recognised from both types of displays. Atkinson et al. (2004) found that for disgust, fear and anger affect was more readily recognised in the full-light rather than the point-light displays.

Also, in a slightly different approach, Pollick, Paterson, Bruderlin and Sanford (2001) showed that affect could be perceived from a single body part. Actors were employed to produce knocking (e.g. as if knocking on a door angrily) and drinking arm movements. Subjects were able to recognise the affects of anger, happiness and sadness easily which suggests that arm movements alone enable emotional attribution (Pollick et al., 2001). The data from this study were also subjected to a multidimensional scaling technique which produced a two-dimensional model of the affects. One dimension was defined as ‘activation’, which accounted for 70% of the variance, and the other dimension was defined as ‘pleasantness’, which accounted for 17% of the variance. Pollick et al. (2001) argued that the psychological space relating to this model conformed to a circumplex model of affect. The activation axis was correlated strongly with the physical dimensions of the movements which suggests that information from movement alone may be sufficient for the recognition of affect (Pollick et al., 2001).

Taken together, the results of these studies support Darwin’s (1872/1999) thesis that emotion can be recognised from bodily expression. Further, the view that information from body movement per se is paramount for recognition of affect is augmented. It appears that the emotions of anger, joy, sadness and fear can easily be recognised from biological motion, irrespective of whether the movements are expressed by dancers or actors. There is less evidence that this is the case for disgust. Affect can also be recognised from arm movements in isolation from other parts of the body.
Table 1: Summary of the major studies where the perception of affect from point-light displays has been investigated

<table>
<thead>
<tr>
<th>Study</th>
<th>Mode of expression</th>
<th>Method of investigation</th>
<th>Type of display</th>
<th>Emotions investigated</th>
<th>General findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walk and Homan (1984)</td>
<td>Dances and mime</td>
<td>Free labelling and structured response lists</td>
<td>Point-light displays</td>
<td>Anger, Happiness, Sadness, Fear, Surprise, Contempt</td>
<td>In free labelling only anger was identified. In matching task all emotions were recognised.</td>
</tr>
<tr>
<td>Dittrich et al. (1996)</td>
<td>Stylised dance movements 2 dancers</td>
<td>Observers’ ratings</td>
<td>Full-light Point-light Upright Inverted</td>
<td>Anger Joy, Sadness, Fear Surprise Disgust</td>
<td>Best recognition for full-light movies and lower recognition for point-light displays for all emotion categories but of these joy and grief had the highest scores.</td>
</tr>
<tr>
<td>Brownlow et al. (1997)</td>
<td>Stylised dance movements 4 dancers</td>
<td>7-point bipolar scale</td>
<td>Point-light displays</td>
<td>Happiness, Sadness</td>
<td>Happy dances rated as happier, stronger, more approachable, dominant and extroverted than sad dances by both groups. Experts rated the happy dances differently to novices. No differences were found between novices and experts for sad dances.</td>
</tr>
<tr>
<td>Pollick et al. (2001)</td>
<td>Arm knocking and drinking performances 2 Actors</td>
<td>Forced choice</td>
<td>Upright Inverted and phase scrambled</td>
<td>Anger, Fear, Happy, Sad, Excited (Strong, Relaxed, Strong, Tired, Weak, Neutral)</td>
<td>Affect can be recognised from arm movement. A 2 D circumplex model of affect with activation and pleasantness dimensions. There was more support that movement kinematics correlated with activation dimension than the pleasantness dimension.</td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------------------------------------------</td>
<td>--------------</td>
<td>------------------------------------</td>
<td>-----------------------------------------------------------------</td>
<td>------------------------------------------------------------------</td>
</tr>
<tr>
<td>Acklinson et al. (2004)</td>
<td>Actors movements 10 actors</td>
<td>Forced choice</td>
<td>Full-light Patch-light Full-light stills Patch-light stills Exaggerated movements—where actors were required to produce more animated expressions</td>
<td>Anger, Joy, Sadness, Fear, Disgust</td>
<td>Best recognition for full-light movies but high recognition for patch-light displays for all emotion categories. Happiness was the most easily discerned emotion in the point-light condition. For all emotions except sadness exaggeration of movement correlated with increased intensity. The opposite was true for sadness.</td>
</tr>
</tbody>
</table>
Watson and Raynor (1920), in a classical conditioning experiment, showed that a fear response could be induced in an infant by making a very loud noise near to the infant’s head. However, such studies could not be reproduced today as one of the great problems for psychologists studying affect are the ethical constraints imposed upon them. To ensure that volunteers do not incur psychological harm, ethical restrictions rarely allow for volunteers to experience naturally produced emotions in laboratory settings at the will of the experimenter. This poses a problem for researchers as, although stimuli could be made from video footage (e.g. news clips) of naturally induced affect, there would be no control over the behaviour of persons expressing the emotion, nor would there be any control in the filming techniques used, e.g. the camera in relation to the emoter’s body.

Therefore, as illustrated by the studies discussed so far, to produce stimuli, researchers capture the emoting behaviour of actors and dancers. This procedure avoids ethical constraints and allows control over the actor’s behaviour (i.e. all behaviour, except for the emotional expressions, remains constant) and also the environmental variables (i.e. camera angle, surrounding lighting/noise). However, in terms of ecological validity, the stimuli sets produced in the past have been equivocal. This is despite the fact that one of the most important decisions for the researcher involves the behaviour sampling (Ambady, Bernieri and Richeson, 2000) and the fact that providing unnatural contexts may produce unnatural results (Planalp and Knie, 2002). For instance, as can be seen in Table 1, Dittrich et al. (1996) and Brownlow et al. (1997) used the movements of dancers to produce their stimuli sets. Such procedures can be criticised as, firstly, dance is not an integral part of normal everyday activities (at least in western cultures) — rather it is a ‘stylized art form that involves cultural and aesthetic conventions’ (Van Meel, Verburgh and de Meijer, 1993, p-119) — and secondly, in these instances, this method disallows for any spontaneity in movement production.

Atkinson et al. (2004) attempted to produce stimuli with high ecological validity, recruiting 10 actors (5 female and 5 male) instead of dancers to develop a set of
stimuli. That 10 actors were used meant that a wide range of idiosyncratic styles could be sampled and that a large corpus of stimuli material would be collected. Importantly, Atkinson et al. (2004) ensured that actors were free to choose how to express the given emotions with only minimal guidance given. However, it can be argued that ecological validity was not achieved as i) each actor wore several pairs of tights over his or her head to obliterate facial features and expressions (this is a very unusual experience obliterating visual sensory information/feedback from the environment) and ii) actors were required to produce affect within a small time window (6 seconds) counted aloud by a researcher during filming. The foregoing points show that the actors employed in this study were isolated from both the physical and social environment during the production of the behaviour samples. Further, the imposed time constraints (in combination with rigid instructions about the finishing position) would have been an attentional constraint that might have interfered with the ability to produce the required emotion.

In the Pollick et al. (2001) study, when arm knocking and drinking movements were produced, the actors were required to read a brief story that allowed them to imagine the emotion to be performed. This study is far better in terms of the behaviour samples in that i) actors rather than dancers produced affects, ii) there appeared to be no time constraints to produce the emotions, nor iii) were they asked to wear clothing that restricted sensory perception. However, this work was not ideal on other grounds: i) it is not natural to map affect onto a physical action such as drinking, ii) using one part of the body (i.e. the arm) to express affect lacks ecological validity as this part of the body may not correlate with the way that affects are normally expressed.

The preceding evidence illustrates the incumbency to develop a new approach to investigate the perception of affect. In order to better understand the perception of affect, it is important, as a starting point, to have a description of the stimulus information to which the person responds (McArthur and Baron, 1983). This cannot be achieved by studying the individual in a social vacuum (Tajfel, 1972). Therefore, it may be profitable to think about how emotions are expressed naturally in everyday situations.
1.4. A naturalistic approach to event perception

Gibson (1950) described psychology as the study of the 'perception and behaviour of animals and men as a function of what the environment affords' (p-22), arguing that the visual world is filled with things which have meaning. Gibson (1950/1979) put forward an ecological theory of direct perception that eschewed cognitive processing to explain perception and, instead, contended that the optic array provided everything required for the perceptual system to achieve its goal of vision. Gibson argued that the environment can specify all the information for perception in the form of affordances and that these will vary depending on the needs of the perceiver. For example, a cup of tea has the affordance of 'drinkability' but only if the perceiver is thirsty. Thus an organism will perceive an object only if it is relevant and important to the organism's well being: Gibson called this 'attunement'. Affordances, according to Gibson, are ecological and relational facts about the environment and behaviour that stay stable over time i.e. they do not vary despite changes in the environment and changes in the retinal image. These invariant features of the optic array, according to Gibson, can be used in combination with each other and are all that is needed for perception. Affordances can be physical or social properties of the environment. The richest affordances, according to Gibson, are those provided by other people (Costall, 1995).

Thus, according to Gibson's theory, stimulus events are information rich; however, accurate knowledge about how we perceive the visual world must be sought from an ecological starting point. In other words, the study of artificial stimulus information is of little utility. Advocates of the ecological approach such as McArthur and Baron (1983) have also argued that the information necessary for perception is contained within the stimulus. Accordingly, observers are sensitive to adaptively relevant information in the stimulus. For example, emotions such as fear and anger should be readily perceived as they are the most essential for survival (McArthur and Baron, 1983). However, their account is not limited to adaptive survival processes as
Accurate social perception also means that personal achievements can be made and unwanted situations can be avoided (McArthur and Baron, 1983).

Advocates of the ecological perspective recognise that not all of social knowing can be explained by direct perception; clearly mental processes by which the sensory information is transformed, stored, recovered and used by the mind are essential for a full understanding of perception (Neisser, 1967). In terms of how the mind might be structured, Fodor (1983, 1985) argued that it (the mind) contained a set of distinct modules. According to Fodor (1985), each module, responsible for a specific function (such as face recognition), is "very fast, mandatory, superficial, encapsulated from ...knowledge, largely organised from bottom to top information flow and largely innately specified" (p-4). Fodor also argued that there must be central systems which complement modular systems and exhibit the opposite characteristics such as slow and flexible processing over widely distributed areas. There has been much interest in Fodor's modularity argument to the extent that his claims about general systems have been largely forgotten; this is despite the fact that little thought has been given to the topic of how the modules might be coordinated (Nakayama, 2001).

The foregoing paragraph illustrates that mental processes are fundamental. However, Gibson's ideas have been influential. In line with Gibson's conceptions, the evidence discussed so far with respect to point-light displays shows that we are able to perceive dynamic properties of the social environment. For example, from displays of biological motion humans can pick up a range of socially relevant information such as a person's sex and acquaintanceship (Cutting and Koslowski, 1977), discrimination of a particular walker (from other walkers) irrespective of viewpoint (Troje, Westhoff and Lavrov, 2005) and true behaviours despite intentions to deceive (Runeson and Frykholm, 1983).
An alternative method of investigation

In studies of affect from biological motion, the verbal channel has been neglected. However, communication is embodied within the various sensory modalities (Streek and Knapp, 1992) such as facial and bodily expressions and words and voice prosody (see Scherer, 2003, for a discussion on the expression of affect from the vocal channel). These sensory modalities have not evolved independently of each other, which therefore means that communication should be investigated by measuring more than one channel at a time (Birdwhistell, 1970). More recently this point has been argued by other scholars: for example, Jorgensen (1998) has postulated that, by studying affect from only one modality (e.g. verbal), researchers are not studying valid communication processes and are only studying parts of the whole. Furthermore, expressing emotion in one modality may facilitate its release in another (Planalp and Knie, 2002). For instance, smiling may facilitate laughter within one person and interpersonally a scream may facilitate another person’s desire to crouch in a position indicative of fear.

Thus, it has been argued that i) emotions are inherently social, ii) body movements and speech are interconnected and iii) consideration of the natural environment is important for perception. Therefore, based on the foregoing discussions, and in an attempt to take a more naturalistic approach in examining the exquisite ability to perceive biological motion, point-light displays were made that encaptured the behaviour of two actors interlocuting.

Interpersonal actions (where 2 actors interacted) have been used in point-light research once before (Dittrich, 1993). The actions studied were grouped into 3 categories, only one of which was interpersonal; the others were locomotory (e.g. walking) and instrumental (e.g. hammering). Observers were required to name the behaviour shown in each display and for the social actions recognition rates ranged from 66% for threatening behaviour to 95% for boxing. This seminal study was the first to use interpersonal body movements in point-light displays. However, there were only two examples used that can be considered as naturalistic behaviours.
(greeting and threatening). Furthermore, the actors in this study mimed the interpersonal actions which is not how greeting and threatening behaviour normally occurs.

It is argued that the interpersonal displays of affect used for the studies reported in this thesis are more natural than those used in Dittrich’s (1993) study as mime is not a normal form of interpersonal behaviour. Furthermore, in the present research all of the actions studied were produced in social contexts.

Chapter 2 describes the methods used to generate the stimuli and also details the experimental methods which were used to measure observers’ judgements. In Chapter 3, the emphasis of the research is to take some rudimentary first steps into exploring how interpersonal displays of affect are perceived. This new approach to investigating the perception of affect is then fully exploited in Chapter 4 to consider the importance of the social interaction per se. The aim of the work described in this chapter is to take a first step at examining certain aspects of communication with regard to recognition from biological motion. In the final two chapters several studies are described which were conducted in order to i) gain some more insight into what the specific cues that specify affect might be, ii) investigate further the way the stimuli might be processed and iii) consider how the skill to recognise affect from biological motion develops in childhood.
1.5. Summary

Human movement is an important feature of the world and vision allows us to pick up the movements of other conspecifics to make sense of our social environment. Movement of the body is a foundational characteristic of the communication system, yet this channel of expression is seldom investigated. The point-light method (Johansson, 1973) is an invaluable tool for investigating human movement in isolation from pictorial cues.

Emotions are also a fundamental aspect of being human. Several studies in the past have investigated the perception of emotion from body movements (by employing the point-light technique) but they have not considered the natural environment as a starting place for their investigations. This has potentially important consequences for the validity of findings in these studies. This thesis investigates a range of interdependent themes concerned with the perception of affect in natural communication behaviours from biological motion. Fundamentally, how the perception of emotion is affected by the social context is a primary area of research. The mechanisms responsible for perception of affect from biological motion also provide an avenue of enquiry. Finally, the way in which these skills develop in childhood is examined. It is hoped that this research will also further our understanding about what is known about the topic of emotion.

These questions are addressed in four empirical chapters. The work in the first empirical chapter (Chapter 3) investigates how well affect is recognised in point-light displays of communicating actors. Effects of display manipulation in the form of inverting and rotating (by 90°) the displays in the image plane are also investigated. The work in Chapter 4 addresses the importance of considering the social context in event perception. Specifically, it is asked how important the communicative context is for the perception of affect. Further, it is discussed what other aspects of the social exchange can be determined from biological motion independent of emotion, i.e. information about interlocutor identification. In Chapter
5, the mechanisms that are responsible for the perception of emotion (from biological motion) are discussed. The first study in this chapter explores the effect of reversing the motion in the display. The second study explores systematically the effects of reducing the information in the displays by removing specific elements of the body. Finally, in Chapter 6, I discuss how the mechanisms may develop during childhood.
Chapter 2

Methods

2.1. Overview

In all of the studies reported in this thesis, performance for perceiving emotions was assessed by analysing judgements made by observers when they were shown point-light displays of actors expressing emotions in a communicatory context. This chapter describes the methods used to generate the stimuli, including the rationale for the script used and how actors performed it. This chapter also details the experimental methods, which were used to measure observers' judgements.

2.2. Actors

Actors are able to display different emotions with variations in skill. For instance, one actor may be able to encode joy with a high level of accuracy but not sadness and for another actor the opposite might be true (Wallbot and Scherer, 1986). In many studies only one or a very few actors are used (e.g. Dittrich et al., 1996) and this limits acting styles. Thus, to ensure variation in acting styles, ten actors (5 male and 5 female) who were recent graduates from The Guildford School of Acting were employed to create the stimuli. All of the actors were aged between 20 and 30 years. Each actor was sent a script to learn 2 weeks before recording. Whilst it is true that acting variations exist at an intrapersonal level (Wallbot and Scherer, 1986) it is reasonable to assume that styles vary at an interpersonal level too. Therefore, to ensure interpersonal variations, the actors were informed that they would be acting together in female and male pairs and that every male actor would work with every female actor. They were told that their task would be to recite the script whilst portraying 6 different emotions: hot anger, joy, sadness, romantic love, fear and disgust.
2.3. Script

A basic script was chosen for the recordings rather than a scenario approach (where actors are provided with a short scenario, but no script, and are instructed to envisage the scenario): the use of a script ensured that a standard dialogue was used for all recordings. Wallbot and Scherer (1986) have argued that the scenario approach allows actors to produce more realistic expressions of affect compared to monologues, which are more likely to produce a declamatory style. However, the use of dyadic interlocution here was a method used to overcome such shortcomings as declamation. Wallbot and Scherer (1986) argued that the use of scenarios allows for between actor consistency in terms of emotions expressed; however, this is not necessarily the case as it depends on how a scenario is interpreted by an actor and also on an actor’s skill in the interpretation of it. Offering no scenario here allowed each actor free inspiration for the production of affect. However, actors were instructed not to touch their acting partner or express themselves by using overt symbolic gestures such as shaking the fist for anger or reaching towards each other with open arms for love.

Importantly, the script ensured that equal turn taking was maintained and that the recordings were of a similar length. The script consisted of 8 lines (4 lines for each actor) with approximately 7 words on each line. The actors were asked to treat the script as emotionally neutral. Here follows the script:

Actor A: It's a long way from Mordovia.
Actor B: I didn’t expect it would be so far.
Actor A: But there’s a trip to Moscow thrown in.
Actor B: No-one told me about the extra trip.
Actor A: Well the embassy arranged it all.
Actor B: I thought that our comrades arranged it.
Actor A: If the train could stop especially here.
Actor B: Such a thing could be arranged.
2.4. Movement Recording Equipment

Actors' movements were recorded using a MacReflex motion analysis system (Qualisys AB). This system comprised of a high-resolution camera, sensitive only to infrared light. The camera was connected to video processing hardware, which calculated the 2-D co-ordinates of lightweight spherical markers, 50 mm in diameter, covered with silver reflective tape (3M Scotchlite), which were attached to the actors' bodies. The equipment was connected to an Apple Macintosh computer, equipped with MacReflex software. The computer was connected to a 15 inch monitor that was able to synchronize on 60 Hz: the sampling rate of the camera (one 'frame' every 16.66ms). The camera was able to distinguish the light-reflective markers from the background and reduce the output of the video signal to contain only information about the markers by using 3 mechanisms; these were: an electronic shutter with a very high shutter speed of 1/4000 s (0.25ms) that reduced the effect of ambient light; pulses of infrared light that were synchronised to the sampling rate of the system; and automatic gain control so that, provided there were no 'false' reflections from other objects, the markers were the only detectable infrared light sources in the scene. A videoprocessor connected to the camera calculated the 2-D coordinates of each marker within the camera's field of view by scanning the image pixel by pixel and calculating the centroid of each marker location. As the markers used were spheres their aspect did not change with different orientations; therefore the centroid was a reliable indicator of the marker's position. Several hours before recording, the equipment was switched on and the scene was checked via the monitor of the host computer to ensure there were no other light sources that would provide 'false' reflections.

2.5. Marker and actor positions

The markers were attached to bands of elasticated tape that fitted securely around the actors' ankles, knees, hips, shoulders, elbows and wrists, and one to the centre of the forehead. In particular, the markers were placed over the iliac fossa area of hip and
the humeral joint capsule in the shoulder. The markers were attached to the front of
the actors’ bodies except for those on the elbows; these were positioned midway
between frontal and lateral positions to give them the best possible chance of being
in view of the camera at all times as the actors moved their arms.

The camera was placed at a distance of 4.5 metres from the target where the actors’
heights filled approximately 70-80% of a monitor screen. To keep the actors in view
at all times they were instructed to work within a fixed area (1.3 m wide and 0.8m
deep) which was marked on the floor. Each actor stood facing at approximately 90°
to the other and at approximately 45° to the camera. No sound recording was made.

2.6. Recordings

A room measuring 5.5 metres by 8 metres was used as a temporary recording studio
and another room, adjacent to the first, was used as a waiting room for the actors.
The recording took place on a bright afternoon and so the Venetian blinds covering
the windows in the recording room were tilted to let in a small amount of light. This
measure ensured that no sunlight could cause ‘false’ reflections for the camera to
pick up. No artificial light was used. The author and 2 assistants were present for the
recording session. One assistant was responsible for fetching the appropriate actors
for each recording (every actor was assigned a number badge and everyone had a
copy of the schedule) and the other assistant was accountable for ensuring that the
markers were attached in the appropriate positions on every actor. The author
instructed the actors on which emotion was required and also operated the filming
equipment. On each recording, the actors were requested not to rush, and to give an
indication when they ‘felt’ the emotion. They were then immediately asked to act
that emotion. The author who could also see the actors — (labelled experimenter in
Figure 2.1.) was positioned to be able to see both the recordings as they appeared on
the monitor where the point-lights were displayed as coloured crosses on a white
background. Figure 2.1 shows a plan view of the recording layout.
Every male actor produced 4 different categories of affect with every female actor (time restrictions disallowed the opportunity for every category to be produced within each acting pair; enough samples were produced by using the method chosen — see below). In every recording both actors portrayed the required emotions simultaneously. This arrangement meant that at least 3 clips of each emotion per actor were produced. This made a corpus of ~100 recordings (5 female actors x 5 male actors x 4 emotions), with at least 16 recordings of each emotional category.
Each pair of actors was allowed to repeat a recording if they thought that the first was not good enough; if this occurred then the first recording was discarded. On one occasion the actors were asked to repeat a recording as one actor turned too much — resulting in the camera picking up too few light points. No more than two recordings of the same stimulus were required. It took approximately 4 hours to produce the corpus of vignettes.

2.7. Dependent measures

For all of the experiments reported in this thesis (with the exception of the qualitative study reported in Chapter 3 and the child study reported in Chapter 6) all of the observers' judgements for perceived emotion were measured by the method described in the following paragraph.

In every experiment the stimuli were presented to observers on either a 33cm x 25.5cm Apple monitor set at a refresh rate of 60 Hz with a 1024 x 768 resolution or occasionally (for Experiments 1b and 7) on an Apple G3 iBook laptop computer with a 14-inch LCD display and a 1024 x 768 resolution. A laboratory was used for all of the experiments using the desktop computer (as most of the experiments were conducted in the laboratory, in the empirical chapters that follow, it is only reported when the laptop computer was used). An Anglepoise lamp lit the laboratory which was fitted with a 60 watt bulb. The lamp faced away from the experimental apparatus towards the wall. Otherwise the room was dark. The point light displays were shown in green on a black background. The luminance of the dots was 44.1 cd/m² when measured in the laboratory. The monitor screen subtended 9° of visual angle vertically and 11.7° horizontally when shown on the desktop computer. The displays were presented within a boxed area of this that was the same colour and luminance as the light points and subtended approximately 8.4° of visual angle vertically and 6° horizontally. When presented on the Apple iBook the monitor subtended 8.9° of visual angle vertically and 11.3° horizontally and the boxed area subtended 8.5° of visual angle vertically and 6.1° horizontally. A static frame from one of the displays can be seen in Figure 2.2. Subjects were seated in front of the displays at a distance
of approximately 80 cm when shown on the larger monitor and were slightly less (70 cm) for the iBook presentations. No headrest was used for any of the experiments unless stated.

Figure 2.2. Static frame of a point-light display showing the two actors where each actor is represented by 13 points of light. In this frame the actor on the left has his left arm occluded by the rest of his body.
Subjects were asked to attend to the monitor carefully during each trial and to judge how much of each emotion they thought they perceived in the stimulus. Following the presentation of each experimental stimulus six (sometimes less, depending on which experiment was being conducted) horizontal scales appeared (arranged one above another) on the screen— one marked for each emotion with the limits labelled from ‘very little’ (on the left) to ‘a lot’ (on the right). Subjects used the computer mouse to move a slider along each scale to provide the rating. A picture of the scale before being rated and an example of how observers might have rated a movie (showing 5 out of 6 ratings completed) can be seen in Figure 2.3 a) and b) respectively. The point where each subject clicked was converted to a corresponding number between 0-100 by the computer program. Observers were asked to use the full range of the scale. For each subject the order in which the emotions appeared on the scale for rating was altered so that the emotion appearing at the top for one subject was moved to the bottom for the next subject and so on. The subjects were obliged to work from the top scale to the bottom on each trial. After the last scale had been rated on one trial the next trial started automatically.
Figure 2.3. An image of the dependent measure used for most of the experiments reported in this thesis (a), and (b), an example of how the image might have looked after 5 out of 6 of the scales had been rated.
2.8. Subjects

Almost without exception, all of the adult subjects were either postgraduate or undergraduate psychology students at the University of Surrey. Undergraduates received a course credit for their participation. Very occasionally, other adults such as the staff from the University’s Psychology Department took part. In the developmental study (Chapter 6) all of the children were recruited from one primary school situated in the South of England. Informed consent from the parents or guardians of the children was obtained for every child. The children were tested at school, during school hours, and were taken individually from their classrooms to a quiet area of the school where the children were tested.

All adults and children that took part as subjects had normal, or corrected-to-normal vision. They were all naïve as to the purpose of the experiment in which they took part and no subject had ever previously encountered the experimental stimuli.
Chapter 3

Effects of upright and rotated displays

3.1. Introduction

Wallbot (1998) and Montepare et al. (1999) have investigated the specific link between the production and the perception of affect from body movements. These authors have suggested that there are qualitative patterns of movement to distinguish between emotions. However, in these studies full-light (video) displays were used as stimuli and as such they contain more information than solely cues from movement. For instance, in a full-light display there are fundamental spatial properties of the visual world such as edges, shapes, colour and texture, which are all likely to contribute to vision (Gibson, 1950). As discussed in Chapter 1 an effective method to study how emotions are perceived from body movements is to use Johansson (1973) point-light displays as they remove most of the spatial information that would normally be available in a full-light display: information from motion remains. Although a small but informative number of studies have been conducted to investigate affect by using this method (e.g. Dittrich et al., 1996; Atkinson et al., 2004), so far the social context in which emotions are usually expressed has not been considered as a method by which to explore perception from biological motion of affect.

