The Perception and Cognition of Emotion from Motion

Helena M. Paterson

Thesis submitted for the degree of Doctor of Philosophy
In the Department of Psychology,
Faculty of Social Sciences,
University of Glasgow

August 2002

Abstract

Emotional expression has been intensively researched in the past, however, this research was normally conducted on facial expressions and only seldomly on dynamic stimuli. We have been interested in better understanding the perception and cognition of emotion from human motion. To this end 11 experiments were conducted that spanned the perception and representation of emotion, the role spatial and temporal cues played in the perception of emotions and finally high level cognitive features in the categorisation of emotion. The stimuli we employed were point-light displays of human arm movements recorded as actors portrayed ordinary actions with emotion. To create them we used motion capture technology and computer animation techniques.

Results from the first two experiments showed basic human competence in recognition of emotion and that the representation of emotions is along two dimensions. These dimensions resembled arousal and valence, and the psychological space resembled that found for both facial expression and experienced affect. In a search for possible stimulus properties that would act as correlates for the dimensions, it emerged that arousal could be accounted for by movement speed while valence was related to phase relations between joints in the displays. In the third experiment we manipulated the dimension of arousal and showed that through a modulation of duration, perception of *angry*, *sad* and *neutral* movements could be modulated. In experiments 4-7 the contribution of spatial cues to the perception of emotion was explored and in the final set of experiments (8-11) perception of emotion was examined from a cognitive perspective. Through the course of the research a number of interesting findings emerged that suggested three primary directions for future research: the possible relationship between attributions of animacy and emotion to animate and inanimate non-humans. The phase or timing relationships between elements in a display as a categorical cue to valence and finally the unexplored relationship between cues to emotion from movements and faces.

Acknowledgements

I would primarily like to thank my two supervisors Professor Anthony J. Sanford and Dr Frank E. Pollick for their support and guidance in completing this thesis. To Professor Sanford I am most grateful for assistance in a continuing theoretical development of the research as well as many hours of interesting discussion regarding diverse subjects such as bees, falling rocks and Virtual Reality. To Dr. Pollick I am additionally grateful for the use of and training in the Biological Motion Lab and his assistance with valuable technical development for the research.

This thesis would not have been possible without the help of two additional people. One of these is Armin Bruderlin, who contributed the knocking and drinking movements used in many of the experiments. Additionally Ales Ude made the Humanoid Solid Body Models used in Experiment 6. I would also like to acknowledge the contribution of two honours students who did pilot studies for Experiments 3 and 7. These were Anne-Marie Walker (2000) and Kirsty MacDonald (2001).

Additional acknowledgement should go to the Economics and Social Research Council for the funding that made this research possible. Similarly I would like to thank the Research Committee of the Department of Psychology for funding to attend various conferences where I presented some of the research reported here. In advance I would also like to thank my viva committee consisting of Professor Anne Anderson convenor), Professor Simon Garrod (internal examiner) and Dr. Ian Thornton (external examiner).

A final thanks goes to my husband Andrew Cardwell for his patience and support through many months of stress, hard work and TV dinners.

Web Site

Demonstrations of stimuli are available through the following web site or by contacting Helena Paterson: http://www.psy.gla.ac.uk/~helena/demos/thesis.html

Declaration

I declare that the above-mentioned thesis embodies the results of my own special work, that it has been composed by myself and that it does not include work forming part of a thesis presented successfully for a degree in this or another University.

Table of Contents

Table of Figures

Table of Tables

Chapter 1– Perspectives on Dynamic Emotion

Introduction

In our everyday interactions with conspecifics, humans seem to effortlessly interpret and utilise the dynamic social signals available from the behaviour of those we engage in communication or simply observe. Of the social signals available, emotion is probably one of the most informative since it not only tells us of the emotional state a person is in, but also seems to convey information useful for various other interpretative processes. For instance, emotional expressions in a voice, face or movements can tell us of a person's attitudes to us, to other people or to the topic of conversation. An example would be when a person pulls a disgusted face while talking of modern art – indicating that they dislike it. Expressions of emotion allow us to infer the intentions of a person to deceive (Ekman, O'Sullivan, & Frank, 1999) or predict what they will do next – for instance a person approaching with a smile might greet you, while one approaching with a scowl might hit you. Emotional reactions that we can gauge from others' behaviour also inform us of the appropriateness of our own behaviour. For instance we might try to avoid negative reactions to our actions – such as anger and disgust - and seek out positive reactions - such as joy and admiration.

In keeping with the extensive role that emotions play in our interactions, emotional signals are conveyed in a wide variety of signals (Tagiuri, 1969). The perceptual signals that have been researched most often by psychologists fall into three rough groups: facial expression, prosody and body language. For facial expressions and body language there is a further sub-division into dynamic and static signals. For the most part, human facial expressions have been used as stimuli to investigate the perception of emotion, however, highly impoverished stimuli such as point-light displays of human actions convey sufficient information for the recognition of emotion (Brownlow, Dixon, Egbert, & Radcliffe, 1997; Dittrich, Troscianko, Lea, & Morgan, 1996; Walk & Homan, 1984).

While the everyday actions of humans – such as walking down the road or knocking on a door - are not intended to be communicative, common experience would serve to illustrate that they often carry characteristic and transient person information. Such information is available for use in person perception (Hill & Pollick, 2000; Dittrich, Troscianko, Lea, & Morgan, 1996; Mather & Murdoch, 1994; Runeson & Frykholm, 1983, 1981; Barclay, Cutting, & Kozlowski, 1978; Kozlowski & Cutting, 1978; Cutting & Kozlowski, 1977; Kozlowski & Cutting, 1977) and it is likely to drive perception in the absence of other cues. This is likely to also be the case in the perception of emotion. For this reason we have been interested in better understanding the perception of emotion from dynamic human movements in the absences of form cues such as facial expressions.