To address this, the central question explored in this first empirical chapter was whether emotion could be perceived from the biological motion of two actors engaged in dialogue.

A supplementary enquiry was to ask how the percept was disrupted through rotating the displays on the image plane by 180°, so inverted; and by 90° (clockwise or anticlockwise), thus appearing with the main axis of the body across the width of the image plane. This method of image rotation is commonly used in biological motion
research; inverted displays are most often used (e.g. Sumi, 1984; Dittrich et al., 1996) and, less frequently, other orientations are also shown (e.g. Pavlova and Sokolov, 2000). The popularity of this approach is based on the logic that the pattern of the kinematics within the stimulus remains constant in spite of image rotation. If perception is disrupted by rotation then it cannot depend solely on low level computational processing accounts that have been put forward such as structure from the relative motions of the display’s elements defined by local rigidity (Webb and Aggarwal, 1982). Therefore, the first empirical chapter addressed the ability of observers to perceive emotion from biological motion expressed in the bodies of two people engaged in discourse but also took a rudimentary first step at considering how the visual system generates percepts by examining the effects of orientation manipulation.

In addition to the foregoing enquiries, included in the chapter are several other experiments that controlled for possible timing and learning artefacts that may have existed in the presentations shown in the first experiment and the first control experiment. As a measure of validation for the quantitative data collection method that is used throughout this chapter, and indeed for most of the other experiments reported in this thesis, a qualitative study is also reported that allowed subjects to express freely what they could see in the displays. Finally, this chapter also investigates the issue of whether there were sex differences for the perception of affect.
3.2. Experiment 1

Point-light displays contain few visual cues, however, that they are impoverished is not true: They are information rich. For example, Dittrich, et al. (1996) showed that surprise, fear, anger, grief and joy could be identified above chance when portrayed by dancers in point-light displays. However, some emotions were identified with more success than others. Portrayals of grief and joy were better identified than portrayals of fear and anger. Portrayed disgust received a below chance score and it was confused with the non-portrayals of fear and grief. More recently, other movements have also been investigated. Pollick, Paterson, Bruderlin and Sanford (2001) used point-light displays of knocking and drinking arm movements (i.e. as if knocking on a door angrily), to show that subjects can perceive a range of internal states (e.g. fear, anger, tiredness) from these actions. Claiming a more naturalistic approach, Atkinson et al. (2004) used single actors to portray the bodily expressions of anger, joy, sadness, fear and disgust to show that observers could recognise all of these emotions well above chance.

Interpersonal actions have been used in point-light research once before (Dittrich, 1993). The actions studied were grouped into 3 categories, only one of which was interpersonal, with four examples in each category. The groups were: locomotory (e.g. walking); instrumental (e.g. hammering) and, of particular relevance here, social actions that portrayed dancing, boxing, greeting and threatening behaviours. Observers were required to name the behaviour shown in each display and recognition rates ranged between 42% for stirring to 97% for going upstairs. For the social actions performances ranged from 66% for threatening behaviour to 95% for boxing. This seminal study was the first to use interpersonal body movements in point-light displays. However, there were only two examples used that can be considered as naturalistic behaviours (greeting and threatening). It is somewhat surprising that no study has used interpersonal communication to capture emotional expressions in body motion given that the function of emotions must be, at least in
part, to send information to others so that in return the observer may induce both appropriate cognitive attitudes and reciprocal emotion (Scherer, 1993).

Numerous studies have shown that display inversion impedes the perception of biological motion. Infants cannot discriminate a point-light walker from similar displays when displays are inverted (Bertenthal, Proffitt and Cutting, 1984) and, for adults, image inversion impedes recognition of point-light walkers quite substantially (e.g Pavlova and Sokolov, 2000; Shipley, 2003).

Point-light walkers aside, the recognition of types of human actions and emotions is also disrupted by inversion although the extent to which an action or emotion is disrupted varies (Dittrich, 1993; Dittrich et al., 1996). For locomotory actions such as jumping, Dittrich (1993) found that recognition was reduced from 87% correct for upright displays to 45% correct for inverted ones. Point-light displays of social actions varied in the amount of disruption caused by display inversion. For example, for dancing and greeting couples, recognition was reduced from approximately 80% for uprights to 60% correct whereas for threatening behaviour recognition was reduced by half (Dittrich, 1993). However, the biggest disparity was for boxing behaviour as when upright these displays were recognised with 95% accuracy and yet when inverted recognition diminished to just 5%.

For the case of emotions, Dittrich et al. (1996) found that in contrast to upright point-light displays where, with the exception of disgust, all portrayed emotions received higher mean scores for recognition than the non-portrayed ones, inversions were more confused with other non-portrayed categories: once again portrayed disgust was confused for (fear and grief) non-portrayals but unlike the results for uprights, portrayed joy was confused with non-portrayed surprise, and portrayed fear was confused with non-portrayed grief. Unfortunately, Atkinson et al. (2004) did not use inverted displays for comparisons to be made.

When inverting point-light stimuli on the image plane, intrastimulus kinematics such as velocity of the dots and relative physical positioning remain constant. Therefore, that image inversion causes such a problem for veridical identification of the stimuli
is of interest as this means that simple physical information is not sufficient for recognition. Rather, the visual system must implement additional constraints for veridical perception of the stimuli. Pavlova and Sokolov (2000) have argued for two such constraints: ecological ones derived from the invariants of the environment such as gravity, and knowledge-based constraints that consist of top-down knowledge about the outside world. There is accumulating evidence that ecological constraints are of importance as observers shown 10 second displays of either upright, oriented 45° or 90°, or inverted point-light walkers (primes) followed by camouflaged displays of the same, were influenced strongly by the primes of upright walkers when detecting the camouflaged examples. The 90° and 180° primes had no effect on subsequent identification of the camouflaged examples at all. These results suggest that prior (top-down) knowledge of the displays was not used for recognition (Pavlova and Sokolov, 2000). This finding is corroborated with evidence that perception remains impervious to prior knowledge about inverted displays but that the addition of either a solid or moving line implying the existence of either a ground plane or some other frame of reference improves recognition of the same displays to above chance (Pavlova and Sokolov, 2003). Furthermore, Shipley’s (2003) novel experiment where a point-light walker walked on his feet in one scene and his hands in another showed that subjects were better able to identify the walker when shown in an upright orientation over an inverted one. This was despite the fact that observers are not familiar with people walking on their hands and that for the inverted hand walker, limbs would be in their normal positions. This result suggests that perceivers use perceptual information about the physical world such as the force of gravity more than knowledge-based information and so stresses that features in the natural world are fundamental (Pavlova and Sokolov, 2003).

The present study draws on the design of Dittrich et al. (1996) to examine the perception of emotion from point-light displays and the effects of inversion. However, rather than the use of single dancers to express bodily affect, interpersonal behaviour as expressed by two actors engaged in dialogue was used (see Chapter 1). Based on the results of Dittrich et al. (1996) and Atkinson et al. (2004) it was expected that performance would be good for the upright portrayals. For inverted
displays performance was expected to be impaired to some degree but that recognition would still be above chance was thought likely.

3.3. Experiment 1: Method

3.3.1. Subjects

Eight postgraduate psychology students and 2 members of the psychology department's staff (7 men and 3 women; mean age 29.6 years; range 18 - 42 years) volunteered to serve as subjects.

3.3.2. Apparatus.

A headrest was used, which was set at a viewing distance of 80 cm from the screen.

3.3.3. Design and procedure

Six different examples of each of the six emotions were used. The criteria for choosing the vignettes for Experiment 1 were to use as many combinations of different actors as possible and to use the vignettes that had the least number of occluded light points (caused by actors turning their bodies too far away from the camera). The examples were presented either upright or upside down which gave a total of 72 trials (6 emotions x 6 samples x 2 orientations). Each stimulus was shown for its full duration (17-30) seconds. Subjects were told that they would see some points of light on the screen that represented two people in a social context both acting out one of six possible emotions. The subjects were not informed that some displays would be presented upside down. All of the stimuli were presented in one randomised block for each subject and the experiment lasted for approximately 50 minutes.
3.4. Experiment 1: Results and discussion

Most displays received non-zero ratings on all scales due to the fact that clicking towards the far left of each scale gave a non-zero number corresponding to the writing shown on the monitor saying 'very little' (of that emotion present). Clicking at the very far left end of the scale gave a zero rating.

3.4.1. Upright displays

For each portrayed emotion the means of the five non-portrayed emotion ratings were pooled so that comparisons could be made between just one 'incorrect' (non-portrayed) score and one correct (portrayed) rating. Figure 3.1.a shows that the portrayed emotion received the higher rating in every case. For example, when anger was portrayed in the display it received mean ratings in excess of 70 whereas the other five non-portrayed emotions received an overall mean of 19.
Figure 3.1. Observers' mean ratings for the a) upright and b) inverted presentations of the emotion response categories of the portrayed and non-portrayed emotions. The mean ratings for the non-portrayed emotions were derived from the mean of the combined scores from the five non-portrayals in each case. Error bars show standard errors.
Table 3.1. Observers’ mean ratings for each emotion response category when a) upright and b) inverted point-light stimuli were presented. Figures in bold type denote the mean ratings for the portrayed emotions. Standard deviations are shown in parentheses.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Response</th>
<th>Anger</th>
<th>Joy</th>
<th>Sadness</th>
<th>Love</th>
<th>Fear</th>
<th>Disgust</th>
</tr>
</thead>
<tbody>
<tr>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upright presentation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td>70 (7.5)</td>
<td>7 (7.3)</td>
<td>23 (20.7)</td>
<td>12 (11.1)</td>
<td>24 (18.8)</td>
<td>31 (16.2)</td>
<td></td>
</tr>
<tr>
<td>Joy</td>
<td>26 (13.8)</td>
<td>53 (17)</td>
<td>8 (6.1)</td>
<td>19 (13.8)</td>
<td>15 (11.1)</td>
<td>14 (12.1)</td>
<td></td>
</tr>
<tr>
<td>Sadness</td>
<td>16 (12.5)</td>
<td>6 (4.1)</td>
<td>59 (17.2)</td>
<td>20 (16.1)</td>
<td>17 (11.2)</td>
<td>20 (9.3)</td>
<td></td>
</tr>
<tr>
<td>Love</td>
<td>21 (11.8)</td>
<td>10 (5.9)</td>
<td>24 (11.2)</td>
<td>47 (19.2)</td>
<td>20 (9.9)</td>
<td>15 (10)</td>
<td></td>
</tr>
<tr>
<td>Fear</td>
<td>18 (18.2)</td>
<td>15 (12.2)</td>
<td>18 (13.6)</td>
<td>16 (11.4)</td>
<td>59 (14.7)</td>
<td>19 (9.9)</td>
<td></td>
</tr>
<tr>
<td>Disgust</td>
<td>24 (11.7)</td>
<td>9 (7)</td>
<td>37 (14.2)</td>
<td>14 (10)</td>
<td>22 (11.4)</td>
<td>37 (11.1)</td>
<td></td>
</tr>
<tr>
<td>b)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inverted presentation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td>57 (7.9)</td>
<td>11 (5.1)</td>
<td>15 (8.9)</td>
<td>14 (14)</td>
<td>17 (11.9)</td>
<td>23 (16)</td>
<td></td>
</tr>
<tr>
<td>Joy</td>
<td>25 (16.7)</td>
<td>46 (23.2)</td>
<td>6 (4.7)</td>
<td>20 (11.6)</td>
<td>11 (6.8)</td>
<td>13 (7.5)</td>
<td></td>
</tr>
<tr>
<td>Sadness</td>
<td>17 (8.3)</td>
<td>11 (6.1)</td>
<td>32 (11.3)</td>
<td>24 (9.5)</td>
<td>16 (9.9)</td>
<td>19 (7.7)</td>
<td></td>
</tr>
<tr>
<td>Love</td>
<td>17 (12.9)</td>
<td>13 (6.6)</td>
<td>26 (15.9)</td>
<td>44 (15.8)</td>
<td>19 (8.7)</td>
<td>17 (9.2)</td>
<td></td>
</tr>
<tr>
<td>Fear</td>
<td>17 (11.3)</td>
<td>24 (8.7)</td>
<td>16 (4.5)</td>
<td>20 (8.1)</td>
<td>20 (11.7)</td>
<td>15 (5.8)</td>
<td></td>
</tr>
<tr>
<td>Disgust</td>
<td>19 (6.6)</td>
<td>12 (7.3)</td>
<td>21 (5.4)</td>
<td>17 (7.8)</td>
<td>19 (8.4)</td>
<td>24 (11.6)</td>
<td></td>
</tr>
</tbody>
</table>

The complete set of ratings is shown in Table 3.1. For each portrayed emotion, a one-way analysis of variance (ANOVA) was conducted between the ratings for that emotion and for the five non-portrayed emotions (i.e. along each row). All

1 An important assumption for within-subjects ANOVA is that sphericity should not be violated as, if it is, the possibility of a type I error may be greatly inflated (for a discussion see Howell, 1997, chapter 14). Where the data have failed the sphericity test in the present experiments (as indicated by the symbol *) the more conservative Greenhouse – Geisser epsilon test has been used, which reduces the degrees of freedom of the numerator and denominator.
ANOVA showed significant main effects (portrayed anger, $F_{5,45} = 24.9, p < 0.001$; joy, $F_{1,87,16.8^*} = 14.7, p < 0.001$; sadness, $F_{5,45} = 22.2, p < 0.001$; love, $F_{5,45} = 12.03, p < 0.01$; fear, $F_{2,6,22.7^*} = 18.8, p < 0.001$; disgust, $F_{1,9,17.9^*} = 10.12, p = 0.001$). Tukey’s post hoc tests confirmed that for each stimulus the portrayed emotion was rated higher ($p < 0.001$) than each of the 5 non-portrayed emotions considered individually with the exception of portrayed disgust. Ratings for this portrayed emotion did not differ from those for (non-portrayed) sadness, anger and fear although they were significantly larger than the ratings for (non-portrayed) joy and love ($p < 0.001$).

These results show clearly that the emotions of anger, fear, joy, sadness and love can be identified in normally oriented point-light displays. This suggests that the spatio-temporal information inherent in human movement is sufficient to specify certain emotions. The exception was for the display of disgust, which was confused with sadness, anger and fear. That some emotions were better identified than others corroborates findings in the literature (e.g. Dittrich et al., 1996; Atkinson et al., 2004) although the order in which emotions were most easily identified differs between studies.

### 3.4.2. Inverted displays

Figure 3.1.b shows the mean ratings of the portrayed emotions as well as the combined mean scores for the non-portrayed emotions when the displays were presented upside down; Table 3.1b presents the complete set of ratings.

A two-way ANOVA between display type and emotion with two levels of display type (the bold figures in Tables 3.1a and 3.1b) showed that ratings for inversion were significantly reduced ($F_{1,9} = 32.2, p < 0.001$) relative to the upright stimuli. A significant interaction between emotion and orientation ($F_{2.3,20.5^*} = 5.5, p < 0.01$) showed that inversion affected the ratings of some portrayed emotions more than others. Post hoc Tukey’s tests following the ANOVA revealed that the ratings for
anger (p < 0.005), sadness (p < 0.005), fear (p < 0.001) and disgust (p < 0.05) were significantly reduced by inversion. Joy and love were unaffected by inversion although these results may have been due to a lack of power given the small number of subjects (n=10) that took part in the experiment.

For the inverted stimuli one-way ANOVAs that compared the ratings for the portrayed and all of the five non-portrayed emotions showed that there was a significant main effect for portrayed anger, sadness, love and joy but not for fear or disgust (portrayed anger, $F_{5,45} = 27.74$, p < 0.001; joy, $F_{1,4,14,7} = 11.13$, p = 0.002; sadness, $F_{5,45} = 6.49$, p < 0.05; love, $F_{2.1, 18.1} = 8.87$, p = 0.002; fear $F_{5,45} = 1.71$, n.s and disgust $F_{2.4, 22} = 3.02$, n/s). Post hoc Tukey’s tests confirmed that for portrayals of anger, joy or love, the portrayed emotion was rated significantly higher than each of the 5 non-portrayed emotions individually (p < 0.01). This was also true for ratings of portrayed sadness (except that the rated degree of sadness did not differ significantly from that of love, although this result may also have been due to a lack of power). Ratings for disgust were confused with each of the 5 non-portrayed emotions except for joy. Fear was indistinguishable from all other (non-portrayed) emotions when inverted.

Inversion reduced the salience of the displayed emotions to varying extents although overall performance was still good for most emotions. This result is in line with previous findings for inverted biological motion displays of dancers emoting (Dittrich et al., 1996). However, that the identification of fear was completely destroyed when displays were inverted was an unexpected result.

3.5. Experiment 1a

The stimuli for both upright and inverted presentations used in Experiment 1 were shown in one randomised block. This meant that there was no control over the orientation order of the stimuli, which gave rise to the possibility that if subjects saw an upright display before seeing an inverted version of the same display, they may have been able to remember the features from the first showing for subsequent
identification. Furthermore, in Experiment 1 each movie varied in presentation time based on the amount of time actors used for the production of each vignette; there were also differences in mean presentation times for each emotion category. The fastest mean time was for the anger vignettes (17.6 seconds), followed by joy (17.9), fear (19.4), disgust (21.6), then love (22.7) and the slowest time was for sadness (24.8). Therefore, a second experiment was devised that had two main aims: the first was to control for any learning that might occur on presentation of the upright stimuli, the second was to control for the possibility that subjects were discriminating the stimuli by stimulus length. Importantly, this experiment also tested the replicability of the results found in Experiment 1.

3.6. Experiment 1a: Method

3.6.1. Subjects

Fifteen undergraduates and 3 postgraduates participated in the experiment and undergraduates received course credits for their participation (3 males and 17 females; mean age = 21.5 years; age range 19-33 years).

3.6.2. Apparatus, design, stimuli and procedure

The experiment employed a mixed factorial design. As in Experiment 1, there were six different examples of six emotions used. These were presented both upright and upside down which gave a total of 72 trials (6 emotions x 6 samples x 2 orientations). There were 2 blocks of trials containing either the upright or inverted stimuli. The order in which they were completed was counterbalanced so that half of the observers saw the upright stimuli first and half saw the inverted ones. This was a between-subjects factor. Within each block, trials were randomly ordered. Each block lasted for approximately 25 minutes.

The same stimuli were used as in Experiment 1 except that each movie was shortened so that only the first 10 seconds (600 frames) were shown.
3.7 Experiment 1a: Results and discussion

3.7.1. Results

The ratings were recorded and the non-portrayed scores were pooled, in the same way as in Experiment 1. Thus, comparisons were made with just one correct (portrayed) and one 'incorrect' (non-portrayed) rating. Figure 3.2 shows these results. Figure 3.2.a shows that the portrayed emotion for the upright presentations received the higher mean ratings in every case. Figure 3.2.b shows that ratings for the inverted portrayals received the higher ratings in each case except for fear, which received similar ratings for incorrect and correct judgements irrespective of which blocked viewing condition was shown first.
Figure 3.2. Observers’ mean ratings as a function of viewing block order for the a) upright and b) inverted presentations of the emotion response categories of the portrayed and non-portrayed emotions. The mean ratings for the non-portrayed emotions were derived from the mean of the combined scores from the five non-portrayals in each case. Portrayed UF and portrayed IF denote whether upright (UF) or inverted (IF) stimuli were shown first to observers. Error bars show standard errors.
Table 3.2. Observers’ mean ratings for each emotion response category when stimuli were blocked and counterbalanced by orientation. The ratings have been collapsed by presentation order and show a) uprights and b) inversions. Figures in bold type denote the mean ratings for the portrayed emotions. Standard deviations are shown in parentheses.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Response</th>
<th>Anger</th>
<th>Joy</th>
<th>Sadness</th>
<th>Love</th>
<th>Fear</th>
<th>Disgust</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Upright stimuli</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td></td>
<td>70 (14)</td>
<td>7 (6.9)</td>
<td>24 (25)</td>
<td>7 (7.1)</td>
<td>19 (15.4)</td>
<td>39 (25.6)</td>
</tr>
<tr>
<td>Joy</td>
<td></td>
<td>28 (19.6)</td>
<td>37 (16.5)</td>
<td>14 (11.9)</td>
<td>18 (15.6)</td>
<td>14 (8.5)</td>
<td>22 (20)</td>
</tr>
<tr>
<td>Sadness</td>
<td></td>
<td>15 (13.6)</td>
<td>9 (8.3)</td>
<td>49 (20.8)</td>
<td>17 (12.6)</td>
<td>19 (12.1)</td>
<td>20 (15.5)</td>
</tr>
<tr>
<td>Love</td>
<td></td>
<td>16 (11.8)</td>
<td>13 (14.6)</td>
<td>19 (12.4)</td>
<td>41 (14.3)</td>
<td>14 (11.1)</td>
<td>17 (15)</td>
</tr>
<tr>
<td>Fear</td>
<td></td>
<td>18 (15.3)</td>
<td>12 (9)</td>
<td>24 (14.4)</td>
<td>13 (10.9)</td>
<td>44 (19.3)</td>
<td>20 (16.1)</td>
</tr>
<tr>
<td>Disgust</td>
<td></td>
<td>21 (15.9)</td>
<td>10 (6.9)</td>
<td>36 (19.9)</td>
<td>14 (10)</td>
<td>17 (16.7)</td>
<td>32 (15.5)</td>
</tr>
<tr>
<td>b) Inverted stimuli</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td></td>
<td>61 (16.6)</td>
<td>9 (7.3)</td>
<td>25 (21)</td>
<td>8 (6.4)</td>
<td>18 (14.4)</td>
<td>38 (22.2)</td>
</tr>
<tr>
<td>Joy</td>
<td></td>
<td>30 (19.9)</td>
<td>36 (18.1)</td>
<td>14 (13.4)</td>
<td>19 (14.5)</td>
<td>13 (7.4)</td>
<td>19 (16.2)</td>
</tr>
<tr>
<td>Sadness</td>
<td></td>
<td>16 (15.7)</td>
<td>9 (6.3)</td>
<td>38 (18.9)</td>
<td>19 (10.5)</td>
<td>22 (17.4)</td>
<td>22 (17.7)</td>
</tr>
<tr>
<td>Love</td>
<td></td>
<td>13 (14.8)</td>
<td>13 (12.1)</td>
<td>22 (17.4)</td>
<td>36 (15.2)</td>
<td>13 (8.8)</td>
<td>16 (17.5)</td>
</tr>
<tr>
<td>Fear</td>
<td></td>
<td>20 (17.6)</td>
<td>17 (10.4)</td>
<td>26 (18.9)</td>
<td>20 (14.1)</td>
<td>24 (14.4)</td>
<td>22 (18)</td>
</tr>
<tr>
<td>Disgust</td>
<td></td>
<td>21 (16.9)</td>
<td>13 (7.6)</td>
<td>33 (19.6)</td>
<td>16 (11.5)</td>
<td>19 (13.3)</td>
<td>31 (21.4)</td>
</tr>
</tbody>
</table>

A mixed design ANOVA (the bars for the portrayed emotions only in Figure 3.2. a) and b) with 2 within-subjects factors: emotion (6 levels) and orientation (2 levels), and one between-subject’s factor (blocked condition order of either uprights or inversions first), confirmed that inverted ratings were significantly reduced compared to upright stimuli ($F_{1,16} = 9.9, p < 0.001$). A significant interaction between emotion and orientation showed that some emotions were more affected by inversion than others ($F_{5, 80} = 4.6, p = 0.001$). However, there was no significant main effect for blocked orientation presentation order ($F_{1,16} = 0.95, p > 0.05$). This confirmed that it
made no difference whether observers viewed the upright or inverted blocks of stimuli first. Therefore, the results for orientation viewing order are considered no further here. To find out how recognition was affected by inverting the stimuli overall, the corresponding ratings from both blocking conditions (inverts first and upRights first) were collapsed together and a two-way within-subjects ANOVA was performed. This showed main effects for emotion \( (F_{5, 85} = 23.8, p < 0.001) \) and orientation \( (F_{1, 17} = 9.4, p < 0.01) \), which was qualified by an interaction between them \( (F_{2, 105} = 4.2, p < 0.05) \). Post hoc tests showed that fear was the only emotion with ratings significantly reduced by inversion \( (p < 0.005) \).

For each portrayed emotion, a one-way analysis of variance (ANOVA) was conducted between the ratings for that emotion and for the five non-portrayed emotions (i.e. along each row in Table 3.2). For all of the upright displays all ANOVAs showed significant main effects \( (\text{portrayed anger}, F_{2.34, 39.84} = 58.8, p < 0.001; \text{joy}, F_{2.18, 37.14} = 7.3, p < 0.005; \text{sadness}, F_{2.7, 46.36} = 27.4, p < 0.001; \text{love}, F_{2.67, 45.68} = 16.2, p < 0.001; \text{fear}, F_{1.92, 32.66} = 15.66, p < 0.001; \text{disgust}, F_{3.33, 56.67} = 15, p < 0.001) \). Tukey’s post hoc tests confirmed that for each stimulus the portrayed emotion was rated higher \( (p < 0.005) \) than each of the 5 non-portrayed emotions considered individually with the exception of portrayed joy and portrayed disgust. Joy portrayals were distinguished from the non-portrayals of sadness, love, fear and disgust \( (p < 0.05) \) but did not differ from the non-portrayals of anger. Disgust portrayals were rated differently to all non-portrayals \( (p < 0.005) \) except for sadness.

For the inverted stimuli ANOVAs that compared the ratings for the portrayed and all of the five non-portrayed emotions showed there was a significant main effect for all of the portrayed emotions except fear \( (\text{portrayed anger}, F_{2.72, 46.19} = 54.1, p < 0.001; \text{joy}, F_{1.89, 32.22} = 8.3, p = 0.001; \text{sadness}, F_{5.85} = 12.4, p < 0.001; \text{love}, F_{5.85} = 9, p < 0.001; \text{fear}, F_{2.77, 47} = 1.07, n/s \) and \( \text{disgust}, F_{2.69, 49.13} = 7.7, p < 0.001 \) \). Post hoc Tukey’s tests confirmed that for portrayals of anger, sadness or love, the portrayed emotion was rated significantly higher than each of the 5 non-portrayed emotions individually \( (p < 0.05) \). Portrayals of joy were rated significantly higher than for (non-portrayals) of sadness, love, fear and disgust \( (p < 0.05) \), but they did not differ significantly from non-portrayals of anger. Ratings for portrayed disgust were higher
than the non-portrayals of anger, joy, love and fear (p < 0.05) but were confused with sadness. As in Experiment 1, fear was indistinguishable from all other (non-portrayed) emotions when inverted.

3.7.2.Comparison between Experiments 1 and 1a

The ratings for the portrayed emotions in Experiment 1a, when presentations lasted for 10 seconds, were combined together irrespective of block order so that they could be compared with the results from Experiment 1 when presentations lasted for the full length of the vignette. Figures 3.3.a and 3.3.b present the ratings from both experiments for the portrayed emotions and the 5 combined non-portrayed emotions for the upright a) and inverted b) presentations.
Figure 3.3. Observers’ mean ratings as a function of experiment for a) upright and b) inverted presentations of the emotion response categories of the portrayed and non-portrayed emotions. The mean ratings for the non-portrayed emotions were derived from the mean of the combined scores from the five non-portrayals in each case.
For the upright presentations in Experiment 1, when the full-length vignettes were shown, ratings for most of the portrayed emotions appeared higher than for the 10-second presentations (Experiment 1a). For the portrayed, inverted presentations from both experiments some ratings were higher for Experiment 1, e.g. joy and love, and some were higher when the shorter movies were shown (Experiment 1a). The non-portrayed emotion ratings varied by a maximum of 3 between the two experiments in every case for each orientation except for the case of inverted displays of non-portrayed anger, where ratings were slightly lower in the first experiment. An omnibus ANOVA with 3 within-subject variables (emotion x error x orientation) and a between-subject variable (experiment) that compared the ratings made in Experiments 1 and 1a for only the non-portrayals showed significant main effects for both emotion $F_{5, 130} = 7.9, p < 0.001$ and the non-portrayed ratings $F_{4, 104} = 13.8, p < 0.001$. However, there was no effect for orientation $F_{1, 26} = 1, p > 0.05$ and importantly, no effect for experiment $F_{1, 26} = 0.3, p > 0.05$. This suggests that there were no differences between the experiments for the incorrect ratings made.

A mixed design ANOVA with 2 within-subjects factors (emotion (6 levels) and orientation (2 levels)), and one between-subject’s factor (experiment) — (the bold figures in Tables 3.1 and 3.2) was carried out on the data. There was a main effect for emotion ($F_{5, 130} = 26.3, p < 0.001$) and orientation ($F_{1, 26} = 42.8, p < 0.001$), qualified by a significant interaction ($F_{5, 130} = 8.5, p < 0.001$). However, no significant between-subject difference between Experiments 1 and 1a was found ($F_{1, 26} = 2.3, p > 0.05$).

### 3.7.3. Discussion

As in Experiment 1, perception of some emotions was affected by inversion. Portrayals of joy (when uprights were shown first) and love (when uprights were shown second) were confused with one or more non-portrayals here whereas in Experiment 1 these emotions were not affected by inversion. Once again, inversions of portrayed sadness were confused with love (only when uprights were shown first)
and portrayed disgust, irrespective of order of showing, was confused with at least two other non-portrayals. The most interesting finding for inversion displays, as in Experiment 1, was that portrayed fear was indistinguishable from all other emotions (except joy when inverts were shown first).

That there was no significant between-subjects difference between the shown order of orientations suggests that subjects did not learn to categorise the emotions by seeing examples of upright stimuli first. Furthermore, that there was no significant between-experiments’ difference between Experiments 1 and 1a shows that there were no differences between seeing full-length vignettes or shortened, 10-second examples of them. Therefore, there was sufficient information in a 10 second clip for recognition of the emotions (with the exception of disgust).

3.8. Experiment 1b

Experiment 1a has shown clearly that the overall results from Experiment 1 were reliable in that emotion can be perceived from displays of interpersonal communication and that the perception of fear, and to some extent other emotions too, is impaired by inversion. Furthermore, Experiment 1a shows that there were no learning effects transferred from upright to inverted versions of the same displays. It has also revealed that discrimination of the stimuli was not based on duration of the presentations as all the vignettes in Experiment 1a were cut to show only the first 10 seconds and there were no differences in ratings between these and the full length stimuli (Experiment 1). However, the consequence of this action meant that, although each film clip was the same absolute length as all the other clips, the fact remained that, due to some emotion types taking longer to act out, there may have been more information (due to more turn-taking behaviour being included) contained in the 10 second presentations for some emotion categories relative to the others. For example, anger presentations lasted for an average of 17.6 seconds when they were shown in full. Therefore, by reducing them to 10 seconds removed less material than for sadness presentations, which were on average 24.8 seconds long. Therefore, Experiment 1b was designed to eliminate this possible artefact. To achieve this, subjects were required to rate examples from each emotion category where the
vignettes shown took the actors similar amounts of time to act out. Figure 3.4.a) gives the ratings for each vignette shown in Experiment 1 and the amount of time that each vignette lasted. Figure 3.4.b) shows the 12 vignettes that were chosen for the present experiment. As the full vignette (with all 8 turn-takes) was shown the ratings given for these examples could then be compared with those from Experiment 1. Therefore, the amount of time that the vignettes were played for, and the amount of turn-taking contained within each one was constant. Therefore, if no between-subject differences were found, it could be confirmed that subjects were not using information based on, or as a consequence of, display duration for recognition of emotions.
Figure 3.4. a) Observers' mean ratings for each of the upright stimuli used in Experiment 1 as a function of total showing time; b) the subset of portrayals used for Experiment 1b.
3.9. Experiment 1b: Method

3.9.1. Subjects

Fifteen subjects participated in the experiment (3 males and 12 females; mean age = 36.7 years; age range 21-58 years).

3.9.2. Apparatus, design, stimuli and procedure

Two full-length examples from each of the six emotion categories were used from the Experiment 1 corpus. The examples chosen were those that were the most similar in length. The mean duration (in seconds) for each emotion category was: anger 19.5, joy 20.9, sadness 22.4, love 19.2, fear 20.65 and disgust 19.3 — a difference of only 3.2 seconds in mean duration between the slowest (sadness) and the fastest (love). Three emotions, anger, love and disgust differed by less than or equal to 3/10ths of one second.

Subjects were seated in front of a table where the displays were shown on the 14-inch G3 Apple Macintosh (laptop computer - described in Chapter 2). They were required to watch each display and rate them in the same way as described in Experiment 1. The stimuli were shown in random order and the experiment lasted for approximately 12 minutes. The ratings were recorded, and the non-portrayed scores were pooled, in the same way as in Experiment 1.
3.10. Results and discussion

Figure 3.5. shows the portrayed and non-portrayed mean ratings for each emotion category for this experiment along with the mean ratings for all the examples shown in Experiment 1. The non-portrayed results (as they are similar between both experiments, \( F_{1, 23} = 0.8, p > 0.05 \)) are considered no further here. A between-subjects ANOVA for the portrayed ratings only revealed no significant between-experiment effect \( (F_{1, 23} = 2.35, p > 0.05) \). Post hoc tests for each target emotion further confirmed that there were no significant differences between experiments \( (p > 0.05) \). Thus, these results establish that neither the varying duration of the vignettes nor the amount of turn-taking communication shown within each display affected observers ability to discriminate between emotions.

Figure 3.5. Observers' mean ratings for the upright presentations of the emotion response categories of the portrayed and non-portrayed emotions shown in Experiments 1 and 1b. The mean ratings for the non-portrayed emotions were derived from the mean of the combined scores from the five non-portrayals in each case. Error bars show standard errors.
3.11. Qualitative analysis

The sliding scales dependent measure that is used in most of the experiments reported in this thesis allows for the possibility that observers might have been able to discriminate between emotions by a process of elimination. Therefore, by allowing observers to express freely what emotion(s) they perceived in the stimuli would validate the quantitative scales if the portrayed emotions were spontaneously recognised by observers.

3.12. Qualitative analysis: Method

3.12.1. Subjects

The subjects were 5 postgraduates and 7 undergraduates from the Psychology department at the University of Surrey (5 male and 7 female; age range 19 to 57 years; mean age 25.75 years).

3.12.2. Apparatus, design, stimuli and procedure

The laboratory and computer set up was the same as that used in the previous experiments reported in this chapter except that a chin rest was not used. Two different examples of each of the six emotions were used from the Experiment 1 corpus and were shown to subjects in random order. Therefore, there were 24 responses for each emotion (one response for portrayed disgust was missing due to experimenter error). The criteria for the stimuli selection were 1) to use stimuli that received medium ratings (for example the anger stimuli chosen were those that received the second and the fourth highest ratings out of six in Experiments 1 and 1a) and 2) to include at least one example from each of the 10 actors. Displays that received the very highest ratings were not included as the high ratings may have been indicative of actors over-exaggerating the body movements made in their performances of the vignettes (see Chapter 5 for a discussion on exaggerated movements). The stimuli were presented in upright form and each stimulus was
shown for its full duration. The observers were seated in front of the monitor and were told that they would be viewing some point-light displays where people expressed emotions. They were informed that they were required to try to describe the displays in terms of any emotional content that could be discerned. The experimenter sat next to the observer and noted the emotion being shown along with descriptions given. The subjects were unable to see the experimenter’s notes as they were asked to watch the monitor constantly and the room was dimly lit. As a back up measure all descriptions were also recorded on audiotape. The experimenter controlled the computer to show each trial once the observer indicated that they were ready. The process for showing all 12 displays and obtaining the relevant descriptions of them lasted approximately 15 minutes.

3.13. Qualitative analysis: Results

The observers’ main responses can be seen in Table 3.3. All of the words used as attempts to describe emotions are included. Words used to describe physical movements are only included when they are salient to the observers’ dialogues or when an observer said little else.

On every occasion apart from 3, when observers expressed that they could not identify emotion, each observer made insightful comments about the displays being watched, although they did not always specifically describe emotions. For example, “intense in something, something quite deep” (HM) was used to describe a love portrayal and “turning away, in a huff with each other” (DN) was used to describe an anger portrayal.

Each of the emotion categories is considered in turn below.
Table 3.3. Table of observers’ free responses to the point-light displays. Words are typed in capital letters when the portrayed emotion is expressed directly.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Anger</th>
<th>Joy</th>
<th>Sadness</th>
<th>Love</th>
<th>Fear</th>
<th>Disgust</th>
</tr>
</thead>
<tbody>
<tr>
<td>M.J.</td>
<td>Shouting. ANGER. Angry. HAPPY. Jumping around indicates happier.</td>
<td>Calmer again, hard to interpret. SADNESS.</td>
<td>Ashamed or relaxed.</td>
<td>Two kids waiting for third to smash a window.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-----------------</td>
<td>---------------------</td>
<td>---------------------------</td>
<td>-----------------------------------------------</td>
<td>-------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>D.N.</td>
<td>Turning away, in a huff with each other.</td>
<td>Annoyed or fooling around or having a good time.</td>
<td>Hanging heads so SAD and not happy.</td>
<td>Confrontational.</td>
<td>Ducking or waving. Annoyed or frustrated.</td>
<td>No sense. Disappointment or dejection but not sure.</td>
</tr>
<tr>
<td>M.P.</td>
<td>One on right pissed off and one on left not apologetic either.</td>
<td>Angry, or they’re very HAPPY.</td>
<td>Upset and anxious, bowing their heads.</td>
<td>Interested in each other, flirting perhaps.</td>
<td>Nervous.</td>
<td>Friendly. Bored. Annoyed. Frustrated.</td>
</tr>
<tr>
<td>Time</td>
<td>Emotion</td>
<td>Mood</td>
<td>Emotions</td>
<td>Socialization</td>
<td>Intimacy</td>
<td>Warmth</td>
</tr>
<tr>
<td>------</td>
<td>---------</td>
<td>------</td>
<td>----------</td>
<td>---------------</td>
<td>----------</td>
<td>--------</td>
</tr>
<tr>
<td>A.M.</td>
<td>ANGER.</td>
<td>JOY.</td>
<td>SAD.</td>
<td>Don't know.</td>
<td>Sad.</td>
<td>Angry.</td>
</tr>
<tr>
<td>P.H.</td>
<td>ANGER.</td>
<td>Aggression, Frustration.</td>
<td>Little interaction. Emotionally low key. Not much.</td>
<td>Sharing excitement, not sure if negative or positive.</td>
<td>Intimacy. Warmth. Don't know it confused me.</td>
<td></td>
</tr>
<tr>
<td>H.P.</td>
<td>ANGRY</td>
<td>Again anger.</td>
<td>A bit upset.</td>
<td>Not much out of that.</td>
<td>Building up to something but can't think of emotion. Deflated. A bit...yeh whatever.</td>
<td></td>
</tr>
</tbody>
</table>
3.13.1. Anger

This was the emotion category most easily identified by observers as on fifteen occasions observers mentioned the word anger in their descriptions of the anger displays. On the 9 occasions that anger was not mentioned, that an argument was taking place was claimed 5 times, usually backed up with descriptions such as ‘not happy’ (MJ), ‘enraged’ (KF) or ‘confrontational’ (AB). The remainder of the time observers made comments such as “in a huff with each other” (DN), “pissed off with each other” (MP) or “aggressive” (JS). On no occasion was there a comment made that could be attributed to the perception of positive affect.

3.13.2. Joy

On 9 occasions out of 24, observers described the stimuli as ‘happy’ and on one occasion ‘joy’ was mentioned. On 7 other occasions ‘excitement’ was used as a descriptive. Observers were ambivalent over their judgements in 8 instances where they made comments about the displays showing both anger and happiness. However, within this category, it was apparent that observers’ were sometimes changing their minds about what emotion they were seeing as the displays unfolded. For example, “angry... happy... jumping around indicates happier” (MJ), or, “aggressive, excited, excited actually” (MJ). Other descriptions included phrases such as ‘friendly aggression’ (PH). On only 3 occasions, and therefore for only 12.5% of descriptions, were the joy portrayals described as negative affect per se.

3.13.3. Sadness

For almost 50% of the time, the word ‘sadness’ was mentioned as either the only, or one of a few similar (e.g. miserable or depressed) feelings. On two occasions, observers described the figures as looking upset and on one occasion each, the words “solemn” (AB) and “depressed” (HP) were used. For the rest of the time the stimuli were described as looking bored (twice), insecure, confused, angry and curious (once each). On two occasions subjects expressed either no knowledge of what was
being displayed ('don’t know') or that there was little emotional interaction to be perceived.

3.13.4. Love

The word ‘love’ was used twice to describe the love portrayals. Other words and phrases were used that may be considered as subordinate categories of love (see Fehr and Russell, 1991) such as “embrace” (MJ), “kind” (CJ), “care” (CJ), “intense” (HB), “friendly” (CJ), “flirting” (MP), “gentle” (PH) and “lust” (MJ). On 10 occasions the descriptions had no relation to the concept of love, for example, “confrontational” (DN), “don’t know” (AM), or “rowing” (HP).

3.13.5. Fear

The word ‘fear’ was used on two occasions. Words that can be associated with this emotion such as “anxiety” (CJ), “afraid”, “scared”, “apprehension — flight or fight” (MJ) were used on another 10 occasions. On another occasion, a scenario approach was used by a subject when they said “like 2 kids waiting for a third to smash a window” (MJ). This response could be interpreted as the subject’s perception of a fearful event about to take place. On 6 occasions the subjects did not get an impression of fear; for example, they perceived dancing or being cross with each other. Twice, the observers were either not sure of the emotion, e.g., “not sure if it’s positive or negative” (PH), or, “building up to something but can’t think of the emotion” (HP).

3.13.6. Disgust

Not once did any of the observers use the word ‘disgust’ or any word that could be associated with this emotion (e.g. ‘revulsion’ or ‘repugnance’). Instead, a wide variety of adjectives were used to describe the stimuli such as “depressed”, “bored”, “happy”, “angry” or “sad”. On some occasions, observers made several discordant descriptions of the stimuli at once, for example, one observer described for one stimulus, “sadness, boredom, annoyance and empathy”. There were 3 instances
where observers said that the stimulus made "no sense, "was odd" or that they were "confused" by it.

3.14. Qualitative analysis: Discussion

This analysis was designed as a measure of validation for the quantitative, sliding scales dependent measure that has been used in most of the experiments reported in this thesis. When the sliding scales were used, it was possible that observers might have been able to use a process of elimination to identify the emotions. Here it has been demonstrated that in many instances observers are freely able to name either the emotions directly or use descriptions of the stimuli which are broadly indicative of the emotions shown for most of the categories. As in the quantitative results, the degree to which observers were successful in freely identifying affect varied depending on emotion category and this variation largely coheres with the quantitative ratings given in experiments 1, 1a and 1b. For example, both the quantitative and qualitative results show that anger was an easily identified emotion as it received ratings of around 70% or more in the former and, in 15 out of 24 showings, the word 'anger' was attributed to the displays, which is also high (62.5% correct). The rest of the time observers gave descriptions that implicitly implied an angry scene. The success rate was also good (42%) for the free identification of joy. That this emotion was also sometimes confused with anger, was consistent with the quantitative results. Interestingly, anger was never confused with joy here and nor was it in the quantitative data (except in one case in Experiment 1a). Sadness was identified freely for about 50% of the time. It was surprising that the observers only identified love explicitly on two occasions. This may be due to the fact that love is an emotion rarely talked about in a scientific context and therefore observers, who were all psychology students, were not expecting love to be the correct answer. In support of this view, when observers were debriefed after the test many expressed surprise and amusement that love was the correct emotion term required and acknowledged that the relevant point-light displays were indicative of love. Despite the few times that the word 'love' was used as a descriptor, many terms were used that are considered as subordinate categories of this emotion (Fehr and Russell, 1991). As in
the case of love, fear was identified explicitly only twice. However, that fear was identified frequently with words, or on one occasion with a scenario, associated with fear shows that fear was implicitly recognised for more than 50% of the vignettes shown.

Of particular interest was the finding that on no occasion here was disgust identified explicitly, nor could any description of these stimuli be associated with anything that might be interpreted as disgust. Instead, a range of descriptions were given such as 'happy', 'angry', 'boredom' and 'lethargic'. It was, in some cases, as if the observers were guessing at what the stimuli might be and, in other cases observers admitted to being confused by the displays. These findings concur with the quantitative data, which shows that disgust was very confused with other emotion categories.

Overall, the qualitative findings corroborate and validate the quantitative method used throughout this thesis. Therefore, the rating scale procedure was used with confidence as it was measuring observers’ perceptions of the specific emotion categories portrayed and the ability to recognise emotions is not an artefact of the sliding scales method.
3.15. Experiment 2

Experiments 1 and 1a have shown that for some emotions inverting the stimuli on the image plane severely impedes perception. This replicates findings by others for both non-emotive, point-light walkers (e.g. Pavlova and Sokolov, 2000) and emotive stimuli alike (Dittrich et al., 1996). Few studies have looked at how rotated stimuli (other than 180° inversions) affect perception. One exception was the study by Pavlova and Sokolov (2000) who found that, with orientations of 90° – 180°, spontaneous recognition of a point-light walker was seriously impaired compared to upright displays. There have been no studies where emotive stimuli have been rotated by 90°.

From studies where either actors or dancers express emotions for observers to classify into specific movement categories (Wallbot, 1998) or, where specific movements are made for observers to decide which emotion they indicate (de Meijer, 1989; Sawada, Suda and Ishii, 2003), there is evidence to suggest that specific movements and postures indicate certain emotions. For example, lateralised movements of the hands have been found to specify anger but not joy (Wallbot, 1998). In contrast, vertical movements, where dancers stood on their toes, stretched the trunk of the body and lifted their heads upwards were indicative of joy but not anger (de Meijer, 1989). When stimuli are turned on the image plane by 90° lateral movements will appear vertical and vertical movements will look horizontal (although they do not change in terms of the actor’s egocentric co-ordinates). Therefore, by rotating anger and joy portrayals on to their sides veridical perception of the stimuli may be impaired: anger portrayals may be confused for joy and vice versa (if these directional cues are critical for perception). The main purpose of this experiment was to show observers rotated presentations to see how important vertical and horizontal movements were for the recognition of joy and anger respectively.
It was found that portrayed fear was severely degraded when the stimuli were inverted (Experiments 1 and 1a). Therefore, a secondary enquiry was made to see how fear was affected when the stimuli were turned on their sides. Although the interest here was in only 3 out of the original 6 emotion categories, all of the stimuli were shown to observers in both upright and rotated displays. This was done partly so that observers would not have too few categories to discriminate between and partly to generate more data for upright portrayals so that the data (for upright portrayals) from Experiments 1, 1a and 2 could be combined in order to look at sex differences.

There are many studies that examine sex differences for emotions but given that the literature is divided between both expressivity and experience of emotion, and that the methods by which assessments are made vary greatly, actually there are few studies that measure sex differences in recognition of emotion from movement of the body. Although women are suggested to be the more ‘emotional’ sex (Kring and Gordon, 1998) the evidence for this is far from conclusive. Indeed, Dittrich et al. (1996) found no sex differences in the recognition of affect from body movement. Therefore, the examination of sex differences here will shed some more light on this issue.

### 3.16. Experiment 2: Method

#### 3.16.1. Subjects

A total of 20 observers (10 males, and 10 females; age range 18 - 41 years; mean age: 21.4 years) participated in the study. Eighteen were undergraduates and 2 were postgraduates and the undergraduates received a course credit for their participation.

#### 3.16.2. Apparatus, stimuli, design and procedure

The experimental set up and the stimuli used were the same as in Experiment 1. The stimuli were presented in the normal upright orientation and horizontally; half of the
horizontal presentations were rotated clockwise (and half anticlockwise), thus the light-point heads would appear on either the far left or the far right side of a boxed area of the monitor. The boxed area subtended approximately 8.9° of visual angle vertically and 9.9° horizontally. There were 72 trials in all (6 emotions x 6 samples x 2 orientations) and stimuli were identical to those used in Experiment 1a. Each stimulus was shown for its first 10 seconds only. They were shown in one randomised block.

Observers were told that they would see some points of light on the screen that represented two people in a social context, both of which would be acting out only one of six possible emotions. The subjects were informed that in some trials the orientation of the stimuli would be altered. As in Experiment 1, a chin rest was used and subjects were instructed not to tilt their heads (in the preparation of the stimuli the experimenter had noticed that there was a natural tendency to turn the head to one side in order to make out the stimuli and so this was an important instruction that was not necessary for Experiments 1 and 1a). The experiment lasted for 50 minutes.
3.17. Experiment 2: Results

For each portrayed emotion the means were pooled so that comparisons could be made between one incorrect (non-portrayed) score and one correct (portrayed) score. Figure 3.6. shows the mean ratings of the portrayed emotions as well as the combined mean scores for the non-portrayed emotions when the displays were presented at 90° from vertical. Table 3.4. presents the complete set of ratings for both upright and rotated stimuli.