Although it is unusual to see a person without their face, there are circumstances in which facial expressions may be obscured while dynamic movements are still available to perception. For instance, in the case where a person dons a mask (during a bank robbery or a masked ball) or when they are viewed from behind. Apart form person perception; there are also other circumstances in which the emotions of a perceived entity are important. For instance, it is important to recognise the aggression of a pet dog or a predator. Similarly, recognising the fear in a herd of cattle can prevent your presence during a stampede. In all of the examples above it is not possible to perceive regular facial expressions, however, intuitively emotions can still be perceived from other signals.

It seems that humans automatically attribute feeling and emotions to animals, birds, insects and abstract objects for instance Heider and Simmel (1944) showed that abstract objects such as a 2D circle, square and triangle were attributed with feelings, moods and personalities from the way in which they moved. This has been recently extended by animators (Lasseter & Rafael, 1987) to make ordinary objects such as a desk lamp appear to have emotions - such as in the case of Luxo Jr. (Lasseter, 1986). Additionally, it is easy to imagine an *angry* swarm of bees or birds. In the cases of Luxo Jr. (Lasseter, 1986), birds and bees movements would be the only cues to emotions. Other animals – primates for instance - do have faces, but their expressions may not correspond to those made by human faces and movements or posture may offer clues as to their emotions[1].

We have been interested primarily in further investigating the perception of emotion from movement. To this end we have used point-light displays (Johansson, 1973) to present human movements. This technique allows

[1] It is not universally accepted that animals and insects have emotions, however, we have use the term "emotion" loosely to encompass an inference that some entity has emotion and what that emotion might be in a given situation. Luxo Jr. certainly has no emotion, but emotion is still inferred from its behaviour.

human movements to be presented devoid of conventional form cues. In an attempt to better understand how emotion may be recognised from movements, we have asked a number of questions:

- How accurate are humans at recognising emotion from motion?
- What stimulus properties might they use to achieve recognition?
- Is it possible to manipulate these stimulus properties in such as way as to modulate perception?
- What role does high level cognitive processes play in the perception of emotion?

The rest of this chapter is devoted to short reviews of the literature on perception of emotion and of point-light displays, it concludes with a more detailed development of the main theme in the current thesis.

1.1 Perception of Emotion

Introduction

In the psychological enquiry into the perception of emotion, two broad themes have emerged over the last 100 years. The first of these concerns the biological and evolutionary underpinnings for the perception and communication of emotion while the second has concerned the more phenomenological and introspective aspects of experienced emotion. The common aims in both has been to systematically develop and test theories of emotion with the ultimate goal being to classify and more fully research the nature of human and animal emotion. In terms of perception of emotion, this has resulted in two rivalrous theories of how facial emotion expressions are represented – either as discrete categories or alternatively as points along a small number of dimensions that also describe the underlying representations of experienced emotion.

In light of recent research on mirror neurones (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996) and the dichotomous nature of perception and action, the problem of how perception and experience of emotion may be related has become interesting. A first step in conducting such research has been to establish a relationship between perception and experience. For this reason some of the emphasis in the current thesis was placed on the perception of emotion and how it may relate to the experience of emotion from a behavioural perspective. To refine the focus, there has been an emphasis on establishing the existence of a common representation space for perceived and experienced emotion and this issue was explored in Experiments 1 and 2.

Another aim of the research was to understand better the perception of emotion from human movement. Although there has been a large body of work on the perception of emotion from static faces, the corresponding work on perception of dynamic faces and human movement has been minimal. While static faces represent the end-state of an emotional expression, these are not the only available signals of emotion in the social world. Other visual signals from humans include the dynamic formation of a facial expression, static and dynamic posture (stance or body language, for instance) and emotion signals encoded in movements. In order to gain a full understanding about the manner in which humans perceive emotion it is important to understand all the visual as well as non-visual signals that may be available during social interaction/observance.

Brief History of Developments

Perception of emotion from facial expression was popularised in the late 19th century by Charles Darwin in his book The Expression of Emotions in Man and Animals (1872). In this book Darwin postulated a minority held belief that both humans and animals experience emotions and that the emotional experience of both have a common evolutionary and motivational underpinning. Darwin made detailed observations of various species of animal and provided sketches of their facial and postural expressions to illustrate his text. From this he concluded that the same expressions are found in both humans and animals. He also theorised about the biological and social underpinnings of emotional expression concluding that that emotional expressions were vital for the regulation of social interactions and the welfare of species living in social groups – such as apes and humans.

The basic premises of Darwin's research contradicted the popular belief that human emotion and expression are a result of their superiority over other animals. A main proponent of this supremacist view had been Descartes who believed that animals were automatons, compared with humans whose emotions were a result of rational thought, gifted by divine intervention. Descartes related emotions to biology, but still maintained that emotion was controlled by reason. The basis for emotion as biological can be traced back to Aristotle who claimed that emotion resulted from a mixture of the bodily humours. Darwin's analysis of the biological nature of emotion concluded that emotion was coded in genes through the evolutionary process, giving a continuance between human and animal emotions.

In no small way the study of emotional expression in the 20th (and now the 21st) Century can be seen to have its roots in the ideas of Darwin. The advent of late 19th and early 20th century saw the research of Darwin causing a paradigmatic shift in the way the natural world, and in consequence human behaviour, was investigated. A main consequence of this was to launch systematic research into the way human expression is generated and perceived. In such research there has been a particular emphasis on the evolutionary origin of expressions and in particular the innateness of emotional expression and perception.