![Bar chart showing mean ratings for portrayed and non-portrayed emotions.](image)

**Figure 3.6.** Observers' mean ratings for the 90° rotated presentations of the emotion response categories of the portrayed and non-portrayed emotions. The mean ratings for the non-portrayed emotions were derived from the mean of the combined scores from the five non-portrayals in each case. Error bars show standard errors.
Table 3.4. Observers mean ratings for each emotion response category when stimuli were a) upright and b) rotated 90°. Figures in bold type denote the mean ratings for the portrayed emotions. Standard deviations are shown in parentheses.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Response</th>
<th>Anger</th>
<th>Joy</th>
<th>Sadness</th>
<th>Love</th>
<th>Fear</th>
<th>Disgust</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Anger</strong></td>
<td><strong>Joy</strong></td>
<td><strong>Sadness</strong></td>
<td><strong>Love</strong></td>
<td><strong>Fear</strong></td>
<td><strong>Disgust</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upright presentation</td>
<td>72 (16.4)</td>
<td>4 (4.1)</td>
<td>8 (6.5)</td>
<td>3 (2.5)</td>
<td>13 (3.2)</td>
<td>27 (22.1)</td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td>23 (16.9)</td>
<td><strong>47 (17)</strong></td>
<td>6 (4.2)</td>
<td>12 (9.1)</td>
<td>12 (10.2)</td>
<td>9 (9)</td>
<td></td>
</tr>
<tr>
<td>Joy</td>
<td>7 (5.7)</td>
<td>4 (4.3)</td>
<td><strong>59 (13.7)</strong></td>
<td>10 (7.5)</td>
<td>18 (14.9)</td>
<td>12 (11.9)</td>
<td></td>
</tr>
<tr>
<td>Sadness</td>
<td>10 (8.7)</td>
<td>8 (14.5)</td>
<td>15 (12.7)</td>
<td><strong>41 (29.5)</strong></td>
<td>12 (16.6)</td>
<td>12 (11.6)</td>
<td></td>
</tr>
<tr>
<td>Love</td>
<td>11 (10.1)</td>
<td>11 (10.5)</td>
<td>21 (13)</td>
<td>6 (8.8)</td>
<td><strong>48 (22.7)</strong></td>
<td>15 (16.4)</td>
<td></td>
</tr>
<tr>
<td>Fear</td>
<td>13 (11.2)</td>
<td>9 (10.2)</td>
<td>37 (25.1)</td>
<td>5 (6.2)</td>
<td>19 (20.1)</td>
<td><strong>25 (18.8)</strong></td>
<td></td>
</tr>
<tr>
<td>b)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rotated presentation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td>72 (13.1)</td>
<td>6 (6.5)</td>
<td>10 (11)</td>
<td>7 (6.4)</td>
<td>14 (14.6)</td>
<td>24 (17.6)</td>
<td></td>
</tr>
<tr>
<td>Joy</td>
<td>30 (20.8)</td>
<td><strong>39 (24.5)</strong></td>
<td>9 (9.6)</td>
<td>8 (7.5)</td>
<td>11 (8.7)</td>
<td>12 (14.8)</td>
<td></td>
</tr>
<tr>
<td>Sadness</td>
<td>9 (9.9)</td>
<td>9 (10.7)</td>
<td><strong>40 (18.5)</strong></td>
<td>19 (15.7)</td>
<td>14 (12.5)</td>
<td>16 (13.7)</td>
<td></td>
</tr>
<tr>
<td>Love</td>
<td>9 (10.8)</td>
<td>8 (11.3)</td>
<td>18 (15.4)</td>
<td><strong>41 (25.8)</strong></td>
<td>15 (16.2)</td>
<td>9 (9.3)</td>
<td></td>
</tr>
<tr>
<td>Fear</td>
<td>16 (12.9)</td>
<td>15 (12.3)</td>
<td>18 (13.4)</td>
<td>9 (7.8)</td>
<td><strong>29 (18.5)</strong></td>
<td>20 (15.4)</td>
<td></td>
</tr>
<tr>
<td>Disgust</td>
<td>17 (12.7)</td>
<td>9 (15.7)</td>
<td>29 (16.7)</td>
<td>10 (9)</td>
<td>15 (12.2)</td>
<td><strong>25 (13.5)</strong></td>
<td></td>
</tr>
</tbody>
</table>

An ANOVA with 2 variables: display type (2 levels, vertical and horizontal) and emotion (6 levels), revealed a significant main effect for emotion ($F_{2,76.5239} = 20.8$, $p < 0.001$) and for rotation ($F_{1,59} = 49.1$, $p < 0.001$) showing that ratings for rotation were significantly reduced relative to the upright stimuli. A significant interaction between emotion and orientation ($F_{5, 295} = 5.52$, $p < 0.001$) showed that rotation affected the ratings of some portrayed emotions more than others. Post hoc Tukey’s

83
tests following the ANOVA revealed that the ratings for joy (p < 0.05), sadness (p < 0.001) and fear (p < 0.001) were significantly reduced by rotation while anger, love and disgust were unaffected. There was no significant difference found for the way that subjects rated the 90° left and 90° right stimuli (F_{5,35} = 1.52, p > 0.05).

One-way ANOVAs that compared the ratings for the portrayed and all of the five non-portrayed emotions when the stimuli were rotated showed that there were significant main effects for all of the portrayed emotions (portrayed anger, F_{3,59,54} = 93.74, p < 0.001; joy, F_{1,71,33,74} = 13.12, p < 0.001; sadness, F_{2,86,54,33} = 16.33, p < 0.001; love, F_{2,69,51,13} = 12.99, p < 0.001; fear F_{5,95} = 4.79, p = 0.001 and disgust F_{2,83, 53,33} = 8.13, p < 0.001). Post hoc Tukey's tests showed that for portrayals of anger, sadness or love, the portrayed emotion was rated significantly higher than each of the 5 non-portrayed emotions individually (p < 0.01). This was not true for ratings of portrayed fear as although ratings for portrayed fear were higher than anger, joy and love (p < 0.05), they were no higher than ratings for (non-portrayed) sadness and disgust (p > 0.05). Ratings for (portrayed) disgust were confused with (non-portrayed) sadness (p > 0.05) but were larger than anger, joy, love and fear (p < 0.001). Of particular interest was the result that portrayed joy was confused with (non-portrayed) anger (p > 0.05) but was not confused with any other emotions (p < 0.005).

3.18. Experiment 2: Discussion

The main finding of this experiment was that the portrayals of joy and fear were severely impeded by showing the displays at 90° orientations. In comparison to Experiments 1 and 1a, rotated presentations of fear portrayals were not as degraded as when inversions were shown when, almost without exception, fear was confused with all of the non-portrayals. In Experiment 2 fear was confused with only two emotions: sadness and disgust. However, that this is so shows that bodily expressions of fear appear to be particularly sensitive to effects of showing non-veridical orientations *per se*. It may prove interesting to investigate the showing of fear
stimuli at orientations of less than $90^\circ$ to see how close to upright positioning the stimulus needs to be before veridical perception is reinstated.

It was expected that recognition of anger would be impaired to some degree as it has been proposed that lateralised body movements are indicative of this affect (Wallbot, 1998). However, that perception of this emotion remained good shows that either horizontal movement is still easily recognised despite rotation or that it is not a critical cue. Bodily expressions of anger do consist of other movements and positioning as well (e.g. arms stretched out frontal, Wallbot, 1998). Moreover, Pollick et al. (2001) found that affect (including anger) could be identified better than chance from point-light displays (of arm knocking movements) when displays were both inverted and phase scrambled; therefore detail of specific form is lost. This suggests that some low-level aspects of the displays such as velocity of the dots are useful for recognition of affect, and so quality of movement may not be an essential cue.

Whilst horizontal movement is not necessary to specify anger, it appears that vertical movement is important for veridical recognition of joy as the data for joy here are impaired when the stimuli are turned by $90^\circ$. Furthermore, portrayed joy is confused with (non-portrayed) anger in this condition. This result suggests that vertical movement is an important cue for the recognition of joy and thus corroborates with observations in the literature (e.g. de Meijer, 1989) that joy is specified by vertical movement and posture. This observation is also reflected in lay knowledge of this emotion — e.g. in the expression ‘jumping for joy’. That perception of joy remained good when inverted images were shown bolsters this claim (except for one anomalous result in Experiment 1a) as it shows perception is dependent on the vertical direction of movement irrespective of veridical positioning. That portrayed joy was confused with non-portrayed anger suggests that, whilst horizontal movement is not critical for perception of anger portrayals, it is an intrinsic feature of anger.
3.19. Sex differences

The literature is replete with studies investigating sex differences in the expressiveness and response of emotions. Numerous studies have been conducted on how men and women express and experience emotions both psychologically (e.g., self reports) and physiologically (e.g., skin conductance responses). The majority of studies suggest that women are more emotional than men (e.g., Wagner, McDonald and Manstead, 1986; Zuckerman, Lipets, Koivumaki and Rosenthal, 1975) although the findings are not always clear and are often inconsistent. Whether the cause of sex differences is environmental or genetic for emotions is complicated as it is likely that the cause of at least some differences is because of the way men and women are socialized (Grossman and Wood, 1993). Furthermore, men hide their feelings more than women (Gross and John, 1998). To complicate the issue further, the sociality of the experimental situation influences how emotions are expressed, i.e., the experimenter will act as an additional eliciting stimulus (Buck, Losow, Murphy, and Costanzo, 1992). In a solitary context (i.e. viewing a film alone) there is less likelihood that such display rules will be operative (as there is no social stimulus present to activate or inhibit them).

For the particular case of the perception of emotions from full-light showings of body movements (where the actor’s back was faced towards the camera so no facial information could be seen) females identified fear portrayals with more accuracy than men but no other sex differences were found out of the emotions tested, which included joy, surprise, fear, sadness anger and disgust (Sogon and Izard, 1987). However, Dittrich et al. (1996) found no sex differences for emotion recognition in full-light and point-light displays.

Here, the data for the veridical (upright) portrayed emotions were combined from Experiments 1, 1a and 2. This meant that there were 48 subjects included in the analysis (20 male: 28 female, age range 18-41 years). It was possible to combine data from these 3 experiments as the stimuli and experimental design in each were the same and the only differences were that each stimulus in Experiment 1 was played...
for its full duration whereas in Experiments 1a and 2 the stimuli were clipped to 10 seconds in length.

It is evident from Figure 3.7 that there were differences in ratings between the sexes for portrayals of anger, joy, sadness, and fear: men gave higher ratings to the stimuli than women. The target emotions for the upright portrayals only were entered into a 2 (sex: female, male) x 6 (emotion: anger, joy, sadness, love, fear and disgust) ANOVA. This revealed a significant main effect for emotion ($F_{5, 230} = 31.35, p < 0.001$) and was followed by a non-significant interaction between sex and emotion ($F_{5, 230} = 1.59, p > 0.05$). There was a significant between-subjects effect for sex ($F_{1, 46} = 4.18, p < 0.05$). Six, two-way ANOVAs with one between-subject factor (sex) and one within-subject factor—the ratings for the portrayed and the five non-portrayed emotions given for each category—showed no significant differences between ratings for males and females except for portrayed fear ($F_{3, 251} = 0.88, p > 0.05$, joy, $F_{2, 34, 107.5} = 1, p > 0.05$, sadness, $F_{3, 157.1} = 0.73, p > 0.05$, love $F_{3, 131} = 0.85, p > 0.05$, fear $F_{3, 142} = 4.05, p < 0.01$, disgust $F_{3, 146} = 0.86, p > 0.05$). Follow up tests for fear indicated that men gave higher ratings for portrayed fear than women ($p < 0.001$). No other differences were found. The full set of ratings for males and females can be seen in Table 3.5.
Figure 3.7. Observers' mean ratings for the upright presentations of the 'correct' response category for the portrayed emotions as a function of sex. Error bars show standard errors.

Table 3.5. Observers mean ratings for the upright portrayals shown in Experiments 1, 1a and 2. The ratings for males and females are shown separately. Figures in bold type denote the mean ratings for the portrayed emotions. Standard deviations are shown in parentheses.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Anger</th>
<th>Joy</th>
<th>Sadness</th>
<th>Love</th>
<th>Fear</th>
<th>Disgust</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Females</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td>69 (13.7)</td>
<td>5 (3.5)</td>
<td>20 (20.9)</td>
<td>6 (6.2)</td>
<td>16 (15.8)</td>
<td>33 (25.8)</td>
</tr>
<tr>
<td>Joy</td>
<td>23 (17.4)</td>
<td>41 (16.1)</td>
<td>10 (10.4)</td>
<td>15 (12)</td>
<td>14 (9.4)</td>
<td>14 (15.1)</td>
</tr>
<tr>
<td>Sadness</td>
<td>12 (12.4)</td>
<td>6 (5.2)</td>
<td>53 (18)</td>
<td>14 (12.8)</td>
<td>18 (12.6)</td>
<td>18 (14.9)</td>
</tr>
<tr>
<td>Love</td>
<td>14 (11.8)</td>
<td>11 (13.6)</td>
<td>20 (14.4)</td>
<td>42 (23.4)</td>
<td>13 (12)</td>
<td>11 (9.9)</td>
</tr>
<tr>
<td>Fear</td>
<td>18 (14)</td>
<td>12 (9.7)</td>
<td>22 (14.2)</td>
<td>10 (9.9)</td>
<td>42 (20.5)</td>
<td>17 (15.5)</td>
</tr>
<tr>
<td>Disgust</td>
<td>20 (15.2)</td>
<td>8 (7)</td>
<td>33 (19.4)</td>
<td>10 (9.3)</td>
<td>20 (18.7)</td>
<td><strong>30 (15.4)</strong></td>
</tr>
<tr>
<td><strong>Males</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td>73 (14.2)</td>
<td>7 (8.3)</td>
<td>14 (17.7)</td>
<td>8 (8.9)</td>
<td>20 (15.3)</td>
<td>32 (18)</td>
</tr>
<tr>
<td>Joy</td>
<td>29 (16.8)</td>
<td>50 (18.8)</td>
<td>8 (5.9)</td>
<td>17 (14.6)</td>
<td>13 (10.4)</td>
<td>17 (16.4)</td>
</tr>
<tr>
<td>Sadness</td>
<td>11 (9.6)</td>
<td>7 (7.8)</td>
<td>58 (17.2)</td>
<td>16 (11.5)</td>
<td>18 (13.8)</td>
<td>15 (10.6)</td>
</tr>
<tr>
<td>Love</td>
<td>15 (10.9)</td>
<td>10 (13)</td>
<td>17 (9.4)</td>
<td>42 (21.4)</td>
<td>16 (15.5)</td>
<td>19 (14.8)</td>
</tr>
<tr>
<td>Fear</td>
<td>11 (8.6)</td>
<td>13 (11)</td>
<td>20 (12.8)</td>
<td>11 (12.2)</td>
<td>58 (16.7)</td>
<td>19 (14.6)</td>
</tr>
<tr>
<td>Disgust</td>
<td>17 (11.9)</td>
<td>11 (9.7)</td>
<td>41 (22.6)</td>
<td>10 (10)</td>
<td>18 (14.9)</td>
<td><strong>30 (18.5)</strong></td>
</tr>
</tbody>
</table>
Men gave higher ratings than women for four of the emotion categories tested. That men judged portrayals more intensely than women was a surprise given that women are generally reported as the more emotional sex and that the opposite effect was found by Sogon and Izard (1987). However, men have been found to display higher skin conductance reactivity when watching (full-light) fear-eliciting films (Kring and Gordon, 1998). Therefore, it appears that men give higher ratings to bodily expressions of fear and produce more physiological activity to fear-eliciting stimuli than women.

### 3.20. General Discussion

The present experiments were designed to investigate whether emotion can be recognised from the non-verbal cues displayed in point-light displays produced by pairs of actors engaged in dialogue. The prime motivation behind this was the fact that natural social situations have not yet been considered in the perception of affect from biological motion. This is considered a major omission given that emotions are communicative in nature. The first experiment clearly showed that the emotions of anger, fear, joy, sadness and love could be identified in such displays, which suggests that the spatio-temporal information inherent in natural human movement is sufficient to specify certain emotions. The exception in these results was for portrayed disgust, which was often confused with sadness, anger and fear. Qualitative data made the quantitative results more cogent by showing that the psychological process involved was recognition, rather than discrimination, between 6 response alternatives. Inversion reduced the salience of the displayed emotions to varying extents although overall performance was still good for most emotions except fear. In Experiment 2 the effect of 90° rotation was investigated. Results for the rotated displays showed that portrayed joy was confused with non-portrayed anger but the opposite was not true as portrayed anger was identified veridically. This suggests that directional cues are important for some emotions over others. It
was found that men gave higher ratings to portrayals of fear than women did but no other sex differences were found.

That a range of emotions can be recognised in biological motion displays of actors engaged in a dialogue is in line with previous research, which has generally used more stylised movements such as dance (e.g. Dittrich et al., 1996). However, the results reported here do not mirror some other results in the literature. For example, portrayed anger was the most easily perceived emotion out of the six shown here, which is not consistent with the findings of Dittrich et al. (1996) or Atkinson et al. (2004). In both of these experiments portrayals of grief and joy received higher ratings than portrayals of anger, and in the case of Atkinson et al.'s results fear did too. In fact, Atkinson et al. found that all of the emotions, even disgust, received very high scores when presented as point-light displays. In the results presented here only portrayed anger consistently received ratings this high. Such variation in findings may be accounted for by a number of differences in the methods used between studies. One important difference was that Atkinson et al. (2004) did not use point-light displays in the strict sense of the concept. Rather, strips of reflective tape were used, which changes the cues normally available in a point-light display in that the orientation of the limbs and wrists is more apparent, and furthermore, a sense of depth is given, especially to the shoulder joints. Therefore, more information is available to the perceiver than in a true point-light display. Other reasons for such high identification in the Atkinson et al. study may have been due to half of the observers having watched full-light displays (videos) one week before seeing the same displays in point-light. Furthermore, a 5-alternative forced choice dependent measure was used, which meant that observers were forced to make yes/no decisions about the stimuli shown.

The fact that fear was well identified here is consistent with Walk and Homan's (1984) results in point-light displays of an actor engaged in mime, and an actor expressing stereotypical movements (Atkinson et al., 2004) but contrasts with the poor identification of fear with other types of movement (e.g. musicians' movements presented in full-light: Dahl and Friberg, 2004; dancers' movements presented in
point-light: Dittrich et al., 1996). Based on their findings for poor recognition of fear, Dittrich et al. argued that to detect the presence of fear quickly a system would use only spatial information as this is instantaneous (dynamic information is not instantaneous as time has to elapse for it to occur). However, that fear was easily identified here suggests that a system for the perception of fear could be based on dynamic form information. Furthermore, for survival purposes, this emotion must be detected wherever it occurs in the visual field, irrespective of where attention is presently directed (Öhman, 2000).

Movement is detected easily regardless of attention, for example, whilst picnicking in a quiet park on a sunny day visual attention can be quickly diverted from a good book to focus on an animal running across a distant open field. However, the literature is replete with studies showing us that with minor disturbances in the visual array large changes in static scenes are easily overlooked — a phenomenon known as change blindness (e.g. Henderson and Hollingsworth, 1999; Rensinck, O’Regan and Clark, 1997; Simons, 1996). That spatial information is not used is not the point argued here; rather the point is that dynamic information is used and is most effective.

It is fundamental that any effective system is efficient, and so to use information that is veridical even if it takes an instant longer (dynamic information) to process is better than to be momentarily quicker (spatial information) but be incorrect. Öhman (2000) has argued that false negatives (i.e. failing to detect fear and elicit a response when it is present) are more costly than false positives (eliciting a response when there is no threat). Whilst this is true, a good system for detecting threatening situations should not make false positives because there are carry over effects for the next threatening encounter.

Interestingly, Bassili (1978) found that fear and anger were not easily identified in point-light displays of facial emotion but disgust was. However, in the present study disgust was poorly perceived. This is in agreement with Dittrich et al.’s (1996)
finding that disgust was not only poorly discriminated in point-light presentations of dancers’ movements but was also the least well identified in the corresponding full-light presentations of the same. Disgust was better discriminated in Atkinson et al.’s (2004) study but their actors were explicitly guided to avoid interpreting disgust as contempt or moral disgust, which was a measure not undertaken here (it is not known how the dancers interpreted disgust in Dittrich et al.’s, 1996 study). Moral disgust is likely to be mixed with feelings of anger, which may explain why disgust was confused with anger here but not in Atkinson et al.’s research. However, disgust did receive the lowest percent correct out of the emotions tested across all (full-light and point-light, moving and still) conditions in Atkinson et al.’s (2004) study. This finding combined with the present results and those of others (e.g. Dittrich et al., 1996) suggests that neither the body nor indeed other channels such as the voice (see Johnstone and Scherer, 2000) may be good channels to communicate this emotion regardless of whether representations of it are presented in full-light or abstracted body movements. To corroborate with this view, Coulson (2004) has demonstrated that disgust can not be recognised from static pictures of the body (whereas other emotions can). Although this emotion is associated with contempt (Darwin, 1872/1999) and morally objectionable behaviour (Ekman, 2003), it is also food related (Darwin, 1872/1999; Ekman, 2003). Indeed Rozin and Fallon (1987) argued for a narrow definition of disgust in that it is only related to food. That the face is a better channel of expression for disgust makes sense and in fact its distinctive facial expression has been well documented historically (e.g. Darwin, 1872/1999; Ekman and Friesen, 1974). Overall it appears that the face is a good channel for perception of some emotions, and the body for others.

A motivation in the present work has been to fulfil the need for ecological, rather than paradigm driven research. Clearly the differences in results between Dittrich et al.’s findings and those reported here might be at least in part due to the fact that dance movements were recorded in the former. Although Atkinson et al. (2004) did not use dancers in the making of their stimuli, that actors made their performances in isolation of others, and indeed in isolation from their surrounding environment
(given that tights were worn over their heads during recordings) may account for at least some of the reported differences between studies.

Evidence to suggest that the recognition of biological motion from point-light displays is impaired when inverted (Bertenthal and Pinto, 1994; Pavlova and Sokolov, 2000; Dittrich et al., 1996; Shipley, 2003) was supported by our findings as in general lower ratings were given to the inverted stimuli compared to the upright ones. This effect was particularly pronounced for fear. Similarly, Dittrich et al. (1996) found that inverted displays of fear were given higher ratings for non-portrayed grief than for fear. However, the effects for inversion were far more pronounced in this study, which may reflect the differences in types of portrayals used between studies.

The foregoing evidence suggests that there may be special processing constraints that relate to fear and not to other emotions. Further, there is evidence in the literature from brain imaging and experimental studies to suggest that there may be a special module for the processing of fear. For example, recent brain imaging research has revealed that stimuli showing bodily expressions of fear and stimuli showing bodily expressions of happiness activate different brain regions (de Gelder, Snyder, Greve, Gerard and Hadjikhani, 2004; see also Dolan and Morris, 2000). Their findings have shown that fearful expressions depicted in full body videos yielded high activity in areas known to process emotional information (such as the amygdala) whereas this was not the case in similar comparisons of either happy or neutral but meaningful body expressions (e.g. putting on trousers or opening a door). Furthermore, it is also known that the amygdala is active in the processing of biological motion (Bonda, Petrides, Ostry and Evans, 1996; Pavlova, Birbaumer and Sokolov, 2005). In addition to activation for emotion specific areas, de Gelder et al. (2004) found that fear stimuli increased activation in areas responsible for action representation and motor cortex. Thus, the combined activity in these areas could amount to a fast track mechanism for fear contagion designed for preparation of action. Therefore, it appears that bodily expressions of fear could be processed differently in the brain to other emotions. This result is substantiated by a study of NM, a patient with damage to the amygdala, who was unable to recognise fear from photographs of body
posture, although able to perceive anger, joy and sadness as well as controls could (Sprengelmeyer, Young, Schroeder, Grossenbacher, Federlein, Büttner and Przuntek, 1999).

The existence of a special system for fear recognition could explain why observers were unable to see fear when presented in unnatural orientations (both inverted and rotated): Whereas for the other emotions enough high level processing could be carried out for veridical perception from the altered orientations this would not be the case for a fast-acting system where non-veridical portrayals of fear would be meaningless as they are ecologically invalid. Consistent with this hypothesis, it has been postulated that a special low level route for the processing of fear exists in the brain whereby the detection of fear stimuli can be picked up almost instantaneously without awareness and with minimal cortical processing (LeDoux, 2000; and see Öhman, 2005 for a recent review). It is plausible that the existence of a fast-acting system for the perception of fear could explain why such a strong inversion effect was found for boxing behaviour (Dittrich, 1993) if this was interpreted as fear by the visual system — both types of behaviour are likely to be similar in quality of movement style, i.e. crouching and jerking backwards (and sometimes forwards) movements.

In conclusion, these studies clearly demonstrate that emotions, with the exception of disgust, can be perceived from point-light displays of interpersonal behaviour. These findings cohere largely with those of similar studies although some differences are apparent. That men gave portrayals of fear higher ratings than women did was of interest given that the percept of fear appears to hold special status in the visual system; supported by the finding that fear recognition was completely diminished for its non-veridical portrayals. This was an arresting result but one that could be explained if a special system exists for fear identification.
Chapter 4

Effects of social interaction

4.1. Introduction

Of central importance in this thesis is the understanding of how perception of emotion is modified by the social context within the display. The aim of the work described in this chapter was to take a first step at examining certain aspects of communication with regard to recognition of affect from biological motion. Two important issues relating to the perception of a social event were explored. The first pertained to the importance of the natural dyad in the perception of affect by a third person observer. Specifically, the aim of this experiment (Experiment 3) was to determine whether performance of the observer differed when the original social context of the display was manipulated. As well as the original stimuli with two actors, observers were shown two other variations: the first was a single actor (the other was occluded) and the second was a simulated dyad without the rhythmic timing of interlocution. This was constructed by mirror-imaging a single actor on one side of the midline (so that each point-light person made the same movements at the same time). Therefore, by studying the natural dyadic interaction as well as situations where this was artificially removed, the importance of the social context will be investigated.

The second experiment reported in this chapter (Experiment 4) pertained to the communicative interaction independently of affect. It is clear that animate motion is an important source of information for an observer and that socially relevant states such as affect can be recognised. In the past, the use of point-light displays has demonstrated that other important intraindividual characteristics such as sex (e.g. Barclay, Cutting and Kozlowski, 1978) and identity (e.g. Troje, Westhoff and Lavrov, 2005) of an actor can be recognised. It is essential to establish what other characteristics can be picked up from biological motion in order to understand the parameters of the social information available from movement cues. Without
communication, humans would not be social. This experiment investigates a primary aspect of communication, i.e. speech. It takes an important first step to understand something about the interpersonal characteristics that are conveyed via the kinematics of the display, independent of affect. The question asked was whether the speaker can be identified.

4.2. Experiment 3

The social importance of emotions was implied over a century ago in the writings of Darwin (1872; 1999) but it is only now that the social dimension of emotionality is becoming considered as fundamental in some of the current literature (e.g. Campos, Campos and Barrett, 1989; Parkinson, 1996; Ekman, 1999; Blair, 2003; Manstead, 2005) (see Chapter 1). However, there is still a dearth of empirical studies that consider emotion at this level.

Primarily, emotions should be considered at a social level as emotions are always about something and these things are normally people or events (Manstead, 2005). Further, during communication, emotions are frequently shared between people and an emotion expressed by the first person (the actor) informs the second person about the actor’s current state and also gives feedback to the actor about how the actor is perceived. Importantly, the behaviour of the second person will be based on their appraisal of the actor’s behaviour. Therefore, rather than thinking of emotions as states that reside only at the intrapersonal level, emotionality can be thought of as a dynamic relational process that occurs between the individual and the environment (Campos, Campos and Barrett, 1989; Blair, 2003). According to this view the study of emotions at a social level must be of interest as emotions have interpersonal or intergroup regulatory outcomes.

These social appraisals of others’ emotional states are observable from an early age. For example, an infant will crawl across a solid table top and then onto and across a
piece of transparent glass (which creates an illusory drop for the infant) only when the infant’s mother encourages the child to do so by smiling: when the mother appears to be fearful the baby remains on the perceived safe-side of the table (Sorce, Emde, Campos and Klinnert, 1985). The social exchange of affect in mother/child interactions is two-way, for example, toddlers have been shown to use particular displays of affect (sadness) in order to gain support from their social environment (Buss and Kiel, 2004).

Typically, positive emotions (e.g. joy and love) occur in social contexts (Shiota, Campos, Keltner and Hertenstein, 2004). These authors have argued that positive emotions are particularly important at a social level because they serve functional purposes to help with bonding within families, romantic partnerships, friendships and also groups. For example, momentary displays of romantic love serve to promote and maintain long-term commitment behaviours in relationships (Gozanga et al., 2001).

Empirical evidence shows that positive emotions are facilitated by interpersonal interactions such as smiling, which was found to be positively correlated with the presence of a friend when happy films were viewed (Jakobs, Manstead and Fischer, 1999). Smiling was also enhanced when friends watched sad films together (Jakobs, Manstead and Fischer, 2001). Surprisingly then, in the latter case, although the stimulus was a sad film the presence of another person promoted smiling behaviour and inhibited the production of sad expressions. Conversely, sad expressions were produced when the sad movies were watched alone. Here it is important to draw a distinction between the expression and perception of emotions as, although this research provides evidence that the social context is important for the production of positive emotions, little is known about whether this is equally true for the perception of positive affect.

What do we know of the social interactions that serve to communicate affect? Humans are equipped to be able to both send and receive sensory information via multiple modalities simultaneously (Birdwhistell, 1970; de Gelder and Vroomen,
2000) and theories of communication differ on what information is important for the exchange (see Chapter 1). However, in everyday interactions turn-taking is an observed phenomenon that serves to function as an important aspect of communication (Argyle, 1967). Turn-taking develops early and is evident in mother-infant interactions (Trevarthen, 1993; Bernieri, Reznick and Rosenthal, 1988; Baron and Boudreau, 1987; Field, Healy, Goldstein and Guthertz, 1990). Furthermore, imitation within a dyadic exchange by each partner of the other partner’s actions serves to facilitate similar behaviour (Hess, Philippot and Blairy, 1999). For instance, smiles appear to be initiated by one partner and then imitated by the other (Cappella, 1993). In general, mimicry is postulated to serve important social functions such as to communicate empathy, to enhance similarity between partners and to increase liking of the interacting partners (Hess et al., 1999). Furthermore, social synchrony and patterns of entrainment can be readily perceived by third parties not involved in the exchange (McClave, 1994). Moreover, relating to the same, film clips of natural mother-infant interactions which are edited and reconstructed so that the behaviour of one interactant in one time period (e.g. the first minute) is paired with the behaviour of the other interactant in another time period (e.g. the third minute) (pseudointeractions) can be discriminated from real interactions (Bernieri et al., 1988). Listeners involved in the exchange move their bodies but not always in a way that signifies that they want to take a turn to speak: this effect is known as back-channelling and it augments the first person’s speech (Duncan, 1972). Back-channelling is auditor-initiated and related movements have a powerful effect on the speaker as, without them, a speaker may start repeating an utterance or become quiet (Weiner, DeVoe, Rubinow and Geller, 1972). Head nods are a prototypical example of back-channelling but speakers make head nods too. Speaker head nods have been found to have an interactive function in that they trigger back-channels; listeners are very sensitive to such head nods, resulting in almost immediate responses (McClave, 2000).

The aim of the present study was to take a preliminary step at examining how the perception of emotion by third parties was affected by the social interaction seen in the vignettes. With the foregoing evidence in mind, the following predictions were
made: first, it was anticipated that the natural social interaction would augment the perception of emotion; second, it was expected that this effect would be stronger for positive emotions. In order to test these predictions the natural synchrony of the dyad was disrupted in two ways. Three viewing conditions were employed which varied the communicative element of the displays. The ability of subjects to recognise the emotions in the normal dyad condition was compared, firstly with a situation when only one actor was present (i.e. the other was occluded — ‘monads’) and, secondly, with a condition without the rhythmic timing of interlocution, i.e. by showing a single actor on one side of the midline with his or her point-lights reflected about the midline (i.e. the mirror image of the same actor making the same movements at the same time — ‘reflected dyads’). This last measure was necessary for two reasons: first, it provided a control for the natural dyads — the total amount of dynamic information (i.e. number of point-lights) displayed in both of these conditions was the same; second, it introduced a pseudointeraction, in that there were two moving point-light figures facing each other in every display but no natural social interaction was present. The dependent measure was the same as that employed for the experiments reported in Chapter 3.

4.3. Experiment 3: Method

4.3.1. Subjects

Seven postgraduates and 11 undergraduates (11 women and 7 men; mean age 21.5 years; range 18-33 years) volunteered to serve as subjects in the experiment.

4.3.2. Design and procedure

There were three experimental conditions: dyads, monads and reflected dyads (see section 4.2). Owing to the findings in the previous experiments reported in this thesis (see Chapter 3) the stimuli which portrayed disgust were not included. Therefore, in this experiment only five emotions were used. Although six examples of each
emotion were used in the experiments reported in Chapter 3 only four examples of each emotion were used for each conditions in the present experiment. This measure was necessary in order to keep the experiment within a manageable time scale for the participants. The stimuli were chosen from the original corpus used in experiments 1,1a and 2; the discarded portrayals were the ones that received the 2 lowest mean ratings for each emotion in Experiment 1 (Chapter 3) — for anger and sadness these were stimuli with means < 50, for joy and fear < 30, and love < 40. There were twice as many trials in the monad condition as each actor was shown in a separate trial. This gave a total of 80 trials, with an even number of dyads (albeit that some were reflected dyads) and monads. The first 10 seconds of each film were shown. The subjects were told that either one or two people would be seen where both were acting out one of five emotions. The ratings procedure was the same as in Chapter 3 except here there were only 5 emotions shown on the rating scales. The stimuli were shown in one block, in random order for each subject. The experiment was approximately 40 minutes long.

4.4. Experiment 3: Results and discussion

Figure 4.1. shows the mean ratings for each portrayed emotion in each viewing condition and Table 4.1. shows the complete matrix of means from Experiment 3.

A two-way ANOVA between display type and emotion with 3 levels of display type (dyad vs. monad vs. reflection) and 5 levels of emotion was conducted on the ratings given to the portrayed emotions (the bold figures in Table 4.1). 1 Significant main effects were found for display type ($F_{2.34} = 8.2, p = 0.001$) and for emotion ($F_{4.88} = 16.6, p < 0.001$). There was also a significant interaction between display type and emotion ($F_{4.75.4*} = 5.65, p < 0.001$).

---

1 An important assumption for within-subjects ANOVA is that sphericity should not be violated as, if it is, the possibility of a type I error may be greatly inflated (for a discussion see Howell, 1997, chapter 14). Where the data have failed the sphericity test in the present experiments (as indicated by the symbol *) the more conservative Greenhouse–Geisser epsilon test has been used, which reduces the degrees of freedom of the numerator and denominator.
Table 4.1. Observers' mean ratings for each emotion response category when dyads, monads and reflected stimuli were presented. Figures in bold type denote the mean ratings for the portrayed emotions. Standard deviations are shown in parentheses.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Response</th>
<th>Anger</th>
<th>Joy</th>
<th>Sadness</th>
<th>Love</th>
<th>Fear</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dyads</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td>72 (16.1)</td>
<td>5 (3.6)</td>
<td>18 (17.2)</td>
<td>11 (13.9)</td>
<td>15 (7.9)</td>
<td></td>
</tr>
<tr>
<td>Joy</td>
<td>22 (18.9)</td>
<td>56 (23.5)</td>
<td>7 (5.6)</td>
<td>18 (20.4)</td>
<td>7 (5.2)</td>
<td></td>
</tr>
<tr>
<td>Sadness</td>
<td>8 (5.4)</td>
<td>7 (5.6)</td>
<td>45 (23.9)</td>
<td>28 (21.5)</td>
<td>14 (13.5)</td>
<td></td>
</tr>
<tr>
<td>Love</td>
<td>15 (13.9)</td>
<td>13 (11.1)</td>
<td>12 (8.8)</td>
<td>43 (26.7)</td>
<td>9 (8.6)</td>
<td></td>
</tr>
<tr>
<td>Fear</td>
<td>23 (19.6)</td>
<td>16 (13.4)</td>
<td>14 (12.6)</td>
<td>16 (13.7)</td>
<td>40 (19)</td>
<td></td>
</tr>
<tr>
<td><strong>Monads</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td>69 (25.4)</td>
<td>5 (7.3)</td>
<td>14 (12.3)</td>
<td>7 (6.2)</td>
<td>14 (17.7)</td>
<td></td>
</tr>
<tr>
<td>Joy</td>
<td>24 (15)</td>
<td>46 (19.7)</td>
<td>7 (4.5)</td>
<td>11 (9)</td>
<td>11 (6)</td>
<td></td>
</tr>
<tr>
<td>Sadness</td>
<td>8 (6.4)</td>
<td>5 (3.7)</td>
<td>52 (23.5)</td>
<td>9 (7.8)</td>
<td>17 (11.9)</td>
<td></td>
</tr>
<tr>
<td>Love</td>
<td>14 (19.5)</td>
<td>8 (4.9)</td>
<td>26 (16.5)</td>
<td>17 (19.3)</td>
<td>19 (23.3)</td>
<td></td>
</tr>
<tr>
<td>Fear</td>
<td>15 (13.8)</td>
<td>12 (9.3)</td>
<td>16 (11.6)</td>
<td>8 (5)</td>
<td>46 (21.6)</td>
<td></td>
</tr>
<tr>
<td><strong>Reflections</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td>68 (20.8)</td>
<td>6 (6.5)</td>
<td>22 (23.2)</td>
<td>10 (12.6)</td>
<td>10 (9.4)</td>
<td></td>
</tr>
<tr>
<td>Joy</td>
<td>27 (18.7)</td>
<td>42 (22.2)</td>
<td>8 (7.1)</td>
<td>13 (15.6)</td>
<td>8 (8)</td>
<td></td>
</tr>
<tr>
<td>Sadness</td>
<td>5 (3.2)</td>
<td>8 (9.1)</td>
<td>35 (25)</td>
<td>27 (27)</td>
<td>19 (18)</td>
<td></td>
</tr>
<tr>
<td>Love</td>
<td>19 (22.16)</td>
<td>13 (11)</td>
<td>10 (7)</td>
<td>36 (21.4)</td>
<td>16 (18.6)</td>
<td></td>
</tr>
<tr>
<td>Fear</td>
<td>16 (14.9)</td>
<td>13 (10.4)</td>
<td>19 (17)</td>
<td>21 (21.8)</td>
<td>30 (18.9)</td>
<td></td>
</tr>
</tbody>
</table>
Post hoc Tukey’s tests revealed that portrayed joy received higher ratings when displayed as natural dyads ($p < 0.02$) than as monads or reflections. Portrayed love received significantly higher ratings when seen as natural or reflected dyads relative to monads ($p < 0.001$). Interestingly, Tukey’s tests also showed that the monad presentations of sadness and fear portrayals received higher ratings than the reflected ones ($p < 0.05$). No other significant differences were found.

![Graph showing observers' mean ratings for the emotion response category portrayed by the actors when displayed as natural dyads, monads, and reflections. Error bars show standard errors.](image)

**Figure 4.1.** Observers’ mean ratings for the emotion response category portrayed by the actors when displayed as natural dyads, monads, and reflections. Error bars show standard errors.

One-way ANOVAs and post hoc Tukey’s tests were carried out on all of the results across each row in Table 2 and all of the portrayed ratings were found to be statistically higher than the corresponding non-portrayals ($p < 0.05$) except for 3
cases. First, after the one-way ANOVA for the 5 ratings of monad portrayals of love ($F_{2.2, 37.4} = 3.5, p < 0.05$) Tukey's tests revealed that love was confused with all of the non-portrayed emotions ($p > 0.05$). Second, following the one-way ANOVA for the 5 ratings of reflection portrayals of sadness ($F_{1.7, 33.3} = 9.54, p = 0.001$) Tukey's tests showed that ratings for portrayed sadness were no different to ratings of non-portrayed love and non-portrayed fear. Third, for the reflected images of fear ($F_{2.61, 44.45} = 3.07, p < 0.05$), fear ratings were not significantly different from those for any of the non-portrayed emotions except joy.

These findings show that performance was markedly altered by the absence of the second actor in either veridical or non-veridical (mirror image) formats, or both, for all the emotions except anger. Recognition of the emotions displayed by natural dyads was in some cases better than with only one actor (the monad condition) or without rhythmic synchronisation between two actors engaged in conversation (as in the reflected dyads). In particular, the removal of one actor affected the perception of joy and romantic love, in that successful identification of these emotions was significantly impaired when the second actor was not present.

### 4.5. Experiment 4

Experiment 3 showed that the social context makes a difference to how affect is perceived. However, very little is known of the hidden structures and regularities that exist in social interactions per se. The issue is complex as verbal and nonverbal behaviour varies greatly in terms of importance and distribution within the interaction. Furthermore, the grammar (i.e. the hidden regularities) within the exchange may also contain a parallel stream of behaviour from each person that occurs independently of the interaction (Magnusson, 1996). For instance, movements that serve the needs of the body have been termed 'adaptors' (e.g. head scratching or squeezing a part of the body) and these occur with little intentionality or awareness (Ekman and Friesen, 1972). Adaptors can be observed within an exchange but are not used as a deliberate method of communication (Ekman and Friesen, 1972).
Adaptors are obvious examples of parallel behaviour that occur alongside the interaction without being related to it. Other far less salient non-communicative behaviours are likely to exist and therefore it is not difficult to see why an understanding of the patterns and structure within a social exchange remain elusive.

We know that body movement and speech are both fundamental in natural human communication (Bachorowski, 1999; Wachsmuth, 2002) and it has been argued that speech and body movements are closely connected in terms of evolutionary processes (Rizolatti and Arbib, 1998). In everyday conversations verbal communication occurs with nonverbal behaviour. For example, when a person speaks, their arms, fingers and head move in a temporally structured way that coordinates with what is spoken (Ekman and Friesen, 1972; Condon, 1986). Thus, speech and body movement can be considered as two parts of the same process (Birdwhistell, 1970). As communication naturally occurs this way — indeed it is almost impossible to suppress body movements whilst speaking (Dobrogaev, 1932, cited in McClave, 2000) — it means that ideally speech should be empirically investigated in situations when it is accompanied by nonverbal communication.

There is also evidence to show that speech and body movements are connected. For example, from microanalysis of videotaped speech, it was found that during speech the head rarely stopped moving and these movements became more rapid with verbal amplitude (Hadar, Steiner, Grant and Clifford Rose, 1983). Further, discrete speaker initiated head movements appear to carry distinct meanings: for example, side-to-side shakes were found to correlate with verbal expressions of intensification (such as ‘very’, ‘a lot’ and ‘great’) of the speech (McClave, 2000).

For the case of emotion recognition, links have been reported between speech and other modalities. For example, performances on the identification of specific emotions such as happiness, sadness, fear and anger have been shown to correlate across facial, prosodic and lexical channels (Borod, Pick, Hall, Sliwinski, Madigan, Obler, Welkowitz, Canino, Erhan, Goral, Morrison and Tabart, 2000). In another
study, the task allowed subjects to process emotional information visually by studying facial expressions, whilst simultaneously listening to emotion expressed in voice prosody (de Gelder and Vroomen, 2000). These researchers found that the perception of emotion from the face was influenced by the vocalisations that subjects heard; this occurred (although to a lesser extent) even when observers were asked to ignore the facial displays. For example, when the task was to judge a facial expression as happy or sad the subjects were more likely to judge a face as happy when they heard a simultaneously presented happy utterance and this happened even when participants were asked to ignore the utterance. Furthermore, the converse effect was observed, i.e. auditory perception of emotion was influenced by visual information (de Gelder and Vroomen, 2000). These authors suggest that observers’ judgements were influenced by the unattended modality (the one ignored). Therefore, evidence exists for the mandatory cross-modal integration of information processing.

That cross-modal processing exists should not be a surprise as emotions are most often expressed through more than one modality at a time. For example, a person may express hot anger whilst shouting, punching the air with a tightly clenched fist whilst also tightly gritting his teeth, his eyes fixed and with his eyebrows raised — thus combining vocal, bodily and facial expressions. Furthermore, there is evidence to show that the neural structures involved with the processing of affect are shared by different modalities of expression (Hadjikhani and de Gelder, 2003).

Although links have been found between facial and vocal expressions of affect, there appears to have been no research conducted to explore directly the bimodal connection between body movement and vocal expressions of emotion. Therefore, the aim of the present study was to investigate the relationship between the visual detection of a speaker and (full body) biological motion when actors were expressing affect.
If speech and nonverbal behaviour are related then it might be expected that an emoting articulator can be distinguished from an emoting non-articulator from biological motion alone. Here it is investigated whether the body movement of the emoting actor that opens the conversation in each point-light movie is sufficient for specification of the articulator and if so whether success in identification is constant for each of the emotion categories. Subjects were shown the same movies as those used in previous experiments (Experiments 1, 1a and 2) reported in this thesis (without the inverted or rotated displays) but this time they were asked to decide which person talked first. As the aim of this experiment was to investigate the relationship between speech and biological motion rather than emotion per se, disgust portrayals were reincorporated into the experimental design.

4.6. Experiment 4: Method

4.6.1. Subjects
Five postgraduates and 12 undergraduates (4 men and 13 women; mean age 25.7 years; range 19 - 46 years) volunteered to serve as subjects.

4.6.2. Design and procedure
A total of 36 trials were used (6 emotions x 6 samples) in all. For every original recording it was always the actor on the left that initiated the conversation, therefore, to show an equal number of trials where the person on the left and the right talked first half of the stimuli (3 for each emotion) were transformed in Matlab so that the original positions for each actor were reversed (the actor on the left appeared on the right and vice versa). Each stimulus was shown for the first 10 seconds and all were shown in one randomised block. Subjects were told that they would see some points of light on the screen that represented two people in a social context. They were asked to judge which person talked first in each presentation and that therefore it would be important for them to be particularly attentive at the beginning of each clip. Throughout each presentation the trial number was displayed in the bottom right hand corner of the monitor. At the end of each trial the subject recorded this number.
as well as the identity of the person (the one on the left or the right) that they thought initiated the conversation. A sheet of paper was provided for this. Once the subject had recorded their answer they pressed a key to be shown the next trial. The experiment lasted for approximately 20 minutes.
4.7. Experiment 4: Results and discussion

For each individual, an accuracy score for correct identification of the first speaker was calculated where a score of 1 was given for each correct choice, therefore a range from 0 to 36 was created, with a mean score of 18 expected by chance. The average accuracy score was 22.3, which was shown to be significantly above chance when a one-sample $t$ test was used, $t(16) = 7.25$, $p < 0.001$. Generally then, observers were able to extract from biological motion information about who was talking.

To enquire whether this ability was similar for each of the six emotions portrayed the separate scores for each emotion category were calculated for each participant. This amounted to six separate scores for each observer with a range of six and a mean score of three to be expected by chance. Figure 4.2. shows the accuracy of speaker identity for each emotion category. One-sample $t$ tests were used to compare actual performance with that expected by chance. For anger, joy, sadness and love, these tests showed that speaker recognition was above chance (anger, $t(16) = 11.14$, $p < 0.001$; joy, $t(16) = 6.13$, $p < 0.001$; sadness, $t(16) = 2.52$, $p < 0.05$; love, $t(16) = 2.28$, $p < 0.05$) but for fear and disgust portrayals this was not the case (fear, $t(16) = 2.04$, $p > 0.05$; disgust, $t(16) = 0.20$, $p > 0.05$). Thus, speaker recognition from biological motion appears to depend on which emotion category is portrayed.
A one-way ANOVA was conducted between the emotion categories which revealed a highly significant result, $F_{5.86} = 15.46, p < 0.001$. In order to understand where the differences lay, post hoc Tukey’s tests were carried out that revealed for anger portrayals, performance of identifying who was talking first was better than for all the other groups individually ($p < 0.001$). For joy portrayals, speaker identification was better than for portrayals of love, fear and disgust ($p < 0.05$) but for portrayals of sadness, speaker identification was only better than for portrayals of fear ($p < 0.001$).

Generally, these results show that observers were able to tell who was talking at the beginning of each interaction although the ability to identify the speaker varied depending on which emotion was being portrayed and, when anger was being portrayed, recognition was the best. This is a considerable achievement given that discrimination was based purely on movement information specified in point-light displays where facial information was not included. The fact that the interlocutor was not identified correctly when fear was expressed is of interest as this shows that in not all cases can the interlocutor be identified even when the emotion can be
recognised. This was not the case for disgust, as although the articulator was not identified nor was the emotion (see Chapter 3).

4.8. General discussion

The experiments reported here were designed to investigate two independent yet interrelated phenomena relating to the social interaction. The first enquiry asked how the social context affected the perception of emotions. The second question asked whether a speaker could be identified in the point-light displays. In Experiment 3 it was found that the social context affected perception of the emotions in most cases. Joy and love were better recognised when natural dyads were observed and the perception of love was particularly impaired when monads were shown. Furthermore, reflected displays impaired recognition for all of the emotions except for portrayals of anger and love.

Joy and love are positive emotions that are particularly expressive socially and, in Experiment 3, they were recognised better when displayed in a two-actor context which was in line with predictions. Joy is typically expressed more in social contexts (Jakobs, et al., 1999; Jakobs, et al., 2001; Schaver, Schwartz, Kirson and O’Connor, 2001) and romantic love is important for the attachment process and plays an informative role in indicating to each partner that the other is trustworthy and can be depended on (Shiota, et al., 2004). Importantly then, not only are positive emotions (love and joy) expressed more strongly in social contexts, the results reported here show that the third party perception of them is more powerful in social contexts too. It is noteworthy that fear, joy and sadness were less easily perceived in the reflected dyads and joy and love less so in the monads. In general, these results show that subjects process details about the social context of a scene and not just the intrinsic movements of an individual actor.

The best example of an observers’ sensitivity to context cues was the fact that portrayals of romantic love were unrecognisable without the presence of a second
actor despite the fact that the stimulus properties for each actor were no different between monad and dyad conditions.

Context cues have been shown to be important in other research. Shipley and Cohen (2000) showed observers point-light displays of sportsmen playing a variety of sports (e.g. basketball, tennis and football). These were made from video footage of real games where the information was reduced to point-lights that corresponded to the sportsmen's hands, feet and head. A typical interaction showed, for example, one person dribbling the ball whilst another was competing for possession of it. These authors found that when the ball was specified in the display (as a single point of light) observers made accurate judgements regarding the coordinated action shown in the displays. However, when the ball was missing, performances were qualitatively impaired as subjects could not make sense of the coordinated relationship between the sportsmen.

Further, if an implied ground plane (in the form of a single or moving horizontal background line) is added to the top of a display when inverted point-light walkers and moving dogs are shown, observers were able to explicitly identify (above chance) the otherwise unrecognised walkers and dogs (Pavlova and Sokolov, 2003). Thus, it appears that when contextual information is available specifying physical or social information beyond the relations that exist within a single point-light human, emergent properties of the display unfold.

Shipley and Cohen (2000) have argued that their results might be better understood if they are considered in an affordance based framework. This suggestion means that properties of the environment surrounding the point-light human in the displays supply information for the third person observer to take appropriate action. This argument could be extended to include the social environment. For example, in the case of Experiment 3, observers might have picked up information that allows them to avoid or meet other individuals or couples. Recent evidence has shown that the way observers perceive human walking varies from when the observer is still
compared to when the observer carries out the same actions as the display (Jacobs and Shiffrar, 2005). These authors asked subjects to view 2 simultaneously presented point-light walkers where gait speeds differed between the displays. The subjects were required to compare the gait speeds of the walkers. This task was completed whilst the observers stood still, walked, or bicycled. More accurate discriminations were made when observers were either still or when they cycled. In other words the production of walking movement interfered with walking perception. Further differences were demonstrated depending on the observer’s position relative to the stimuli (Jacobs and Shiffrar, 2005). This shows that, by providing more realistic environmental conditions than those used traditionally in a laboratory, the results are altered markedly — observer movement interferes with visual perception. The challenge is now to incorporate variations to the third person’s social behaviour within the experimental design rather than or as well as variations of the visual stimuli in order to comprehend whether an affordance based framework is appropriate for further understanding of these results.

Experiment 3 strongly demonstrated that emergent properties unfold within interpersonal interactions; these transmit socially relevant information that may provide affordances for the third person to take action. Experiment 4 also showed that there are qualities inherent within the social exchange that advise the third person observer of important communicative characteristics. Here it was demonstrated that a speaker can be identified even if the only information available is a point-light display of two people engaged in conversation where, importantly, facial information is not available. Therefore, speech and body movement are related. In this experiment articulator recognition varied depending on which emotion was presented and for anger displays speaker identification was at its best.

What aspects of bodily expression could have been informative for speaker recognition here? At this time only speculative answers can be offered as identification could be linked to qualitative or quantitative aspects of the movements and it is unknown what contributions spatial information had per se. However, whatever dimensions were used for speaker identification some useful clues have
been revealed. Speaker identification was significantly above chance for most emotions and so some general feature that is common to all displays is likely to be operative. This may be information from amount of movement as anger and joy typically feature a lot of movement and these were the emotions where best articulator recognition occurred. The speaker was not identified for displays of fear and disgust and, in contrast to the possible explanation for portrayals of joy and anger, it is unlikely that these non-significant results were due to little movement activity as low amounts of movement also existed for portrayals of love and sadness and yet the speaker was identified when these emotions were acted out. Indeed, in some of the disgust portrayals movement activity was reminiscent of portrayed anger, with large and expansive, fast-moving arm movements. Therefore, some other cues must be responsible for interfering with identification — quite possibly these were nonverbal back-channel effects from the interacting partner. In other words the movements made by the listener competed with, or were related to, those made by the speaker.