Perception of Facial Emotion

Darwin's ideas that facial expressions were innate and could be used to study emotion inspired a number of researchers to replicate his early study on the perception of expression. A major issue in Darwin's work was that expressions were well recognised and interpreted by humans and animals for social communication. In consequence early researchers showed pictures of facial expression to large numbers of observers who were asked to name the expressions. Bruner and Tagiuri (1954) reviewed this research and concluded that there was no invariable pattern in the expression of emotions. In general the early research on perception of expression, taken as a whole, seemed undecided on a primary question: can emotions be recognised from faces?

Any conclusion that emotion cannot be recognised from facial expression is in conflict with everyday human experience and a number of researchers set out to show conclusively that humans could reliably recognise the expressions of others. The approach taken by this second wave of researchers was to look for innate or basic human emotions that could be expressed in the face. This approach persists today and has two inherent assumptions, firstly that in human cognition emotions are represented as discrete categories and secondly that emotional expression are the reflection of human feelings, so by studying them the researcher can make assumptions about the existence of the same emotions in observers. This research into basic, human expressions came into its own in the 1970's when Izard (1971)and Ekman, Friesen, and Ellsworth (1972) published their anthropological studies on the innateness of facial expression of discrete emotions. Izard presented research to show that fundamental (basic) emotions were found in humans across different cultures and from the very early weeks after birth. The fundamental emotions used by Izard had previously been defined by Tomkins (1962)and Tomkins and Izard (1965) as surprise, enjoyment, disgust, shame, distress, interest, fear, contempt and anger with both mild and intense examples of each. In their work, Ekman and Friesen reported visits to western and eastern cultures as well as an isolated tribe in New Guinea, also concluding that there are a small number of innate emotions that are expressed and recognised by all humans. These discrete emotional categories were happiness, sadness, anger, disgust, surprise and fear. These last emotions are at present the most commonly used in research on facial expression.

In parallel to this categorical account of emotion perception, a dimensional account has also been developed. As with the categorical account, the dimensional account developed in answer to the disparate findings from early research. In this account, emotions are represented along a few dimensions that would allow the classification of expressions. Schlosberg (1952) proposed 2 major dimensions along which emotions were represented according to the confusions observers made regarding them. The first of these were pleasantness-unpleasantness and the other was attentiveness-rejection, later he also suggested a third dimension coding for activation (Schlosberg, 1954). The basis of these findings were the confusions between emotions and the results were replicated in a number of studies by both Schlosberg and other researchers (Tagiuri, 1969).

In a review of the universality of emotional expressions, Russell (1994) criticised proponents of basic/fundamental emotions from the perspective that they did not consider other explanations for the data that was collected. The two review articles published along with the Russell's critique (Ekman, 1994; Izard, 1994) show the extent to which issues about perception of facial expression are still a hotly debated subject.

In an attempt to settle the question of whether expressions are represented categorically or along dimensions, Young, Rowland, Calder, Etcoff, Seth and Perrett (1997) recently used modern morphing techniques to address the question of whether the representation of emotions were categorical or dimensional. They presented morphed facial expressions and asked participants to judge the images according to the labels attributed to prototype faces. Their findings were that facial expressions were represented categorically. This was because between emotion categories there was not a smooth transition in the responses to morphs of prototype expressions. Instead there was a sharp change in the perception of morphs where a category boundary was located. Takehara and Suzuki (2001), however, used the morphing technique in a similar study and found that a two-dimensional structure was best for representing recognition of facial expressions. The critical difference in this study was that Takehara and Suzuki (2001) requested participants to rate each expression according to all the prototype labels.

A striking feature of the research that has contributed to this debate has been that the methods used by each of the groups are substantially different. All studies have similarities, such as the presentation of photographs to observers and then eliciting a response from them. However, the techniques used for analysing the data are different. Proponents of categorical perception, often analyse data in terms of the proportion of correct recognition according to pre-specified labels. Since labels are pre-specified, it may be the case that a participant's views are biased to use the best available label for the situation rather than responding according to their initial beliefs. Additionally this research suffers because photographs of expressions have been very carefully pre-selected (Russell, 1994) to provide prototypes of only a few emotions. It certainly does not seem that any researchers ever reported negative results for non-basic expressions. For this reason the lack of confusion between their expressions would be because they selected expressions far apart on the dimensional scales suggested by proponents of the dimensional account.

Proponents of the dimensional account have a range of techniques for eliciting responses, such as picture sorting, similarity ratings, free response and forced choice, and the participants' responses are normally analysed according to the confusion between different expressions and analysis involves multi-dimensional scaling. In this research pre-selected faces are also often used and the tasks used to elicit responses need to be in a form that can be transformed to dissimilarity ratings for multi-dimensional scaling analysis. Whenever the dimensional model is used as an account of emotion recognition, it also the case that a categorical account is not excluded by reporting the data in this way.

The methodological differences that have been used to separately show the categorical and dimensional nature of facial expression perception has made it hard to find correspondence between the two research areas in order to compare results and findings. In order to understand the underlying structure of emotion, it is therefore necessary to find correspondence between studies for the two approaches. Dailey, Cottrell, Padgett, and Adolphs (in press) have addressed this issue in recent research. Dailey et al. (in press) simulated the findings from human perception in a biologically based neural network. They found that after learning to classify facial expressions of Ekman and Friesen's 6 basic emotions (Ekman, Friesen, & Ellsworth, 1972) the model matched a variety of data from the literature on perception of emotional expression, without additional parameter tuning. This data included those of Young et al. (1997) as well as MDS analysis of data presented by Ekman and Friesen (1972).

Hence the results showed that data from both the dimensional and categorical accounts of emotion representation were described by the data. However, rather than these representing an underlying psychological representation, the data supports a fully perceptual representation. This is particularly important for the proponents of a dimensional account, since it would indicate that a dimensional structure derived from perceptual data may not have a relationship to an underlying psychological space that in turn is derived from underlying human emotional physiology (Dailey et al., in press). The reasoning behind this is that a machine with no underlying physiology or psychology of emotion spontaneously generated the same dimensional perceptual space as human participants. The human representation is therefore likely to be perceptually driven and Dailey et al. (in press) hypothesised that the development of the expression-to-emotion mapping evolved in tandem with the need to communicate emotional states effectively.

Structure of Experienced Emotion

In exploring the way in which humans recognise affect, it is possible not only to look at perception of other people's expressions, but also to examine the experience of emotion. There is recent evident to suggest that experience and perception interact when observing another person's actions (Decety & Grezes, 1999; Rizzolatti et al., 1996). Additionally biological motion relies on both specialised bottom-up processes of motion detection (Neri, Morrone, & Burr, 1998; Mather, Radford, & West, 1992) and interactions between these and top-down processes (Thornton, Pinto, & Shiffrar, 1998; Shiffrar & Freyd, 1993, 1990). In the case of emotion perception such a top-down process may well originate from the influence of an internal structure of affect.

A number of models for the structure of experienced affect have been proposed recently (Thayer, 1996; Larsen & Diener, 1992; Watson & Tellegen, 1985; Russell, 1980). The basic and consensual finding from this research has been that experienced emotion is represented in a two-dimensional structure. The main divergences between models have been the rotation of the axes and labelling of dimensions (Yik, Russell, & Feldman-Barrett, 1999). This similarity between the models has been used by Russell and colleagues (Barrett & Russell, 1999; Yik, Russell, & Feldman-Barrett, 1999) to propose an integrated structure for the representation of experienced emotion. This structure is based on the circumplex model (Barrett & Russell, 1999), a two-dimensional model with a circular structure first proposed by Russell (1980). The two dimensions in the model are independent and bipolar with one dimension representing valence (for instance hedonic tone, pleasant – unpleasant) and the other, arousal or activation (arousal – sleep/ activated – deactivated). The models are also continuous, with affects falling on a circle, centred on the origin of the psychological space defined by the two dimensions. This is in contrast to a "simple" structure in which the items would fall in tight clusters.

Between the psychological spaces for perceived (as proposed by proponents of a dimensional account of perceived affect) and experienced affect there is a lot of similarity. For instance the dimensions proposed by Schlosberg (1952) for facial expressions were also valence and activation.

1.2 Perception of Emotion from Dynamic Stimuli

Identity and Expression Recognition from Dynamic Faces

In all the research discussed above, photographs of faces that show the end-point of an emotional expression were used as stimuli for judgements of emotion. Using static pictures of facial expressions is a problem since it excludes all the dynamic cues that are encountered in normal human communication. It is unusual for us to see still faces, rather there is a large number of facial motions that we encounter during communication. Recently this issue of whether motion information would enhance recognition has been addressed for the perception of identity with mixed results (Thornton & Kourtzi, 2002). For instance rigid rotations of a 3D head enhanced identity recognition (Pike, Kemp, Towell, & Phillips, 1997). Non-rigid motion, however, seemed to add little information (Bruce & Valentine, 1988) apart from when faces were well known (such as celebrities) and under poor viewing conditions such as viewing photographic negatives (Lander, Christie, & Bruce, 1999; Knight & Johnston, 1997) or threshold manipulated images (Lander, Christie, & Bruce, 1999). Lander and Bruce (2000) showed a general benefit for recognising moving faces along with the results of enhanced recognition for familial faces. The enhancing benefits of dynamic stimuli were best observed under characteristic motion signals compared with reversed motion or temporally disturbed signals. Recently Hill and Johnston (2001) showed further that rigid motion benefited recognition of identity while non-rigid motion was useful for categorising gender.

Thornton & Kourtzi (2002) used a priming paradigm with a matching task to show that unlike previous research (Lander, Christie, & Bruce, 1999; Christie & Bruce, 1998; Knight & Johnston, 1997; Bruce & Valentine, 1988) non-rigid motion enhanced recognition of unfamiliar faces as measured in reaction times. In their final experiment, Thornton & Kourtzi (2002) changed the task to be a matching for facial expressions. In this case no advantage for dynamic over static information was found. This agrees with findings by Kamachi, Bruce, Mukaida, Gyoba, Yoshikawa, & Akamatsu (2001) in which the temporal properties of morphed expressions were manipulated in order to investigate the role played by dynamic information on expression recognition. They presented morphing sequences in which a face changed from a neutral expression to an emotional expression in a sequence of 6 (fast), 26 (medium) and 101 (slow) frames and measured the accuracy and intensity ratings of judged emotion. Kamachi et al (2001) found that that there were complex interactions with the time taken to present expressions so that happy and surprised dynamic movements were best recognised from short temporal sequences. Sad movements were best recognised from long sequences and angry movements were best recognised from medium length sequences. However, only sad movements were better recognised from dynamic compared with static stimuli. For ratings of intensity, dynamic expressions were never rated to have greater intensity than static displays.

The findings of Thornton & Kourtzi (2002) and Kamachi et al (2001) is in opposition to the conclusions by Edwards (1998) who suggested that the temporal characteristic of dynamic facial expressions were important in themselves. Edwards used static poses of an expression sequence that had to be sorted from being scrambled to the correct sequence. She found that participants were sensitive to the dynamic expressive information encoded in static frames. One criticism of Edwards (1998) may be that she did not directly compare moving and static stimuli. In all the research with expressive faces that have been reported so far, it seems to have been the case that stimuli were pre-selected for their effectiveness to communicate emotion. It may be the case, as with facial identity, that only under adverse conditions would the dynamic information play a role. It certainly seems that in normal conditions where a good exemplar emotional expression is available, no benefit is gained from additional dynamic information, but that a human observer is still sensitive to the dynamic characteristics of facial expressions.