In summary, the experiments reported here are the first to demonstrate that the social context within a biological motion display affects how emotions are perceived. These findings provide clear evidence that perception of emotion is altered for some emotions when two actors are displayed or when non-veridical contexts are shown. In addition these data show that speaker recognition is possible from the minimal cues provided in (full-body) point-light displays depicting interlocution. This is quite a remarkable achievement, given the fact that there was no facial information present in these minimal cue stimuli. If an ecological framework is employed these effects may be better understood.
Chapter 5
Further analysis of the mechanisms

5.1. Introduction

The main experimental method of investigating how biological motion is processed by the perceptual system is by observing how subjects respond to displays where the images are distorted. For example, in the first empirical chapter it was shown that by rotating point-light displays by 90° and 180° perception was impeded thus demonstrating the existence of orientation based processing constraints for biological motion. Other research has demonstrated similar effects for inversion (e.g. Dittrich et al., 1996). However, in general there is a paucity of data in biological motion research that investigates emotion. This is surprising given that the number of studies exploring biological movements for action is so high.

In this chapter the stimuli are manipulated in several ways in order to give some insight into what the specific cues for affect might be and to investigate further the way they might be processed. The first experiment (Experiment 5) explores the processing constraints of the perceptual system when the stimulus is a complete (13 point) display. However, in the second experiment (Experiment 6) the stimuli are altered so that fewer points are shown. This last manipulation provides a way of investigating whether there are particular characteristics within a display that specify affect.

Specifically, Experiment 5 investigates whether affect can be recognised in displays where the information is distorted by reversing the motion (playing the displays backward). This is a relatively neglected procedure but an important one to consider as although the information in forward-played and backward-played displays is equal in terms of its physical properties, it may be different from an ecological point of view (Pavlova, Krägeloh-Mann, Birbaumer and Sokolov, 2002). Therefore, observers are shown displays where the stimuli are played backward and the results
of this experiment are compared with the same stimuli being played normally (Experiment 3).

Experiment 6 pertains to the attributes of the stimulus in a way that questions whether information from the whole stimulus is necessary for the perception of affect. A theme that runs through all of the results reported in the thesis so far is that categories of affect appear to be processed differently. For example, the perception of affect was augmented by the presence of a natural dyad (Experiment 3) but only for particular emotions (love and joy); and the recognition of fear (Experiment 1) was at chance when the stimuli were turned upside down but this was not observed for other emotions. As the evidence shows that particular emotions are processed differently to others then might there be distinctive characteristics that specify each of them? Authors such as Darwin (1872, 1999) Wallbot (1998), and Atkinson et al. (2004) have argued that there are distinctive movements and postures associated with specific emotions. For example, Wallbot (1998) has postulated that lateral hand and arm movements are indicative of anger and forward facing shoulders indicate fear. However, all of these researchers made their observations based on movement patterns of the full body. Therefore, in order to investigate the possibility that more information is present in a full-body stimulus than required, and to specify what configurations are fundamental for perception of affect, the stimuli are broken down in order to explore some of their specific components.

5.2. Experiment 5

The perception of reversed motion displays of walking point-light dogs was investigated by Pavlova et al. (2002). The dogs walked as if they were on a treadmill, i.e. there was no net translation. When the displays were played backwards it was found that observers, in a free verbal response task, were able to recognise some form of biological motion (e.g. 2 walkers or a cat in motion). Interestingly, in reversed motion displays subjects perceived the motion as forward facing rather than backward facing (Pavlova et al. 2002). For instance, one interpretation was 'a giraffe moving to the right' (in the forward-played condition) and, for the same display, as 'a cat with a tail lifted' (in the backward-played condition). Therefore, observers
appeared to be able to identify some form of biological motion in both conditions but interpretations altered depending on the direction of movement. Performance was impaired compared to normal (forward-played) displays as Pavlova et al. (2002) documented that, for the reversals, there was a slower response time, a greater number of interpretations as to what the displays may have been and a number of interpretations of the displays being more than one object, for example, 'two workers carrying a tube' (in the forwards played displays they were always interpreted as a single entity). The correct responses were given more often when the displays were played normally.

Performance tends to also be impaired in the recognition of human gaits played in reverse. For instance, in a task where observers were required to judge a non-translating point-light walker's direction of motion (in comparison to a vertical counterphase grating backdrop) it took longer to identify the motion of backward-played walkers than those that were forward-played (Fujimoto, 2003). However, observers were able to correctly identify the direction of motion despite the increased reaction times. In another study, where (non-translating) walking motion was reversed, Vanrie, Dekeyser and Verfaillie (2004) documented a similar effect to Pavlova et al. (2002) in that observers who were asked the orientation of the figure perceived the walker as facing the observer. However, in this study observers had been informed of the backward-play manipulation prior to viewing the displays (observers were not informed of the manipulation in the Pavlova et al. (2002) study nor the Fujimoto, 2003 study). Therefore, the forward-facing interpretation was incongruous with what would be perceived normally as it implied that the walker was seen as moving away but also facing the observer (the most logical interpretation when knowing of the backward-play manipulation would be to perceive the walker as facing away from the observer). This effect is of interest as it suggests that observers may be using high-level top-down processes to interpret the image (Vanrie et al., 2004). This phenomenon might be thought of as analogous to the hollow mask illusion (Gregory, 1973) where a concave mask of the human face appears to be convex as it appears that the visual system attempts to make some ecological sense of the mismatched information. However, whereas the hollow mask illusion may be accounted for in terms of effects of experience (faces are never really
hollow), Vanrie et al. (2004) have argued that familiarity cannot account for the forward-facing effect as people are often seen facing away from the viewer. However, it is reasonable to argue that observers are more familiar with seeing a facing view of a conspecific over a backward view of the same. Therefore, experience could account for the forward-facing phenomenon.

Clearly, compared to normal displays, perception is altered when observers view reversed motion displays of quadrupeds and walkers. However, the percept is not so degraded that observers cannot make some sense of the images. The present study examines whether observers can recognise affect from backward-played point-light displays. Affect and human (or animal) locomotion are qualitatively different. One salient dissimilarity is that there is a structured, pendulum like, pattern of movement when humans walk which is different to that exhibited when displaying affect. For purposes of efficiency, moving from one place to the other involves moving the limbs in a regular, predictable way but emotions appear not to be expressed by such regular movement of the limbs.

An important function of perception must be to detect in which direction the animal is moving or facing (to avoid collision). In contrast, when a human detects affect, direction of movement is not an essential feature. Therefore, the perceptual system is likely to operate in qualitatively different ways to perceive affect and to perceive objects. However, to know which way an emoting actor is facing is fundamental. Therefore, the goal of Experiment 5 was to establish whether and if so how backward-played displays would affect recognition of emotion.
5.3. Experiment 5: Method

5.3.1. Subjects

The participants were 13 volunteers: three postgraduates and 10 undergraduates (1 male and 12 females; mean age 27 years; range 19 - 47 years).

5.3.2. Design and procedure

The same stimuli were chosen as those used in Experiment 3: this amounted to 4 examples for each of the 5 emotions (owing to the findings in Experiments 1, 1a, 1b and 2 the stimuli which portrayed disgust were not included). The movies were transformed so that they could be played in reverse by the computer program. The first 10 seconds of each film were shown in reverse. Subjects were told that they would see some points of light on the screen that represented two people in a social context both acting out one of five possible emotions. However, they were not informed of the reverse nature of the stimuli. All of the stimuli were presented in one randomised block for each subject and the experiment lasted for approximately 20 minutes. The ratings procedure was the same as in Experiment 3.

5.4. Experiment 5: Results and discussion

For each portrayed emotion the means of the five non-portrayed emotion ratings were pooled so that comparisons could be made between just one ‘incorrect’ (non-portrayed) score and one correct (portrayed) rating (Figure 5.1). The (correct) rating for the portrayed emotion was higher in every case than for the (incorrect) non-portrayals. Table 5.1 shows all of the values individually as well as the values for the same stimuli when they were played forwards (Experiment 3). The correct ratings for backward-played anger and backward-played sadness were higher than the equivalent ratings for the forward-played displays. It is not clear why this should be so although the backward motions, unlike the normal displays, did not start from a neutral posture which may have made recognition of the backward-played stimuli
more easy. However, if this is the case then it is unclear why only some emotions were more easily recognised. For each portrayed emotion, an ANOVA was conducted between the ratings for that emotion and for the four non-portrayed emotions (i.e. along each row for the backward-played displays in Table 5.1). All ANOVAs showed significant main effects (portrayed anger, $F_{2.50, 30.94} = 94.21$, $p < 0.001$; joy, $F_{1.41, 16.95} = 10.98$, $p < 0.005$; sadness, $F_{1.96, 23.51} = 24.98$, $p < 0.001$; love, $F_{2.4, 28.66} = 15.06$, $p < 0.001$; fear, $F_{1.8, 21.67} = 13.84$, $p < 0.001$). Tukey's post hoc tests confirmed that for each stimulus the portrayed emotion was rated higher ($p < 0.05$) than each of the 4 non-portrayed emotions considered individually.

Table 5.1. Observers' mean ratings for each emotion response category when forward played (Experiment 3) and backward played point-light stimuli were presented. Figures in bold type denote the mean ratings for the portrayed emotions. Standard deviations are shown in parentheses.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Response</th>
<th>Anger</th>
<th>Joy</th>
<th>Sadness</th>
<th>Love</th>
<th>Fear</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forward played</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td>72 (16.1)</td>
<td>5 (3.6)</td>
<td>18 (17.2)</td>
<td>11 (13.9)</td>
<td>15 (7.9)</td>
<td></td>
</tr>
<tr>
<td>Joy</td>
<td>22 (18.9)</td>
<td>56 (23.5)</td>
<td>7 (5.6)</td>
<td>18 (20.4)</td>
<td>7 (5.2)</td>
<td></td>
</tr>
<tr>
<td>Sadness</td>
<td>8 (5.4)</td>
<td>7 (5.6)</td>
<td>45 (23.9)</td>
<td>28 (21.5)</td>
<td>14 (13.5)</td>
<td></td>
</tr>
<tr>
<td>Love</td>
<td>15 (13.9)</td>
<td>13 (11.1)</td>
<td>12 (8.8)</td>
<td>43 (26.7)</td>
<td>9 (8.6)</td>
<td></td>
</tr>
<tr>
<td>Fear</td>
<td>23 (19.6)</td>
<td>16 (13.4)</td>
<td>14 (12.6)</td>
<td>16 (13.7)</td>
<td>40 (19)</td>
<td></td>
</tr>
<tr>
<td>Backward played</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td>82 (15.3)</td>
<td>5 (3.7)</td>
<td>22 (16.7)</td>
<td>8 (6.8)</td>
<td>22 (16.6)</td>
<td></td>
</tr>
<tr>
<td>Joy</td>
<td>33 (25.9)</td>
<td>54 (23)</td>
<td>8 (10.2)</td>
<td>19 (14.8)</td>
<td>13 (20.4)</td>
<td></td>
</tr>
<tr>
<td>Sadness</td>
<td>8 (5.2)</td>
<td>9 (7.4)</td>
<td>63 (27.1)</td>
<td>24 (16.7)</td>
<td>16 (20.8)</td>
<td></td>
</tr>
<tr>
<td>Love</td>
<td>9 (7.2)</td>
<td>11 (11.4)</td>
<td>28 (17.3)</td>
<td>46 (22.8)</td>
<td>17 (18.4)</td>
<td></td>
</tr>
<tr>
<td>Fear</td>
<td>13 (12.3)</td>
<td>9 (8.4)</td>
<td>18 (17.7)</td>
<td>12 (10.7)</td>
<td>45 (21.2)</td>
<td></td>
</tr>
</tbody>
</table>
Figure 5.1. Observers' mean ratings for the backwards presentations of the emotion response categories of the portrayed and non-portrayed emotions. The mean ratings for the non-portrayed emotions were derived from the mean of the combined scores from the four non-portrayals in each case. Error bars show standard errors.

To compare the backward-played portrayals with the forward-played ones a two-way ANOVA with 5 levels of emotion and 2 levels of display type was conducted on the ratings given to the portrayed emotions. A significant main effect was found for emotion ($F_{4, 48} = 16.34, p < 0.001$) but no significant main effect was found for the direction of motion ($F_{1, 12} = 1.95, p > 0.05$). Further, there was no significant interaction between display type and emotion ($F_{4, 48} = 1.07, p > 0.05$). Therefore, no further analyses were carried out.

---

1 An important assumption for within-subjects ANOVA is that sphericity should not be violated as, if it is, the possibility of a type I error may be greatly inflated (for a discussion see Howell, 1997, chapter 14). Where the data have failed the sphericity test in the present experiments (as indicated by the symbol *) the more conservative Greenhouse – Geisser epsilon test has been used, which reduces the degrees of freedom of the numerator and denominator.
The results reported here show that observers are able to extract the cues available in point-light displays that enable them to recognise affect irrespective of the fact that the displays were played backwards. This shows that the visual system uses information independent of direction and therefore does not rely on the dynamic of the pattern of movement. Further, these results suggest that walking movements are qualitatively different to those that express affect.

### 5.5. Experiment 6

Experiment 5 has shown certain aspects of the temporospatial dynamic are not crucial for the perception of affect. In Experiment 6 the point-light displays are manipulated in another way in order to explore the mechanisms that underlie the perception of affect.

Evidence has been provided to show that as few as 4 or 5 moving points that relating to the knees and the ankles (Ahlström, Blake, and Ahlström, 1997) or knees, ankles and hips (Johannson, 1978) supply the observer with a vivid percept of a human walker. This suggests that there is considerable redundancy in a full-body point-light walker. In another study where the configurations of a point-light walker were manipulated so that either the extremities (wrists and ankles), mid-limb elements (elbows and knees) or central components (shoulders and hips) were removed the detection of a point-light walker (presented in a mask of visual noise) varied depending on which elements were missing (Pinto and Shiffrar, 1999). For example, although there was no difference between normal viewing and the condition where the extremities were missing, task performance markedly deteriorated when the elbows and knees and the shoulders and hips were removed. These findings are suggestive of a system that categorises components of the body (Pinto and Shiffrar, 1999). Moreover, that performance was significantly impaired when the mid-limb elements were missing suggests that the dynamic symmetry (i.e. the equal and opposite movement of the limbs) was important for discrimination. Further, that removal of the central limb components impaired discrimination suggests that the
principle axis of organisation (the torso) (see Marr, 1982) is also a fundamental feature.

As already mentioned (see 5.2) walkers move in a predictable way in that as one limb moves forwards the neighbouring limbs move backwards, in anti-phase to the first limb (Pinto and Shiffrar, 1999). Emotions are not expressed like this although there are specific movements that indicate affect. For example, lateral and fast-moving arm movements are associated with anger (Wallbot, 1998). Therefore, it may be expected that particular elements of the human body will specify affect but it is likely that these will be different to the ones that the visual system exploits to recognise a human walker.

Roether and Giese (2005) have taken a first step to explore the components of the body that are fundamental for specifying affect. These researchers produced biological motion stimuli of emotional gaits for anger, joy and sadness, which were split into upper body (arms, shoulder and head) and lower body (hip and legs) components. Subjects were asked to distinguish between emotional and neutral walks and to provide expressiveness ratings. Results showed that the lower and upper body appeared to contribute independent information about affect. The upper body was more informative than the lower body for expressions of anger and sadness but not for fear; this was identified more equally from both upper and lower body components. This evidence supports the view that all of the information contained in a full point-light display may not be necessary for recognition of affect. However, Roether and Giese (2005) only manipulated 2 body components (upper body and lower body). Moreover, they used simulated rather than naturally produced stimuli.

The aim of the present experiment was to examine whether the visual system exploits particular information from the human body. Varying exemplars of point-light displays that omitted particular points were created. Thus the displays appeared to have missing features of the body. Specifically, the questions asked were whether some information was redundant in a full point-light figure. If so, then the next enquiry was to see which subconfigurations were characteristic of human emotion and whether these were different for the various emotions. It was expected that there
would be some redundancy in the full 13 point-light figure. However, it was not known what characteristics of the human form would be important for recognition.

Five subconfigurations were employed in the design. The number of point-lights shown for each of the manipulations ranged between 3 and 8. Whilst these 5 combinations of elements are by no means exhaustive, they are designed to incorporate major features of the body that may be particularly important for the specification of affect. It was reasonable to assume that the head and shoulders may convey important information about affect as the face and voice are fundamental channels for emotional expression (e.g. face, Ekman, 2003; voice, Scherer, 2003). Therefore, showing only the head and shoulders made up the first configuration. Most objects have a principle axis which is important for recognition (Marr, 1982) and the principle axis by which the human body is recognised is the torso. Following Pinto and Shiffrar (1999) removal of the elements that define the torso, i.e. the shoulders and hips, was the second configuration produced. Although it was not expected to be critical for perception of affect the third condition was the removal of the interlimb joints, that is the elbows and knees. The latter condition eliminated pair-wise rigidity of the major joints and has been found to be important for accurate recognition of biomechanical motion (Pinto and Shiffrar, 1999). Movement of the arms may be an important feature for anger (e.g. Wallbot, 1998) and arm movements have been argued to be a fundamental human feature in terms of evolution (Rizzolatti and Arbib, 1998). Therefore, the last 2 subconfigurations were a) where all of the elements remained except for the arms and b) where only the arms were shown.
5.6. Experiment 6: Method

5.6.1. Subjects

Twelve females and eight males (mean age 21.3 years; range 18-30 years) volunteered to serve as subjects. The majority (17) were undergraduates.

5.6.2. Design and procedure

All of the stimulus displays were modified examples of the stimuli used throughout this thesis. As the design of this experiment was to establish what would be the minimum information required to perceive emotion, monads rather than dyads were used. Owing to the fact that monads were presented, portrayed love was excluded from the design as this emotion was not recognised in this format (Experiment 3). Four monads, for each of the remaining 4 emotions (anger, joy, sadness and fear) that had been used in Experiment 3 were used.

The stimuli were shown in the original (13 point-lights) format as well as in 5 other subconfigural combinations, which can be seen in Figure 5.2:

*Figure only head and shoulders.* (Fig. 5.2.a). All of the points apart from the head and shoulders were removed. Thus the least information in terms of light points was made available (3 points).

*All of figure except torso.* (Fig. 5.2b). The points that corresponded to the hips and shoulders were removed. This eliminated information from the centre of the body (8 points).

*All of figure except mid-limb elements.* (Fig. 5.2c). The knee and elbow points were removed. This eliminated the pair-wise rigid relations among the limb joints (8 points).
All of figure except arms. (Fig. 5.2d). The arms, specifically the wrists and elbows, were removed (8 points).

Figure only arms. (Fig. 5.2e). This figure contained only information from the wrists and elbows (4 points).

For all of the above subconfigurations except for 'figure only head and shoulders' the point-light that corresponded to the position of the head was also removed.

There were 96 trials in all (4 emotions x 4 samples x 6 configurations). Each stimulus was shown for 10 seconds. The experiment was conducted in two blocks. The first block comprised 80 trials (the subconfigurations) and the second block comprised 16 trials (the full body condition). The fact that the full point-light displays were played second eliminated any possible learning effects although this was unlikely (see Experiment 1a). The same rating scale procedure that has been used in the other studies of this thesis so far was employed as the dependent measure.
Figure 5.2. Illustrations of the presentations shown where a single frame of an anger monad is displayed in each of the 5 subconfiguration formats. a) Only the head and shoulders remained. b) All elements were shown except the torso. c) All elements remained except the interlimb joints. d) All elements remained except the arms. e) Only the arms were shown.
5.7. Experiment 6: Results and discussion

The means for each subject were calculated for every stimulus configuration shown. All of the results (with the exception of portrayed joy) for the full body condition were rated higher than the same (monad) condition in Experiment 3. There is no clear reason why this should be so although it is most likely due to individual differences within each group of subjects. Figure 5.3 summarises the ratings given for each (correct) portrayed emotion and it illustrates that observers' performances varied as a function of the missing element combinations. It also shows that the pattern of recognition for each combination was not uniform across emotion categories.

![Figure 5.3. Summary of results for each of the (correct) portrayed emotions for every stimulus configuration. Error bars show standard errors.](image)

To ascertain whether observers' performances varied depending on the element combinations shown, a two-way repeated measures analysis of variance with 4 levels of emotion and 6 levels of display type was conducted on the ratings given to the portrayed emotions. Significant main effects were found for display type ($F_{5, 95} =$
23.97, p < 0.001) and for emotion ($F_{3, 57} = 12.23, p < 0.001$). There was also a significant interaction between display type and emotion ($F_{15, 285} = 9.04, p < 0.001$). Planned comparisons confirmed that there was a decrement in performance compared to the whole body (13 point) condition for some of the reduced configurations but this effect varied according to emotion category. For portrayed anger recognition was impaired when only the head and shoulders remained ($t (19) = 7.04, p < 0.001$) and when all of the elements remained except for the arms ($t (19) = 11.16, p < 0.001$). For portrayed joy recognition was also impaired when all of the elements remained except for the arms ($t (19) = 2.32, p < 0.05$). The most disruption in performance was observed for portrayed sadness as for every subconfiguration, recognition was significantly impaired (head and shoulders remaining: $t (19) = 5.16$, $p < 0.001$, all parts remained except torso: $t (19) = 5.07$, $p < 0.001$, all parts remained except knees and elbows: $t (19) = 3.07$, $p < 0.01$, all parts remained except arms: $t (19) = 4.24$, $p < 0.01$, arms only remained: $t (19) = 4.15$, $p < 0.001$. Like sadness, recognition of portrayed fear was impaired for all of the missing element combinations except for when only the arms were removed: head and shoulders remained: $t (19) = 2.24$, $p < 0.05$, all parts remained except torso: $t (19) = 3.28$, $p < 0.005$, all parts remained except knees and elbows: $t (19) = 2.62$, $p < 0.05$, only arms remained: $t (19) = 6.52$, $p < 0.001$. No other differences were found.

Portrayed Anger

![Bar chart showing observers' ratings for different remaining elements for portrayed anger, joy, sadness, and fear.](image)

a)
Portrayed Joy

- **Severity**:
  - Joy: 90%
  - Sadness: 20%
  - Anger: 0%
  - Fear: 0%

Portrayed Sadness

- **Severity**:
  - Joy: 0%
  - Sadness: 80%
  - Anger: 10%
  - Fear: 0%

**Portrayed Joy (Remaining elements)**

- Head & shoulders: 40%
- All except torso: 50%
- All except elbows & knees: 60%
- All except arms: 70%
- Arms: 80%

**Portrayed Sadness (Remaining elements)**

- Head & shoulders: 10%
- All except torso: 20%
- All except elbows & knees: 30%
- All except arms: 40%
- Arms: 50%
Figure 5.4. Observers mean ratings for presentations of the emotion response categories (a) anger, (b) joy, (c) sadness and (d) fear) of the portrayed and non-portrayed emotions organised by stimulus configurations. Error bars show standard errors.

Figure 5.4 (a-d) illustrates the full set of (correct) portrayed and (incorrect) non-portrayed results for each portrayed emotion in every condition. For each emotion category one-way ANOVAs and post hoc Tukey’s tests were carried out for the 5 subconfigurations and emotions separately. All of the portrayed emotions were found to be statistically higher than the corresponding non-portrayals except for 7 cases (2 for each emotion category except only 1 for anger). Therefore the disturbances in recognition appeared to vary fairly evenly across emotion categories but there were differences in the type of information that were important for each emotion. The head and shoulders combination caused confusions to be made for 3 of the emotion types: after the ANOVA for the 4 ratings of portrayed anger \( (F_{1.79, 34^*} = 16.04, p < 0.001) \) Tukey’s tests revealed that anger was confused with fear \((p > 0.05)\), after the ANOVA for joy \( (F_{3.57} = 5.97, p = 0.001) \) joy was mistaken for anger and after the same for sadness portrayals \( (F_{1.82, 34.56^*} = 12.26, p < 0.001) \) sadness was found to be confused with fear \((all p > 0.05)\). Following the one-way ANOVA for the 4 ratings of portrayed joy in the condition when only the arms were shown, \( (F_{1.55, 29.41^*} = 20.06, \)
ratings for joy were found not to differ from non-portrayed anger. After the ANOVA for fear portrayals in the same (only arms) condition \((F_{3,57} = 5.97, p = 0.001)\) fear was confused with sadness (both \(p > 0.05\)). In contrast, after the ANOVA for sadness when all of the elements remained except for the arms \((F_{2.08,39.6} = 10.03, p < 0.001)\), sadness was mistaken for fear (\(p > 0.05\)). Finally, after the ANOVA for the 4 ratings of fear when the torso elements were left out \((F_{3,57} = 8.52, p < 0.001)\) fear was confused with sadness (\(p > 0.05\)).

The aim of the current study was to determine how systematic removal of components in the displays affected the recognition of anger, joy, sadness and fear. The experiment showed that as few as 3 or 4 points were sufficient for correct recognition of some emotions. However, an uneven pattern of results was revealed in that in some conditions recognition was impaired for some emotions whereas perception of others, in the same condition, remained unaffected. Furthermore, for sadness portrayals performance was impaired compared to the full point-light condition regardless of which subconfiguration was displayed. For portrayed fear a similar result was found (except for when only the arms were missing where recognition remained equivalent to the full stimulus condition). Taken together, these results suggest that the contribution of different components of the body varied depending on which emotion was displayed; for sadness and fear perceptibility was significantly better when the full 13 point-light display was available. However, performance for full-figure displays of anger and joy was seldom enhanced compared to the corresponding subconfigurations. Having said this, in many instances even when perception was impaired to some extent emotions were still perceived veridically. Therefore, it seems that the visual system is adaptable in its ability to extract information about affect from body movement. Some of the information in a full point-light figure is redundant for the perception of affect.
5.8. General discussion

The present experiments were designed to investigate further the mechanisms that underlie the perception of affect and the stimulus information required by the mind for veridical processing. Experiment 5 investigated what happened when the motion was reversed. The results showed that observers were able to recognise affect despite the disruption of reversing the movements. Further, when the results were compared to those of Experiment 3 (in the condition when the same displays were played normally) it was demonstrated that there was no difference between backward-played and forward-played movement.