From this perspective the data presented by Kamachi et al (2001) (presented in Figure 3, pp. 883 of Kamachi et al, 2001) should be re-examined since non-preferred temporal sequences seemed to cause some interference in recognition. So not only were angry movements recognised best during short or medium temporal sequences, but when an angry sequence was presented in a long temporal sequence, it reduced the effectiveness of the expression to communicate anger. Similarly, when sad movements were presented at the medium or long rate they were recognised at levels comparable to static faces, but at the fast rate they were not as effective at communicating sadness. This would agree with the suggestion by Edwards (1998) that humans are sensitive to the temporal qualities of facial expressions.

Recognition of Facial Expression from Point-light Displays

A way in which to more closely examine the role motion might play in the perception of expressive faces, is to isolate the dynamic motion signal from static form information. One technique for achieving this is to use point-light displays (this technique is described more fully in on page 23). Bassili (1979; 1978) used point-light displays of facial motion to investigate whether facial motion alone would be sufficient for recognition of expression. He found that although participants recognised faces best when normally illuminated, emotions could still be recognised from point-light displays of facial motion. In comparison to full-light displays, point-light displays generally resulted in worse recognition, however, for happy surprised and disgust the two types of displays were closely comparable. Additionally, different parts of the face were important for recognition of different emotions (Bassili, 1979), so that happiness and sadness were best recognised from the bottom half of the face while surprise was equally well recognised from both top and bottom halves. Anger was best recognised from the top half of the face and fear and disgust were recognised best when the whole face was available. Static point-lights were not as informative for face recognition and in a full-light dynamic motion condition expressions were recognised with greater accuracy. Generally the same pattern of results for regional importance were found for the full light condition as for point-light displays.

Apart from the research by Bassili, there have been no other attempts to study human facial motion in isolation for expression recognition. The exception is recent research by Pollick and colleagues who investigated the role of rigid head movements for expression recognition (Frank E Pollick, Hill, Folwaczny, Gray, Holmes, & Love, 1999) as well as spatially compared with temporally modified movements (Frank E. Pollick, Hill, Calder, & Paterson, submitted). They found that positive spatial exaggeration of facial expressions increased ratings of intensity of perceived emotion as well as enhancing recognition. This was not the case for duration manipulated movements. They concluded that for their stimuli, duration was not diagnostic of expression recognition since the range of duration for different actors (800-1266 ms) was actually greater than the range of duration for different emotional expressions (858-1100ms). In comparison, for the artificially dynamic faces of Kamachi et al (2001) there was no variation between actors and additionally the range of duration for expressions (200-3357 ms) was much greater than those recorded by Pollick et al (submitted) from human actors.

From the above discussion, it seems that dynamic information for faces plays a role in recognition of facial emotion expressions. However, it is unclear what this role may be or in what way form and motion information would interact. Additionally the findings by Pollick et al (submitted) raised questions about the usefulness of artificially created human motion since artificial motion would necessarily exclude some of the variation that may naturally occur in the human population.

1.3 Biological Motion

Introduction

The term "biological motion" in vision science relates generally to any stimuli that comprises human or animal movement. However, it is often used interchangeably with the term "point-light displays" to denote stimuli in which motion is shown independent of form cues. A number of techniques have been developed to capture such formless human motion and these have been summarised in the sections below.

History of Motion Capture

For much of the work in the following thesis, biological motion sequences have been used as stimuli. This type of display was pioneered by Johanson (1973), however, one of the basic premises that lead him to developing point-light displays (PLD's) – to study human motion independent of form – had been developed in the 1800's, but could not be implemented with the same effect. Here I will trace the development of motion capture techniques from their early beginnings during the development of photography to the modern-day use of digital computers.

Static Beginnings – Muybridge and Marey

The motion of humans and animals has been studied extensively throughout history by researchers and philosophers. Early studies tended to focus on the mechanics of motion in an attempt to improve the training of animals and athletes. Motion had been studied from the perspective of gaining information about the physical properties of a movement and much of this research has advanced to physiotherapy, physical rehabilitation, prosthetics and ergonomics. Advances in human movement science have relied heavily on technological advances such as photography, film and video and more recently, computer animation. In the following paragraphs, the efforts of the first photographers of human movement – Muybridge and Marey - will be summarised.

Muybridge and Marey are popularly best known for their roles in developing cinema technology, but within the biomechanical sphere they are famous for being amongst the first researchers to use photographic techniques to study the motion of humans and animals. Muybridge's first claim to fame was, through the use of sequential photographs to resolve a centuries old debate about the motion of horses galloping. This debate had been centred on whether all four of a horse's legs left the ground during a gallop. Muybridge's photographs showed that this was indeed the case, however, the results also showed the error of artists in depicting a gallop with the horse's legs splayed in mid-air.

Once his photographic techniques had been refined, Muybridge (1899; 1895; 1893) went on to carry out time and motion studies of other animals as well as humans. He also made detailed analyses of the gait of humans and animals, other activities of men, women and children, birds in flight and large birds walking, in addition to other unusual animals. Apart from the impressive array of subjects photographed by Muybridge, that were later animated into movie sequences, there was little theoretical development about human movements in the work of Muybridge.

At about the same time that Muybridge was working in America, Marey was working in France on similar problems. He also wanted to catalogue the mechanics of motion in humans and animals with an extension to the motion of inanimate objects. Whereas Muybridge used successive photographic plates exposed in sequence to produce distinct images, Marey used chronophotographs (Marey, 1895). The technique he developed allowed the repeated exposure of a single photographic plate at pre-specified intervals, allowing the representation of a complete sequence on a single picture. With this technique it was also possible to capture motion at greater velocity than Muybridge had been able to achieve.

It soon became evident to Marey that in a typical photograph of dynamic events, there is too much information to make any useful study of motion. To combat this, he developed the technique of dressing his actors in dark clothing with a white cloth string stitched along the line of each limb. As the actor moved, a camera was used to make repeated exposures in a single photographic plate, causing the motion to be captured in a static photograph. These photographs look very strange, a little like stick aliens moving and leaving a glowing train of spots to show where they had been. This was the first attempt to limit the information available to the essentials that were being studied i.e. the position of limbs at certain times, exclusive of the features of the body. In his work Marey was concerned with the mechanics and ergonomics of motion and used his findings in military applications to the French army.

Marey found a great benefit in isolating motion from form in that it allowed him to analyse movement for a number of applications. In vision research there is also a benefit to using movement separated from form since only the motion signal can be presented to an observer. Form has often been presented in isolation through the use of still photographs and the combination of form and motion is available in video displays, but techniques for presenting real human movements were only recently developed (Johansson, 1975, 1973).

Point-light Displays – Johansson

Johansson (1975; 1973) may arguably be regarded as the first researcher to isolate human movement from form cues for use in perception experiments. Previously perception of human movements had been studied by Wolff (1943) who disguised identity and other form cues by dressing actors in baggy clothes and hoods while recording their movements on film, however, this method could not fully separate form and motion. Heider and Simmel (1944) solved the problem by making animations of abstract objects, however, in this case the movement was artificial. Johansson approached the perception of human movement from the perspective of event perception and used it to further investigate a vector analysis theory that was developed to account for event perception.

The video-based techniques developed by Johansson (1975; 1973) for the capture of human movement were relatively simple and can be traced back, in biological motion research, to the research of Marey and the earliest experiments in event perception (Michotte, 1963). The basic idea was to remove all visual cues to form so that only dynamic information from a display stimulus would be made available to observers. In this way shape and form cues play no part in the perceptual processes resulting in only movements being studied. Johansson (1973) described 2 techniques for motion capture. In the first technique light bulbs were fixed close to the joints on the clothing of actors and a 16mm film was made of the lights as the person moved. Some technical problems with using light bulbs resulted in him developing a second, but related technique. In the second technique new video technology was used to tape actors with reflexive tape attached to their joints. To present the stimuli, a TV was monitor set to maximum contrast so that only spots of light were visible on the screen.

A startling observation about the point-lights were that when presented statically, they were judged by naive observers to represent non-human stimuli such as star constellations or Christmas tree lights. However, within a very short period of time after movement commenced, observers perceived 8-10 points of light as a person and could identify the action that was being performed (Johansson, 1973).

Modern Motion Capture Techniques

Recently a number of digital motion capture techniques have been developed to cope with the growing range of commercial uses of captured human movement data in the entertainment industry. In vision science and motor control research the need for human motion capture data has also grown beyond the use of Johansson's simple video recordings of human movements and have started to use the range of new motion capture technologies. Such specialisation in vision research ranges from fine movement measures of eye movements (Yang, Dempere-Marco, Hu, & Rowe, 2002) and grip aperture (Franz, 2001; F. E. Pollick, Chizk, Hager-Ross, & Hayhoe, 2001; Aglioti, Desouza, & Goodale, 1995) to planar arm movements (Ghafouri & Feldman, 2001; Pollick &

1997), posture (Freeman, Avons, Pearson, Meddis, & IJsselsteijn, 2000) and computer visual analysis of human movement (Aggarwal & Cai, 1999; Gavrila, 1999).

For the purposes of the current thesis, we have been mostly interested in tracking technology that digitally measures human body movements. These techniques range from digitised video to optical and magnetic tracking systems.

Digitised video is probably the oldest of the three techniques and for this reason it is available in a number of formats. The most labour intensive video capture technique is to record an actor's actions in video and then to digitally encode elements of the display on a frame by frame basis. For point-light displays, individual frames from a video sequence or a number of sequential photographs (Mather & West, 1993) would be scanned into a computer. For each frame the joints would be defined and captured as individual elements that could be extracted using signal processing of the image. The individual frames could then be played back at the correct frame rate, resulting in a point-light display of biological motion. The benefit of this technique is that it allows any video or photographic motion to be digitised – for instance Mather and West (1993) used the sequential photographs captured by Muybridge (1899; 1893) to make point-light displays of animals in motion.

Digitising by hand is a laborious process and recent video-based motion capture technology has concentrated on image processing and computer vision to provide a shortcut to the process. Gavrila (1999) has supplied a comprehensive review of these techniques and while the applications for these techniques are still very much in the domain of computing science and computer vision, it is probably the case that they will become more widely used in research on perception. This is mostly because they allow the surreptitious recording of human movements in a natural setting (Deutscher, Blake, & Reid, 2000). Potentially any human actions and interactions could be changed into point-light displays or other forms of controlled stimuli. For instance in the applied case where police officers must make online identification of potential cases of violence from closed circuit television (Troscianko, 2002). In this case there is a great benefit to reducing the information made available to observers so that only task relevant content must be judged. Point-lights potentially offer such a data reduction scheme.

In comparison to the naturalistic situation above, it is sometimes desirous to have controlled stimuli where, in the end result, there would be correspondence between different actors in the same situation (Freeman et al., 2000; Burton, Wilson, Cowan, & Bruce, 1999). Three techniques for capturing human movement data in such situations are used most regularly in psychological research. The first of these are video-based techniques that track markers fixed to the clothing of an actor. These markers can be isolated from the rest of the scene so that only the human motion would be captured. These systems are in general very expensive and are hence only commercially available such as in the Edinburgh Virtual Environment Centre (EdVEC, 2002).