Observers have shown some sensitivity to reverse transformations of walking movements for humans (Fujimoto, 2003; Vanrie et al., 2004) and quadrupeds (Pavlova et al., 2002). However, perception was not severely impeded in these cases even though it seems important for a perceiver to know about the direction of a walker’s movements. The results of this experiment suggest that for the case of affect direction of movement is not an important cue.

Pavlova et al. (2002) have argued that limitations in the processing of biological motion are hierarchically nested. For example, these authors have argued that limitations connected to inversions are likely to be more powerful than those for position cues. Certainly, the results from Experiments 1 (and 1a, 1b) and 2 in combination with the results for this experiment support Pavlova et al.’s. (2002) claim. However, processing constraints may also vary depending on the type of motion, i.e. inanimate or animate objects or movements that specify social information.

Experiment 5 examined whether the visual system was sensitive to the emotions when they were played in reverse. But what information in the stimulus does the perceptual system extract to achieve veridical perception of emotions? The aim of Experiment 6 was to answer this question by determining whether the whole body or specific elements of the body were important for perceiving affect. In line with
expectations, the general finding was that there were characteristic features of the body that specify affect from biological motion. Interestingly, recognition varied as a function of the different configurations shown. Further, only a few point-lights were necessary for perception of most emotions.

The findings of Experiment 6 are partly consistent with those of Pinto and Shiffrar (1999) in that specific elements of the human body are important for event perception. Pinto and Shiffrar (1999, p.315) argued that the limbs are a ‘fundamental building block’ for veridical perception of human walking movement. However, this is not the case for the perception of affect as the removal of the interlimb joints (specifying dynamic limb symmetry), which were important features for walker recognition (Pinto and Shiffrar, 1999), made no difference to the recognition of affect. This is not surprising given that walking is regularly structured around momentum of the limbs and the expression of affect does not have this quality. Evidently, the perceptual system is sensitive to these qualitative differences of movement energy between physical actions and psychological states.

A fundamental feature of hot anger was movement specified by the arms: although anger could be perceived without the arms displayed, recognition was greatly impaired in this condition. In contrast, when only the 4 points that specify arm movement were displayed performance was equal to the full body condition. This feature was unique to anger (at least for the emotions tested) as for all of the other emotions when the arms were the only elements available, performance was either partially impaired as in the case of portrayed sadness, or it diminished completely. However, it cannot be the case that positioning of the arms relative to veridical positioning of the body is essential as anger was still easily perceived when the displays were inverted or rotated (see Chapter 1). The fact that there is very little structural information of the human form in the arms — they appear as 2 moving sticks — suggests that kinematic information is likely to be an important cue for anger.

In support of this claim, information from kinematics has been found to be fundamental in the processing of biological motion. For example, Hill and Pollick
(2000) found that observers were better at discriminating identities of actors from the actors' arm movements when time based information was exaggerated (i.e. the movements were segregated into individual parts and then the separate movement durations were increased by 50% and 100%). Importantly, spatial information remained constant in all of the conditions. Furthermore, in a manipulation where temporal information was negatively exaggerated (i.e. by -50%) discrimination of the actor was poorer (Hill and Pollick, 2000). For the specific case of affect, Pollick et al. (2001) asked observers to categorise arm-knocking movements of different affects (as if knocking on a door angrily, for example). A two dimensional model of affect was created from the results by using a multidimensional scaling procedure. Pollick et al. (2001) found that information from the physical movement of the arm correlated highly with the so-called ‘activation’ dimension of the model. This remained true even when form information was disrupted by scrambling the spatial information in the displays. Thus it appeared that observers’ responses were largely based on the kinematics of the actors’ movements.

There is further evidence to suggest that kinematics are important as (for point-light displays produced from actors’ more exaggerated portrayals of affect rather than normal portrayals of the same) there was a general trend for enhancement in recognition for the exaggerated expressions (Atkinson et al., 2004). However, results varied, as this was not true at all for the categories of sadness and joy. A caveat of this approach is that spatial information was not held constant in the exaggerated movement conditions. However, when the best exemplars of exaggerated movement were presented in a static full-light condition it was found that the exaggerations made little difference to perception. This result suggested that movement information was the underlying cue (Atkinson et al., 2004).

It may be that for slower (less movement activity) emotions such as sadness spatial cues could have more status than kinematic cues. For example, it was demonstrated that for sadness, recognition was impaired for all of the subconfiguration categories compared to the normal (13 point-light) displays. It is of interest that in Atkinson et al.’s (2004) study sadness was the only emotion that received significantly lower ratings when movements were exaggerated (by the actors). The most likely
explanation for this effect is that increases in movement are inconsistent with the perception of sadness. Taken together, it appears that sadness may be recognised with the most accuracy when the visual system uses form cues of the whole figure. Certainly, sadness can be easily recognised in static body positions — although, interestingly fear cannot (Coulson, 2004). In another study that illustrates the utility of spatial information, Bientema and Lappe (2002) have demonstrated that observers can recognise a human point-light figure without local image motion. This was achieved by moving every point-light individually to a randomly allocated position of the limb in the display (therefore no point carried the valid motion signal from one frame to the next frame). However, in this study a single frame was not enough for recognition, rather a sequence of frames was required. This effect was robust even when residual motion was controlled for by reducing the presentations to 4 frames a second. Therefore, it may be that spatial information from posture is important for recognition of some emotions but only when the spatial information evolves over time.

Recognition of biological motion could be accomplished by combining spatial and kinematic information. Specialised pathways have been postulated to exist in the visual system that process information from form and information from movement in quasi-independent streams (Milner and Goodale, 1995): the ventral stream, arises from primary visual cortex and projects to the inferotemporal cortex, and the dorsal stream, also arises from primary visual cortex but projecting to the superior parietal cortex (Ungerleider and Mishkin, 1982). It has been proposed that the ventral stream processes spatial information and the dorsal stream processes movement information. Giese and Poggio (2003) have argued that the visual system could encode patterns of body shapes via ‘snapshot’ neurons in the ventral processing stream and movement information could be processed in the sequences of complex optic flow patterns in the dorsal processing stream (see also Casile and Giese, 2005). Therefore, evidence of biological mechanisms in combination with recent computational models supports the utility of both form and movement information for the processing of biological motion and this is consistent with the findings here. Notwithstanding, a number of open questions remain, such as how the spatial and motion information and their relative weighting in the different emotions might be combined.
Portrayed joy was confused with anger when the fewest number of points were shown, i.e. 1) when only the head and shoulders were displayed, and 2) when only the arms were displayed. Sadness and fear were confused with each other in some conditions. Yet these anger/joy and sadness/fear confusions remained mutually exclusive (except for one occasion when anger was confused for fear) and similar confusions have been documented elsewhere in this thesis and by other authors (Dittrich et al. 1996). Anger and joy are expressed by faster movements than are sadness and fear. For example, Pollick et al. (2001), who have documented the velocities for various emotions when specified in actors’ arm movements, recorded average velocities of approximately 450 mm/sec for joy and 550 mm/sec for anger whereas fear and sadness had lower velocities — approximately 300 mm/sec each.

That these confusions occur in certain reduced cue situations could be explained by postulating a processing framework for the biological motion recognition. Massaro and Egan (1996) have proposed a model for bimodal processing of emotion called the ‘Fuzzy Logical Model of Perception’ where information is processed in three stages: evaluation, integration and decision making. Information is evaluated based on a prototype of a particular emotional expression. Secondly, information is integrated across the modalities. Finally a decision is made based on the amount of support for each alternative. A similar framework could explain unimodal processing where information from various cues (e.g. from form and optic flow) is first evaluated, and then integrated. Finally, a decision is made at a higher level of the processing hierarchy. As demonstrated by the results here, when cues are degraded to make recognition difficult (as shown in the confusions) incorrect decisions are often made by the visual system as it is operating with insufficient information. For example, in the case of portrayed joy when only the arms were shown, the visual system may have been able only to detect that the emotion was high in activation (i.e. anger or joy) at the evaluation stage with this limited information. However, the percept of joy can only be recognised as ‘joy’ when more information is available to be combined at the integration stage, after which a correct decision can be made.
The results of Experiment 6 show that the visual system does not use the same constituent parts of the stimulus in all emotions. The visual system appears to be specifically attuned to particular body configurations for veridical perception of specific emotions.

There are limitations to this study in that only 4 emotions were examined here. Likewise, the number and type of body component combinations considered were by no means exhaustive. For example, there were no manipulations that examined movement from only the lower body — it might prove interesting to see if lower body movements, i.e. of the hips, are indicative of fear. For instance, Roether and Giese (2005) have shown that point-light displays of the lower half of the body can specify fear.

In summary, the experiments reported in this chapter demonstrate that backward-played displays do not disrupt perception of affect. Furthermore, the whole body is not necessary for veridical perception of affect, showing that there is some redundancy in biological motion displays. This is not surprising given that in the natural environment people are often partially occluded by other things (Pinto and Shiffrar, 1999). Of interest is the observation that there is considerable variation in the features needed for recognition, depending on emotion type. Form cues appear to dominate for some emotions and movement cues appear to dominate for others. Certainly the experiments reported in this chapter show that the visual processing of biological motion is complex but both robust and flexible.
Chapter 6

Children’s perception of emotion from biological motion

6.1. Introduction

So far it has been established that perception of emotion from biological movement portrayed in the context of interpersonal dialogue can be perceived by adults, is affected by the social context of the displays and, with only minor exceptions, is robust despite systematic degradation of the stimulus. This chapter explores how children recognise emotion from biological motion as veridical perception of socially relevant information is fundamental to the development of interpersonal communication.

The ability to perceive point-light displays develops early in infancy. For example, Fox and McDaniel (1982) used a preferential looking task to demonstrate that 4-month-old infants preferred to look at a point-light walker to dynamic noise. Infants can discriminate non-human movement too as in a habituation task 6-month-olds could categorize vehicles and animals (Arterberry and Bornstein, 2002). Furthermore, 5-month-old infants can discriminate a walker from similar stimuli and yet cannot discriminate the same stimuli when they are turned upside-down (Bertenthal, Proffitt and Kramer, 1987). This finding is important as both images had the same low-level visual properties such as velocity of the dots in the display, and so the preference was not for some primitive aspect of the images.

Clearly infant research has some limitations as it cannot be known what the infant perceives in the displays but children as young as 3 are able to attribute motion verbs such as ‘walking’ and ‘dancing’ to point-light displays of walkers and dancers respectively (Golinkoff, Chung, Hirsh-Pasek, Liu, Bertenthal, Brand, Maguire and Hennon, 2002) and can recognise and name point-light displays of a walking man, walking and running dogs, and a flying bird (Pavlova, Krägeloh-Mann, Sokolov and Birbaumer, 2001). Static versions of the same displays revealed poor performances by the children and therefore showed that movement information was vital for
recognition and by the age of 5 the children were almost at ceiling levels of performance (Pavlova et al., 2001). Interestingly, these researchers found that 3 and 4 year-olds were able to identify a walking point-light dog better than similar displays of a human walker (although this effect was not apparent for a running dog). This result is surprising given that it is more biologically relevant to perceive human biological motion than canine motion. In another developmental study Berry (1991) demonstrated that 5-year-olds could discriminate sex correctly from point-light displays of actors’ faces but only when the actors were engaged in conversation with another (unseen) person. In contrast, in another condition where there was no social interaction and the actors simply recited the alphabet the children’s judgements were at chance performance. This implies that children are sensitive to information in quotidian conversation but are not sensitive to information that has no social relevance.

Our knowledge of what children can perceive in point-light displays has been largely confined to the identification of physical objects and activities and there has been very little research into whether children can identify internal states such as emotion from these unusual displays. In one rare study, Moore, Hobson and Lee (1997) presented 7-year-olds with point-light displays of a single actor engaged in actions to represent emotions (surprise, sadness, fear, anger or joy), for example, jumping and skipping around to depict joy. In 21 out of the 65 verbal responses made, the children were able to identify the emotions. However, the developmental trajectory for acquiring such skills and whether all the emotions were recognised equally well were not examined. The correct responses were made for less than half of the stimuli shown which suggests that the skill to identify affect was not fully developed at this age; although this result may have been due to task difficulty as the children were required to describe verbally what they saw in each display. Therefore, the children might have recognised affect in some displays but may not have been able to verbalise their answers correctly.

There is indeed some evidence that children are sensitive to the emotion conveyed within such stimuli as pictures of faces (e.g. Izard, 1971; Walden and Field, 1982), voice prosody (Stifter and Fox, 1987), music (Cunningham and Sterling, 1988;
Boone and Cunningham, 2001) and static body postures (Tracy, Robins, and Lagattuta, 2005). In a study that investigated the particular case of body movements, Van Meel et al. (1993) tested whether 5-, 8-, 10-, and 12-year-olds could recognize affect (anger, joy, sadness and fear) when it was portrayed in dance movements (facial information was not occluded). These authors argued that, irrespective of whether the dependent measure was a free response or a multiple choice response format, children under 8-years were not able to recognize emotion. However, even the 8-year-olds appeared to perform quite poorly (the authors did not compare the accuracy scores to chance) although their performances were better than those of the 5-year-olds. Differences between individual affects were not examined.

Boone and Cunningham (1998) also investigated the developmental trajectory for the acquisition of this skill when portrayed in dancers’ body movements (facial information was blurred). In contrast to the results of Van Meel et al. (1993) Boone and Cunningham (1998) found that children as young as 4 were able to identify sadness from full-light displays of expressive body movements in dance (even though the facial information was blurred and faded and therefore could not be used as a cue). By 5 years of age, fear and happiness could also be identified. The disparity between the results of the Boone and Cunningham (1998) study and the Van Meel et al. (1993) study is puzzling and cannot be accounted for in terms of cues from facial affect given that the study where performance was particularly poor was also the study where facial information remained. It is possible that differences in the dance samples between the studies could account for the differences in results.

In adults anger was found to be the most easily identified emotion (see Chapter 3) and young children can perceive anger when full-light information from the face and body is available as 4- and 5 year-olds do recognize anger from nonverbal and nonstylized adult portrayals (Cummings, Vogel, Cummings and El-Sheikh, 1989). Furthermore, children as young as 3 can identify anger both from faces and voice prosody (Stifter and Fox, 1987). Therefore, it was of particular interest as to whether young children would identify this emotion when expressed in naturalistic portrayals of social interaction rather than stylized dance movements, of which they are likely to have little experience.
How might such skills develop? One hypothesised mechanism is a link between motoric production and perception. According to this view it is necessary for a child to be able to produce the specific motor actions involved in a cue in order to perceive it (Lockman, 1990; Bushnell and Boudreau, 1993). Boone and Cunningham (1998) argued that young children were unable to perceive anger because they were unable to produce the complex directional changes in the face and torso that are claimed to be associated with this emotion (e.g. de Meijer, 1989). However, this hypothesis was not directly tested.

The aim of the present study was to assess whether children could identify emotion from more natural communication behaviours than dance (Boone and Cunningham, 1998) and more natural types of movement than that expressed by single actors behaving stereotypically (Moore et al., 1997). Another aim of this research was to track the developmental trajectory for the acquisition of this skill. From an ecological perspective and taken with Berry’s (1991) finding that children detected information about the sex of an actor only from natural patterns of movement characterised in interacting faces, the stimuli used in the displays developed here should give children the best possible chance of successful identification of their emotional content.

In this study the youngest group were 5-year-olds — the youngest age that most of the emotion categories expressed in videos of dance movements could be identified (Boone and Cunningham, 1998), and it was reasoned that with more impoverished stimuli younger children than this would be unlikely to perform well here. Seven- and 9-year-olds were included to understand the development of this skill through mid-childhood. Anger, happiness, sadness and fear were the emotion categories selected on the basis of their being successfully identified in previous studies (e.g. Boone and Cunningham, 1998).

Specifically, the aim of this study was to enquire (i) whether children can perceive emotion from point-light displays, (ii) whether this ability improves with age, and (iii) whether the ability to identify emotion develops at different rates for the
different emotions. A control group of adults was also tested to establish the maximum level of performance that could be expected in the task.

6.2. Method

It was clear that the method employed in the experimental design for recording the adult subject's responses was likely to be too complex for young children to understand; therefore a design was needed that would be appropriate for the children to use. Smiley and Huttenlocher (1989) found that young children gave the best performance when using a binary choice technique over open-ended verbal responses or choosing from among verbal alternatives when given an emotion identification task. Other researchers (e.g. Boone and Cunningham, 1998) have also used this method with success. Therefore, a non-verbal binary choice task was implemented in the present experiment. This method involves showing two displays simultaneously and asking the subject to pick the correct one. To show the dyadic displays used for the other studies reported in this thesis was deemed as being too complicated and therefore the displays were altered so that single actors were shown expressing an emotion whilst engaged in discourse with another (unseen) actor. A group of adults was tested prior to the children for two reasons: first, to ensure that the method worked successfully and second, to use as a control group for the children.
6.2.1. Subjects

A total of 68 observers from four age groups were recruited. There were 14 adults (8 women, 6 men, mean age: 25 years 1 month, range: 19 – 36 years); 18 five-year-olds (8 boys, 10 girls, mean age: 5 years 6 months, range: 5 years 5 months – 5 years 11 months); 18 seven-year-olds (9 boys, 9 girls, mean age: 7 years exactly, range: 6 years 10 months – 7 years 2 months); and 18 nine-year-olds (10 boys, 8 girls, mean age: 9 years exactly, range: 8 years 8 months – 9 years 3 months).

6.2.2. Stimulus materials and apparatus.

The stimuli selected for the experiments contained 3 different recordings for each of the 4 emotion categories — anger, joy, sadness and fear. The stimuli used here were chosen as they were found to be good exemplars (i.e. they received high ratings) in the previous studies on adults reported in this thesis. As before, they also comprised as many different actor combinations as possible and at least one performance from each of the 10 actors was used.

Each of the recordings was split into separate stimuli containing just one of the two actors. These individual stimuli were then presented in mixed pairs, with, as an example, one actor portraying anger and the other portraying fear. When viewing such a mixed pair, a child could be asked to identify “who’s angry” or “who’s scared”. To make it clear that the two actors were not directly interacting with each other in a mixed pair a solid vertical line was shown between them as if they were in different rooms. The first 10 seconds of each recording were used.

During the experiments the point-light displays were shown on an Apple iBook computer with a 14-inch LCD display. The point-lights were shown in green on a black background. The children sat approximately 40cm from the screen. Figure 6.1 shows an example of the stimuli shown to the children.
Figure 6.1. Static frame of a point-light display showing the two actors where each actor is represented by 13 points of light. In this frame the actor on the left of the screen is expressing anger and the actor on the right is expressing joy.

6.2.3. Design and procedure.

Children were seated individually in front of the computer monitor and the experimenter was seated next to the child. A flat board, with a vertical line drawn centrally to follow the central line that appeared on the monitor, covered the computer keys and a small toy cat was placed on the centre of the board in front of the child. The children were asked to put the cat to the side of the screen that showed the point-light display of the figure that was expressing the target emotion (e.g. "who’s happy?"). After every response the toy was returned to the centre.
Familial emotion terms were used; specifically, cross was used as well as angry for anger target displays, and scared and frightened were used for fear. All of the children were asked if they understood all of the familiar emotions terms used in the study (e.g. “what do you do when you are scared?”) and, without exception, the children responded appropriately. All of the children responded to every item and they were allowed to see the presentations more than once if they needed to. This occurred approximately 10% of the time for the younger groups and rarely for the oldest group. Only on two occasions did a child ask to see a display more than twice and on no occasion did a child ask to see a display four times or more. The experiment lasted for 10 minutes for each observer. Three examples of each target emotion were shown with foils from each of the other emotion categories used. There were 24 trials in all (four target emotions x three examples of each target emotion x two repetitions of each so that every example would appear once on the left side and once on the right side of the central line in the display). This procedure allowed both halves of each original dyad to be used once in the experiment.

A dimly lit and quiet area of the school was used to test each child individually. The children were told that they would see two people on the screen but that they were people made of dots. An example of one of the displays was shown and the children were asked to put the toy on the side of the screen where the tallest person was. This task was always completed successfully and so demonstrated that the children could understand similar instructions to that of the experimental task. The children were told that they would be asked a question as to the whereabouts of the happy person and so on. Once again, a clear indication that they understood these terms and instructions was always given.

The adults were tested in a dimly lit laboratory at the University of Surrey. They were tested in a similar way to the children but were asked to verbalise ‘left’ or ‘right’ to refer to the side of the screen where they judged the target emotion to be displayed. The experimenter used the terms happy, sad, fearful and angry to express the target emotions that the adult observers were to locate.
6.3. Results

For each individual, an accuracy score for correct identification of the target emotions was calculated where a score of 1 was given for each correct choice, therefore a range from 0 – 24 was created, with a mean score of 12 expected by chance. The average accuracy scores for the four groups were 21.6, 19.2, 17.4 and 15 for the adults, 9-year-olds, 7-year-olds and 5-year-olds respectively. All of these scores were shown to be significantly above chance levels when a one-sample t test was used; adults, \( t(13) = 17.20, p < 0.001 \); 9-year olds, \( t(17) = 9.23, p < 0.001 \); 7-year-olds, \( t(17) = 6.56, p < 0.001 \); 5-year olds, \( t(17) = 4.00, p < 0.005 \). Therefore, at the most general level, the younger children were able to extract at least some information about the emotional content of the actors’ portrayals from biological motion. Some of the required discriminations were sufficiently challenging for adult performance to be high, yet below ceiling — as confirmed when adult scores were compared to a score of 6 in a one-sample t test \( (r(13) = 4.24, p < 0.005) \).

To ask whether the developmental trend was similar for each of the four target emotions, the separate scores for each emotion category were calculated for each participant. This amounted to four scores for each observer with a maximum of six and a mean score of three to be expected by chance. Figure 6.2 shows how recognition improved with age for each of the four emotions separately. The rate of development and level of performance was uneven and differed between the target emotions although in general it appears that monotonic improvement occurred as age increased. One-sample t tests were used to compare actual performance with that expected by chance. The only group-emotion combination not above chance \((p > 0.05)\) was for the 5-year-olds when the target emotion was sadness.
Figure 6.2. Accuracy of emotion identification as a function of age and emotion category. Error bars show standard errors. The results above each bar show results of one-sample t tests (df = 13 for the adults and 17 for the children) that compared actual performance with that expected by chance.

Trend analysis was performed on the data to investigate the shape of the function for perceptual development across age for each emotion. This showed a positive linear trend for each of the slopes that related accuracy to age for each of the emotions (anger, \( t(64) = 4.83, p < 0.001 \); joy, \( t(64) = 3.49, p < 0.001 \); sadness, \( t(64) = 4.85, p < 0.001 \); fear, \( t(64) = 3.75, p < 0.001 \)). This confirmed that recognition of all four emotions improved with age. A negative quadratic trend was also evident for the anger data (\( t(64) = 2.31, p < 0.05 \)). This indicated that perceptual ability for the recognition of anger developed rapidly after the age of 5 (when performance was just above chance) and was at adult levels by age 7 (\( t(30) = 0.16, p > 0.05 \)). No other quadratic trends were detected for the other emotions showing a more even developmental trajectory (\( p > 0.05 \)).
The data also allowed an enquiry to be made as to whether performance in identifying a particular emotion varied depending on the foil emotion. In order to investigate this, 12 separate scores were produced for each participant where the maximum score was 2 and a score of 1 was to be expected by chance. The full set of scores for each correct emotion category broken down by emotion category of the accompanying foil can be seen in Table 6.1. and Figures 6.3 to 6.6 show the performances for each of the emotions for each of the age groups separately.

Table 6.1. Observers' mean scores for each correct emotion category broken down by emotion category of the accompanying foil. Standard deviations are shown in parentheses.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Correct/Foil</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Anger/Joy</td>
<td>Anger/Sad</td>
<td>Anger/Fear</td>
</tr>
<tr>
<td>5-year olds</td>
<td>0.95 (0.5)</td>
<td>1.4 (0.6)</td>
<td>1.5 (0.7)</td>
</tr>
<tr>
<td>7-year olds</td>
<td>1.4 (0.5)</td>
<td>1.6 (0.6)</td>
<td>1.8 (0.5)</td>
</tr>
<tr>
<td>9-year olds</td>
<td>1.5 (0.5)</td>
<td>1.8 (0.4)</td>
<td>2.0 (0)</td>
</tr>
<tr>
<td>Adults</td>
<td>1.6 (0.5)</td>
<td>1.8 (0.4)</td>
<td>1.9 (0.3)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Joy/Anger</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5-year olds</td>
<td>1.6 (0.5)</td>
<td>1.4 (0.6)</td>
<td>1.1 (0.8)</td>
</tr>
<tr>
<td>7-year olds</td>
<td>1.3 (0.7)</td>
<td>1.4 (0.7)</td>
<td>1.6 (0.6)</td>
</tr>
<tr>
<td>9-year olds</td>
<td>1.3 (0.8)</td>
<td>1.8 (0.5)</td>
<td>1.6 (0.5)</td>
</tr>
<tr>
<td>Adults</td>
<td>1.6 (0.5)</td>
<td>2.0 (0)</td>
<td>1.8 (0.3)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Sad/Anger</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5-year olds</td>
<td>1.16 (0.6)</td>
<td>1.2 (0.5)</td>
<td>0.95 (0.7)</td>
</tr>
<tr>
<td>7-year olds</td>
<td>1.5 (0.7)</td>
<td>1.5 (0.6)</td>
<td>1.2 (0.6)</td>
</tr>
<tr>
<td>9-year olds</td>
<td>1.7 (0.6)</td>
<td>1.7 (0.4)</td>
<td>1.3 (0.8)</td>
</tr>
<tr>
<td>Adults</td>
<td>2.0 (0)</td>
<td>2.0 (0)</td>
<td>1.5 (0.6)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Fear/Anger</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5-year olds</td>
<td>1.4 (0.5)</td>
<td>1.3 (0.7)</td>
<td>1.0 (0.8)</td>
</tr>
<tr>
<td>7-year olds</td>
<td>1.4 (0.8)</td>
<td>1.6 (0.6)</td>
<td>0.8 (0.8)</td>
</tr>
<tr>
<td>9-year olds</td>
<td>1.7 (0.7)</td>
<td>1.7 (0.6)</td>
<td>0.8 (0.7)</td>
</tr>
<tr>
<td>Adults</td>
<td>1.8 (0.4)</td>
<td>1.7 (0.5)</td>
<td>1.6 (0.5)</td>
</tr>
</tbody>
</table>
Figure 6.3. shows performance at identifying anger for each age group. When the foil was sadness or fear 5-year-olds were above chance and there was improvement in performance with age. On the other hand, when the foil was joy 5-year-olds performed at chance levels (p > 0.05) but the older age groups were substantially better (p < 0.005). Paradoxically, Figure 6.4 shows that when joy was the target emotion and anger was the foil 5-year-olds performed well above chance (p < 0.001) but 7- and 9-year-olds performed at chance levels (p > 0.05).
Figure 6.4. Mean accuracy of joy recognition as a function of age broken down by foil type. Error bars represent SEM. The figures above each bar show results of one-sample $t$ tests that compared actual performance with that expected by chance.