The second type is magnetic tracing systems such as Flock of Birds® by Ascension Technology Corporation. These systems are also expensive and while commercial studios can afford to track a large number of markers, in research it is more common to track only a small number of markers to for instance track the posture of an observer in a virtual environment (Freeman et al., 2000).

The final and most widely used motion capture technology for research, is an optical motion tracking system - the Optotrak system from Northern Digital Inc. As with video based and magnetic motion capture systems, for an optical system markers are attached to the joints of an actor. Three or more cameras are then used to track point sources of infrared light output by the markers. The cost of this tracking system is within the range of research departments and has become popular in vision research for perception of biological motion and research into action and perception interaction.

Benefits of Digital Motion Capture

Although they are expensive and rely on specialist expertise, the digital motion capture systems above offer a number of benefits to research. The first among these is that they reduce the information available to a researcher and in a stimulus display, to elements that are of primary interest. This feature is also common to the old-style video technology, however, there is an added benefit from 3D captured motion in that is allows a presentation of the displays from any viewpoint. Additionally it makes the measurement of movement properties a relatively easy procedure. Finally since the information is in a digital format with only a few elements that have correspondence, it is easy to perform techniques such as morphing and motion warping.

1.4 Perception of Biological Motion

Categories of Biological Motion

A measure of spontaneous stimulus categorisation occurs during this recognition of biological motion. The most obvious category is that the motion is made by some type of animal and then whether it is human or not (Mather & West, 1993). Another category is the type of motion, in other words whether it is a person/animal walking or running. A third is the recognition of a person's activities, such as dancing (Brownlow et al., 1997; Dittrich, 1993; Walk & Homan, 1984), cycling, climbing and gymnastics (Johansson, 1973) and lifting boxes of different weight (Runeson & Frykholm, 1983, 1981). (Dittrich et al., 1996) showed subjects three categories of activity and recorded their responses in the categories and in the activities. The categories of actions were locomotory (walking, climbing stairs, jumping and leaping), instrumental (hammering, ball bouncing, stirring and box lifting) and social (pairs dancing, boxing, greeting and threatening each other). The locomotory actions were recognised fastest and with most accurately, results for the social activities were also rather good, but for the instrumental activities - which relied on less expansive and more finely controlled motor activities - recognition was with the poorest accuracy and reaction times were longest.

Of greater interest to us has been the identification of human styles of movements. Such styles include the gender, identity and feelings of a person.

Gender and Identity from Point-light Displays

Cutting and Kozlowski (Cutting & Kozlowski, 1977) studied the possibility of point-light displays carrying information about the identity of an actor. Their aim in using Johansson's technique was to strip all familiarity cues, apart from motion, from their stimuli. They found that viewers could recognise themselves and others with above chance accuracy from point-light displays of movement. Viewers tended to use dynamic cues from movement to do the task such as the speed, bounce and rhythm, as opposed to non-dynamic cues such as the height of actors. In a second study they also examined the way in which the sex of an actor can be recognised (Kozlowski & Cutting, 1977) and found that recognition of gender was also above chance. Cutting, Proffitt, and

Kozlowski (1978) continued this research to find a biomechanical invariant that differentiated male and female human walks. They named the invariant the centre of moment and it comprised a perceptually inferred point on the torso of a walker related to the ratio of shoulder width to hip width. This theory was further tested by Cutting (1978) when he generated a synthetic human walker in which the shoulder and hip widths were exaggerated to make hypernormal male and female walkers and found it to be a robust invariant cue that observers use to judge gender.

Other Aspects of Person Perception

Recognition of gender and identity are two types of person perception that has received attention from researchers using point-light displays. However, other aspects of person perception that have been investigated are perception of lifted weight (Bingham, 1987; Runeson & Frykholm, 1983, 1981), deceptive intention (Runeson & Frykholm, 1983) and emotion (Brownlow et al., 1997; Dittrich et al., 1996; Walk & Homan, 1984).

Runeson and Frykholm (1981) suggested that, as with inanimate objects, the kinematics evident in point-light displays of human movement, would have enough complexity for a viewer to directly perceive the dynamics in an event. A particularly potent example of this was observers' ability to accurately estimate a weight that was being lifted by an actor. This ability extended from video and real life displays to point-light displays of action. Runeson and Frykholm (1983) went on to suggest that underlying dynamic patterns are the result of biomechanical necessities of anatomical structure and motor control. These underlying patterns reveal themselves in a person's movement making it possible for an observer to directly perceive person characteristics such as gender, identity and emotion. It also made it difficult for an actor to practice deceptions that would violate this underlying structure. Such deceptions included pretending to be of a different gender and attempts at deceiving an observer about the weight an actor was lifting.

1.5 Perception of Emotion from Biological Motion

The problem of recognising emotion from human movement has been explored for the special case of the interpretation of stylised dance movements (Dittrich et al., 1996; Walk & Homan, 1984). Of the six emotions examined (surprise, fear, anger, disgust, grief/sadness, and joy/happiness) it was found in both studies that anger was the most reliably identified emotion. Other differences among the identifiability of the different emotions were noted, however, they fell into no particular pattern between the two studies. The overall rate of recognition for the six emotions reported by Dittrich, et al. (1996) was 63% for point-lights. While these results provide good evidence that stylised movements can be seen as expressive they do not address the more general case of movements that are not stylised. Additionally, the data collected by Dittrich et al (1996) was only analysed in terms of the accuracy with which emotions were recognised.