Figure 6.5. Mean accuracy of sadness recognition as a function of age broken down by foil type. The figures above each bar show results of one-sample $t$ tests that compared actual performance with that expected by chance.
Figure 6.5 shows that children were poor at picking sadness from fear foils and performance for them all was at chance \((p > 0.05)\). This was also the most difficult judgement for adults, although performance for them was above chance \((p < 0.05)\). When the foils were anger or joy fairly even developmental trends are evident but the 5-year-olds did not perform above chance \((p > 0.05)\). When fear was the target emotion (Figure 6.6), none of the children could distinguish fear if sadness was the foil \((p > 0.05)\). Adults performed well above chance on this task \((p < 0.001)\). Therefore, for the special case of fear and sadness combinations, irrespective of whether fear or sadness was being judged, the skill for perceiving them correctly was not evident until after the age of 9 but before adulthood.

![Fear Recognition as a Function of Foil Type](image)

**Figure 6.6.** Mean accuracy of fear recognition as a function of age broken down by foil type. The figures above each bar show results of one-sample \(t\) tests that compared actual performance with that expected by chance.
6.4. Discussion

This study has shown that young children have the ability to identify affect from biological motion. This result both extends and bolsters the previous findings in this field which shows that young children can identify not only concrete entities such as walkers or dogs (Pavlova et al., 2001) but also abstract constructs such as emotions (Moore et al., 1997).

Although 5-year-olds recognised affect from biological motion, they were unable to recognise sadness. There was a trend for general linear improvement as age increased across all emotion categories but there was variation within each category and age group depending on the type of foil presented. This shows that although recognition improves with age, development of this skill is not a uniform trend.

That improvement in performance with age differed between the emotions is interesting as this uneven pattern of development may inform us about the mechanisms underlying the perception of emotion from biological motion. If recognition depended upon the operation of a single mechanism, devoted to the analysis of biological motion or a general motor skill, then an even developmental picture would have emerged than the one observed. Boone and Cunningham (1998) argued that perception of a particular emotion was related to being able to produce a specific motor action. However, there is no direct evidence of such a link and so it cannot be discounted that the acquisition of this skill is related to visual experience. Experience makes a difference to perception. For example, physically abused children who might well be exposed to anger as an environmental cue on a regular basis accurately identify facial displays of anger on the basis of less sensory input than do controls (Pollak and Sinha, 2002). Importantly, the abused children were not better at perceiving other emotions than were controls. Further, children identify anger and joy when expressed through vocal and facial channels by their mothers more readily than when expressed by strangers (Shackman and Pollak, 2005). For the specific case of identifying point-light displays Pavlova et al., (2001)
argued the reason young children were better able to identify walking quadrupeds over walking humans was because they had little experience of humans from the particular view shown, but as young children are short (as are dogs) they had no trouble in recognising the walking dogs. Motor-skill theories and experience-based theories need not be mutually exclusive (Jacobs, Pinto and Shiffrar, 2004; Loula, Prasad, Harber and Shiffrar, 2005).

On the other hand, if recognition of emotion from biological motion is thought of as supported by a collection of cues picked up by different mechanisms then an uneven developmental trajectory would be likely to emerge. The idea of a collection of mechanisms for picking up a range of cues is supported by other studies reported in this thesis. When adults were presented with the original stimuli of the dyads that had been degraded in various ways the effects tended to be specific to the emotion being judged and the manipulation of the point-light display that was made. For example, inversion selectively impeded the recognition of fear, whereas presenting single actors rather than natural actor pairs selectively impaired the recognition of love but enhanced sensitivity to sadness (see Chapter 3). Artificial pairs created by presenting a single actor and the mirror image of this had further specific influences on judgements.

Substantial evidence supports the notion that separate systems exist to serve different emotions. For instance, in a recent review of the literature Lawrence, Murphy and Calder (2004) argued that neuroscientific evidence shows fear and disgust to be represented in the brain in regionally specialised places, as double dissociations — where normal results for task A but abnormal results for task B are found in one patient and the opposite results are found in the other patient (task B normal, task A abnormal) (Shallice, 1988) — have been found between fear and disgust in brain damaged humans. Thus, patients with amygdala damage show impaired recognition of fear expressions but not disgust (e.g. Sprengelmeyer, Young, Schroeder, Grossenbacher, Federlein, Büttner, and Przuntek, 1999) whereas patients with insula/operculum damage exhibit the opposite behaviours (e.g. Calder, Keane, Manes, Antoun and Young, 2000).
In adults (see Chapter 3) anger was the portrayed emotion most easily recognised even when presented non-canonically. However, the data reported here show that 5-year-olds were only just above chance at recognising anger overall but performance varied depending on the type of foil that was presented. The ability to recognise anger improved rapidly after the age of 5, particularly when the foil was joy. For bodily expression of emotion (full-light) in dance movements Boone and Cunningham (1998) found that sensitivity to anger began to emerge at around the age of 5 and this is largely in line with the results reported here. However, for voice prosody (Hortaçsu and Ekinçi, 1992) and facial expression (Cummings, Vogel, Cummings and El-Sheikh, 1989; Sullivan, Kirkpatrick and MacDonald, 1995) anger could be recognised by children as young as 5 and so it may be that the perception of anger from the movement of the body develops later than it does via these other channels of expression.

Of interest was the result that 5-year-olds could pick out joy when anger was the foil and yet older children were unable to. Paradoxically, 5-year-olds could not recognise anger when joy was the foil and yet older children could. First, this result shows that younger children were not using logical reasoning to decide which of the two stimuli was the correct answer (i.e., when being asked to find anger, knowledge of what joy looks like combined with seeing a joy stimulus means that anger must be the other stimulus). Second, this pattern of results suggests that as a child becomes attuned to one stimulus, sensitivity to other previously mastered stimuli may suffer. Boone and Cunningham (1998) conducted a cue intensity task, i.e. the upward movements which are thought to be a critical cue for joy (de Meijer, 1989) were manipulated so that there were either more or less upward movements than in the original stimulus, and found that 5-year-olds were more reliant than 8-year-olds in the use of the upward arm movement cue in decoding happiness. This could be indicative of children becoming less sensitive or giving less weighting (see Chapter 5) to this cue as they become attuned to other relevant stimulus cues (Boone and Cunningham, 1998). Further, this type of phenomenon could explain the anomalous result of non-monotonic development in the Boone and Cunningham (1998) study as 8-year-olds were better than adults at recognising sadness. Therefore, adults too may be susceptible to a loss of sensitivity in previously mastered cues in order to become
My results show that ability to identify emotions largely improves throughout mid-childhood although there are some pockets of non-monotonic development. These results are largely in line with the findings of Boone and Cunningham (1998) although there are some differences in results, for example, those researchers found sadness to be the first emotion recognised; here it was the last. This study reports that joy was the most easily identified emotion by 5-year-olds thus paralleling the findings that positive emotions are the first to be recognised developmentally in music (Cunningham and Sterling, 1988) and in facial expressions (e.g. Philippot and Feldman, 1990; Walden and Field, 1982). That full-light presentations of dancers were used as stimuli by Boone and Cunningham poses a problem for direct comparisons as it is not known whether the contrasting results between studies are due to fundamental differences between dance movements and actions made in interpersonal communication or differences between full-light and point-light conditions. Notwithstanding, the general finding is true for both studies that the skill to perceive affect from biological motion emerges at or by around 4-5 years and then in most cases matures to almost adult levels by the age of 8 or 9.

Overall, these findings show that young children have the ability to use the isolated movement cues that are conveyed in point-light displays to identify the emotions of others. Emotion identification from point-light displays improves with age largely monotonically but unevenly until late in mid-childhood when near-adult performance is achieved.
Chapter 7
Discussion

7.1. Overview

Four interdependent aims were set in Chapter 1 that relate to emotion and biological motion. These were: (i) to see how the social context is important for the perception of affect from biological motion; (ii) to further our knowledge about the mechanisms responsible for this ability; (iii) to see how the skill to read emotion from body movements develops in childhood and (iv) to further elucidate what is known about emotions. The principle findings relating to the first three questions were discussed according to their relevance in the empirical chapters. This chapter has two main purposes; the first is to answer the final question about how the findings contribute to what is known about affect, the second is to discuss the methods used in this (and other) research and to suggest ideas for future experiments. First the main findings are summarised.

7.2. Summary of the main findings

In Chapter 1, the existing studies that have investigated the perception of affect from biological motion were critically reviewed. It was argued that it would be useful to look at the subject from a social angle and to produce stimuli which captured more natural communication behaviours than those used previously. Methods were discussed in Chapter 2. The work in Chapter 3, using both quantitative and qualitative measures, assessed the utility of these new behavioural samples. The results showed that anger, joy, sadness, love and fear could be distinguished in normally orientated point-light displays. However, disgust was not recognised in either the qualitative or the quantitative task. Follow up experiments established that neither the varying duration of the vignettes nor the amount of turn taking within a display affected observers’ ability to recognise emotions. Effects of orientation were also explored; inversion was found to reduce the salience of the displayed emotions
to varying extents although performance was still good for most emotions except fear. Overall, these findings are consistent with the view that social information is specified in the spatiotemporal dynamics of point-light displays. The results show that the body provides an excellent channel (for most emotions) by which information can be picked up from a distance and is not limited to communication by a small number of actors as oro-facial communication would be.

The experimental designs in Chapter 3 were similar to those conducted in other research laboratories. However, the work in Chapter 4 went beyond what has previously been investigated. Two important issues relating to the perception of a social event were explored. As a first step towards understanding if, and if so how, the social context augments the perception of affect, in Experiment 3 subjects were shown upright versions of the original dyads, a single actor (monads) and a dyad comprising a single actor and his/her mirror image (reflected dyads). In general, recognition was found to vary between the emotion categories depending on the context in which the emotion was shown. In particular, the removal of one actor seemed to impair the recognition of joy and love; these are emotions that are particularly expressive socially. There were other impairments as a result of the non-veridical contexts (as in the reflected dyads). This suggests that observers take into account information from the entire scene when judging the emotional content of a situation.

These findings show that the properties of body movement alone cannot provide a full explanation for the recognition of affect and that the social context is a defining feature for the recognition of affect from biological motion. There are emergent properties within a scene that do not occur as a result of processing the kinematics in the display. This was well illustrated by the fact that portrayals of romantic love were unrecognisable (as love) without the presence of a second actor despite the fact that the stimulus properties for each actor were no different between monad and dyad conditions. Clearly, the demonstration of the social influence on the perception of affect in this study augments the view that it is important to consider emotions at a social level.
Experiment 4 also showed that there are qualities inherent within the social exchange that advise the third person observer of important communicative characteristics. Here it was demonstrated that a speaker could be identified even if the only information available is a point-light display of two people engaged in conversation, where importantly, facial information was not available. Therefore, speech and body movement are connected. This confirms suggestions that speech and body movements are intimately related (see Chapter 1).

In Chapters 5 and 6 several studies were described which further investigated the mechanisms which are involved in the perception of affect. During locomotion, for purposes of efficiency, the limbs move in a regular, predictable way but emotions are not expressed by such regular (pendulum like) movements of the limbs. The experiments in Chapter 5 replicated procedures that had been conducted to investigate locomotory actions to assess whether there are differences between the way that affect and locomotion are perceived. In the first study (Experiment 5) the information contained in the display was reversed to see if this made a difference for recognition (compared to normally played displays). The results showed that observers were able to recognise the displays despite the backward-played motion. The second study (Experiment 6) critically examined what information in the stimulus was required to perceive emotion. This was achieved by occluding some of the information in the displays. The results showed that as few as 3 or 4 points were sufficient for recognition; this demonstrates that there is considerable redundancy in a point-light display of affect.

Taken together, the results of these studies showed that emotion is processed differently to locomotion. This claim is made because in these results, displays played backwards made no difference in recognition compared to displays played forwards whereas for locomotion, performance is impaired when the motion is reversed. Further, the contribution of different components of the body varied depending on which emotion was displayed and these components are not the same as those that are important for the perception of locomotion.
In Chapter 6 a developmental study was documented where children completed a binary choice task to see if they could perceive emotion from point-light displays. It was demonstrated that children as young as 5 are able to perceive affect from biological motion. This ability largely improves throughout mid-childhood then in most cases matures to almost adult levels by the age of 8 or 9.

### 7.3. Implications for the study of emotion

The findings reported in this thesis contribute to answering the question of whether emotions should be conceived in terms of discrete categories or as varying along certain bipolar dimensions such as pleasantness and activation (see Chapter 1). The results are difficult to reconcile with a dimensional account. Rather, they are consistent with an approach whereby emotions should be thought of in terms of qualitatively different categories possibly related to functional specialization in the brain.

This claim is made as dissociations were found between different emotions depending on the way the stimuli were manipulated — the finding of dissociations suggests the existence of an underlying system of functional specialization (Lawrence et al. 2004). For instance, there were big differences in the way that fear was recognised when the stimuli were rotated and inverted compared to the other emotions tested. The results reported in Chapter 4 took this point (of dissociations) further in that the social context affected the recognition of some categories of emotion over others. Furthermore, in Chapter 5 it was shown that different elements of the images were critical for the recognition of different emotions. Moreover, in development, children do not acquire the ability to perceive affect for every emotion at the same rate, which is again supportive of categorical structures.

Therefore, it is suggested that the data reported in this thesis are easier to reconcile with a categorical account rather than thinking of emotions in terms of dimensions.
Other evidence suggests that emotion categorisation is evident from 5-months of age as infants, in a task where they are habituated to seeing different smiles worn by different people, prefer to look at a new emotion category expression (fear) worn by a new person over a never-seen-before smile (see Bornstein and Arterberry, 2003). For adults, support in favour of categorisation comes from a study where adults have been shown to categorise emotional expressions that are conveyed via the vocal channel (Laukka, 2005; for similar effects on facial expressions of affect see Young, Rowland, Calder, Etcoff, Seth and Perrett, 1997).

However, it would be precipitous to conclude that the structure of affect can only be thought of in terms of categorisation. The findings reported in this thesis are also partly consistent with dimensional theories (at least for an account that specifies a pleasantness dimension) as in Experiment 3 both joy and love (positive emotions) were affected by the social context. However, despite the fact that both sadness and fear are considered as displeasurable it was only the percept of fear that was destroyed in the experiments that investigated inversion. A theory that posits pleasantness as a dimension would predict that these emotions would not be dissociated.

Some authors argue that there is room for both dimensional and categorical accounts to describe affect but with superior emphasis on one account over the other. For example, Russell and Barrett (1999) argue that dimensions are fundamental (but acknowledge that the search for categories will continue) whereas Lawrence et al. (2004, p-163) postulate that dimensions such as activation and valence may be important components of emotion but that these should not be considered as 'emotion primitives' (whereas categories can be). These authors suggest that it is likely that the dimensional aspects are processed at higher levels such as the frontal (Rolls, 1999) or prefrontal cortex (Davidson, 1992).

It has been discussed how the data fit into models that endeavour to explain the structure of affect. However, the data also provide evidence for modular and not just categorical processing (see Chapter 1) as inverted fear stimuli were unrecognisable (see Chapter 3). This result is suggestive of a perceptual module for fear. Also, the
processing of concrete activities from biological motion has been shown to be IQ independent, which is indicative of modularity (Moore, Hobson and Anderson, 1995). My data for fear inversions live up to Fodor’s (1983) requirements in that processing is mandatory and encapsulated from knowledge. However, Fodor argues that modules should be innate. The evidence in Chapter 6 — that the skill to recognise affect from biological motion develops at around age 5 — does not support Fodor’s claim for innate modules (unless they are latent in early childhood). Of course, it might be possible that the dependent variable used here was not sensitive enough to show that young children can easily perceive affect from biological motion. Indeed, the youngest children tested were able to recognise joy easily. There have been few developmental studies to investigate the perception of affect from biological motion and so future research in this area is encouraged (see 7.4 for future experiments).

That there is evidence for modules refutes Gibson’s (1950; 1979) argument (see Chapter 1) that all the information required for vision is contained within the stimulus. However, in line with Gibsonian theory, the evidence of modular processing does suggest that information is processed in a bottom-up rather than a top-down way. In line with Gibson’s ideas, the natural environment is a good starting point for our enquiries. The results from Experiment 4 show the importance of considering the social environment in that the same expressive behaviours signal different messages depending on the social context. Thus, it is not enough to consider only the intrapersonal aspects of emotion; interpersonal aspects are equally important.
7.4 Validity, limitations and future research

Validity

This research was conducted to investigate the perception of affect by using more natural stimuli than have been used in the past. Overall, the general results verify those of other workers in the field in that affect can be recognised from body movement. Emotions are complex and cover such a broad class of phenomena that they may comprise more than one entity (Russell and Barrett, 1999). Indeed, the problem with emotion is that we do not really know how to define it; for example, 'anger' covers a range of behaviour — from being slightly annoyed about picking up the wrong milk in the supermarket to the act of murder by a frenzied mob. Some might argue that until the operational variables are known there is no point in investigating a construct. However, it would be imprudent not to investigate something that pervades our everyday lives. Therefore, in the meantime — before the operational definitions are decided upon — the best way forward is to conduct many studies looking at emotion from different angles in order to build up good face validity. The present research should prove useful in this respect.

There is a growing bank of motion capture data available for the purpose of perceptual experiments. For example, Ma, Paterson and Pollick (in press) have produced, from the walking, lifting, throwing and knocking actions of emoting actors, a library of over 4000 movements. These authors have argued that libraries such as this will be helpful in describing the general movement properties of affect that are used for recognition. Further, Ma et al. contend that these properties have so far not been detailed due at least partly to the fact that different research laboratories produce different movement samples which are used for the purpose of a single (or a few) studies. I agree with this contention. However, we must be careful in these attempts until the appropriate operational variables are defined. The present research, for example, has highlighted that recognition of some emotions is dependant on the social situation — this would not be known if all of the work conducted across
laboratories used behavioural samples from one library such as the one produced by Ma et al. (in press).

Therefore, all research that investigates affect is somewhat problematic due to the lack of operational definitions for the subject matter.

Limitations

The samples used in this research had good face validity in that the qualitative results (Chapter 3) showed that the emotions could be named correctly (apart from disgust) on at least half of the occasions. Good test-retest reliability was also demonstrated in the work where recognition was based on the upright and inverted/rotated displays (Experiments 1,1a, 1b, 2 and 3). However, one criticism of this work in general was that only one type of dependent measure (the rating scales) was used throughout the thesis (with the exception of the child study). The measurements made by using this method were not independently produced as judgments were made for every emotion on every trial (rather than having subjects only pick one). This may have some implications in terms of violating the assumptions needed for using parametric testing procedures. Furthermore, one could argue that the measurement scale could be ordinal (data must be at least of interval level for parametric testing). However, ANOVA (which was primarily used throughout this thesis as the primary inferential statistical method) is robust against violations such as these (see Howell, 1997). Furthermore, even if subjects had been asked to pick only one affect (out of the 5, or sometimes 6 or 4 choices available for every trial) they would still have been making implicit judgments about the other emotion categories in their heads. Moreover, it was useful for further understanding to have this error data to know the way in which emotions were confused with each other (for example, the way love was confused for sadness without the presence of the dyad — Experiment 3). Finally, the ANOVA results in this thesis were always reported with due caution in that, if sphericity (see Howell, 1997) was violated, the results from more conservative tests were used.
In Chapter 3 control experiments for obvious confounds such as timing and learning were conducted. However, the dyadic stimuli used in this thesis were quite complex and contained many parameters. More simple stimuli such as a single body part, like those used by Pollick et al. (2001), may be better for ensuring internal validity. There is always a trade-off between taking reductionistic approaches which are more likely to conserve internal validity but may compromise external validity. An attempt here has been to produce results that are high in both external and internal validity. The research here has plotted a middle course.

It is important to note that a shortcoming of this research may be the fact that there was no neutral condition inbuilt into the experimental design. However, given the communicative context in which the stimuli were produced the production of a neutral stimulus might have been difficult. Furthermore, to include a neutral stimulus would have been more important if more quantitative methods had been used such as signal detection. This research supports the view that some emotions might be more functional than others at an interpersonal level. Love is one such emotion; joy is a candidate for being another. The research here suggests that there is little difference in the way that other emotions (e.g. anger) are perceived between intrapersonal and social levels. However, a limitation of this research was that in Experiment 3 the monad condition showed an interlocuting actor with his or her partner occluded. Therefore, in this condition the interpersonal context may have been still observable albeit with one interlocutor hidden. It might be profitable, therefore, to consider a further study where the stimuli are produced both with and without interpersonal interactions in order to assess more fully the contribution of the social context per se. Another idea would be to have a static second actor to imply a social context.

Lastly, the natural interaction (rather than a pseudointeraction as in the reflected dyads condition) was shown to be important for the perception of some emotions such as fear. Why this was so has not been answered in this research. Clearly, social interactions are complex and there appear to be hidden patterns of movements within an exchange that contribute to the perception of affect. Pollick, Lestou, Ryu and Cho
have suggested that neural networks are effective pattern classifiers for complex stimuli (see also Magnusson, 1996). A discussion of such ideas is beyond the scope of the present work. However, Pollick et al. (2002) have used automatic pattern classifiers and thereby demonstrated that there are hidden patterns in point-light displays that are not picked-up by human observers.

Future studies

The production of movement by an observer interferes with their perception of a point-light walker (Jacobs and Shiffrar, 2005; see Chapter 4). This shows that results are altered by providing more realistic environmental conditions than those used traditionally in a laboratory. It was suggested in Chapter 4 that variations in the third person’s (the subject’s) social behaviour should be incorporated in the experimental design rather than, or as well as, variations within the visual display itself in order to comprehend whether an affordance based framework is appropriate for further understanding of these results. Whereas it is not difficult to ask a subject to walk (on a treadmill) whilst they make judgments about a point-light walker it is more difficult to design an experiment where subjects are required to make judgments about affect as well as experience affect. However, it is not impossible: a subject could be involved in a computer based virtual reality task where they become involved in a social event. However, this idea could be criticised. For example, Dennett (1991) has argued that no virtual reality tasks (however realistic they become) will ever be able to fool people. Furthermore, there becomes a point when trying so hard to mimic situations in the real world means that good internal (and therefore external) validity is lost due to introducing complex designs which permit too many potential confounds.

Hadjikhani and de Gelder (2003) have suggested that there are synergies between different processing modalities as the same important neural structures involved with the processing of affect are shared by different modalities of expression; this provides an opening to a new area of research. Further studies could be conducted
that explore the relationships between the various channels of expression such as the face and the voice, or the body and the face.

It can be argued that strong conclusions cannot be made on the basis of the results of a single developmental study (Chapter 6). Very little work has been carried out in the developmental field to explore the perception of affect from biological motion so it would be profitable to carry out some further developmental studies. It has been assumed for too long that the face is the primary channel for the perception of affect and even though it is likely that infants are exposed to faces more than bodies, given the foundational nature of body movement (see Chapter 1) it is possible that infants could discriminate affect from kinematics of the body. Thus, Bornstein and Arterberry (2003) have shown that 5-month-olds can categorise emotions from photographs of faces. Therefore, infants of a similar age would be chosen for a study where they would become habituated to different examples of point-light displays showing one type of affect, fear for example. Then, for the test trials, the infants would be shown examples of fear and happiness in simultaneously presented point-light displays. Another idea is to conduct a preferential looking task to see if infants can match the sound of voice expressing affect to a point-light display (of the body) that depicts the same emotion. This would test for cross-modal processing of affect as well as the ability to perceive emotion from body movements (and sound).
7.5 Conclusion

There are 3 main conclusions that result from the work investigated in this thesis. Firstly, the results show that the body is an excellent channel to convey the expression of affect. Secondly, the social environment is a defining feature of emotion as ability to recognise the stimuli can alter depending on the social context irrespective of intrastimulus kinematics. Lastly, the work has implications for the topic of emotion; different processing constraints exist for different emotions and these are indicative of a categorical structure for affect. Furthermore, the results from Chapter 3 are suggestive of a special module for the processing of fear. The aim of this work was to take a social approach to the investigation of affect from biological motion. It is hoped that this research will motivate future researchers to take into account contextual information, and in particular social cues, in experimental designs.
References


Hadjikhani, N. and de Gelder, B. (2003). Seeing fearful body expressions activates the fusiform cortex and the amygdala, *Current Biology, 13*, 2201-2205


Rensinck, R. A., O'Regan, J.K. and Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8,* 368-373.