In exploring the way in which humans recognise affect, it is possible not only to look at the accuracy with which an affect is recognised, but also at the structure of the representation of affect. A number of similar models for the structure of experienced affect have been suggested (Barrett & Russell, 1999; Thayer, 1996; Larsen & Diener, 1992; Watson & Tellegen, 1985; Russell, 1980). Primarily the similarities between models are that the structure of affect is a two-dimensional and continuous one. The two dimensions are bipolar and independent and one dimension represents valence (for instance hedonic tone, pleasant – unpleasant) and the other, arousal or activation (arousal – sleep/ activated – deactivated). The model is also continuous, with affects falling on a circle, centred on the origin of the psychological space defined by the two dimensions.

One important issue to consider is that the circumplex model has been established as a model of a person's own experiences of affect and thus would not necessarily apply to the perception of affect. However, there is evidence to support the prediction that an internal model of one's own experience could serve in the perception of the movement of others. For example, there are a number of recent findings which indicate that visual and motor processing interact whilst an observer watches an action (Rizzolatti, Fadiga, Fogassi, & Gallese, in press; Fogassi, Gallese, Buccino, Craighero, Fadiga, & Rizzolatti, 2001; Decety & Grezes, 1999; Decety, Grezes, Costes, Perani, Jeannerod, Procyk, Grassi, & Fazio, 1997; Rizzolatti et al., 1996). In this research a typical finding has been that for both humans and monkeys there is a system of mirror neurones that are activated both when the subjects sees an action and when they make the action. There is also good evidence that this finding extends to perceived and experienced emotions (Adolphs, 2002)

Moreover, evidence from the study of biological motion (Thornton, Pinto, & Shiffrar, 1998; Shiffrar & Freyd, 1990) indicates that the perception of biological motion relies not only on specialised bottom-up processes of motion detection (Neri, Morrone, & Burr, 1998; Mather, Radford, & West, 1992) but also on interactions between bottom-up and top-down processes. In their research Thornton, Pinto, & Shiffrar (1998) found that disruption of the low level temporal signature disrupted perception of biological motion, and underlined the role of low level motion filters. However, they also showed that perceptual processing of point-light displays took place over extended spatio-temporal periods and hence engaged high level visual processing. Thus, it appears reasonable to conjecture that an internal model of affect would be part of the top-down processes involved in organising the perception and categorisation of emotion.

In the research for this thesis we have endeavoured to use everyday human actions that may carry information about emotions. Our primary goals were also to explore the interaction between a high-level representation of emotion and the low level kinematics properties that are available in human movements for perception. For a successful interaction between a high level representation and a low-level visual signal, there has to be a mapping of distal stimulus properties onto an internal representation. It is this representation and its possible relationship to stimulus properties, which has been the focus of the following research.

In Chapter 2 (Experiment 1) we will collect perceptual judgements from observers as they view point-light arm movements of humans performing ordinary actions with emotions. These judgements will be used to chart both the proportion correct recognition and the representation of emotion from motion in human observers. In Chapter 3 (Experiment 4) these findings will be further explored through a correlation with the movements kinematics. In Chapters 4 and 5 (Experiments 3 -7) stimulus properties that would inform categorisation and recognition of emotion will be further tested. In particular the role of speed will be tested in Experiment 3 and the role of form cues will be explored in Experiments 4 - 7. In Chapter 6 (Experiments 8-11) a number of control studies will be reported that charts some cognitive aspects in Experiment 1.

Chapter 2 - Perception of Emotion from Motion

Introduction

The principal aim of this chapter will be to introduce the first experimental results for research that was conducted as part of this thesis. In the first part of the chapter the methods for motion capture, movement processing, stimulus generation and experimental procedures will be detailed. Following this will be Experiment 1 – the perception of emotion from human movements tested in a categorisation task. In the analysis of results from Experiment 1 the concepts of a psychological space and multidimensional scaling will be introduced as a means by which to analyse data from a confusion matrix.

General Experimental Methods

In this part of the thesis the general methods that were used in the recording and processing of human movement data and in behavioural experiments, will be described. The purpose of this section is to minimise the amount of repetition in following experimental chapters. For each experiment there *was* variation in the methods and these are described in later chapters, where necessary. However, the information in this part of the thesis should be regarded as the default wherever details are missing from the other experimental chapters.

All of the experiments that have been reported in the current thesis were undertaken in the Biological Motion Lab in the Department of Psychology at the University of Glasgow.

In most of the experiments that are been reported in the following chapters, digital animations of human movements were used as stimuli to which human observers were asked to make judgements. In the execution of such experiments, there are a number of distinct stages and the general methods have been arranged into three parts to reflect and detail these. Firstly the process for digitally capturing the joint movements of human actors will be described. Following there is a description of how the captured three-dimensional (3-D) movement data was processed to make animated sequences for use in experiments. Finally the presentation of stimuli, along with the methodological and procedural considerations that were replicated across different experiments, are described.

Motion Capture

A large part of the preparation of stimuli was the digital capture of human movement data. This process, while relatively straightforward, had to be carefully controlled. A major issue with respect to the motion capture sessions, was the need to make the data measured from different actors and for different emotions as consistent as possible. This was done to ensure a correspondence between different sets of data for later comparison.

Equipment for Motion Capture

To show human movement devoid of form cues, the movement of an actor's joints are isolated and recorded for presentation. Our method for isolating joint movements from form cues was to use an Optotrak (Northern Digital Inc.) for digital motion capture.

To distinguish the joints of a person so that only this motion may be captured, small infrared emitting diodes (IRED's) were attached to a part of the actor's skin in the vicinity of the joint. For the duration of this thesis I will refer to the optical lenses as either "markers" or IRED's. The Optotrak was used to distinguish x, y and z co-ordinates of point sources of infrared light emitted by the IRED's which were then digitally recorded. The point sources of infrared light were brighter than their surroundings and IRED's were adjustable to different conditions of ambient infra red light intensity.

The IRED's were wired to the system hardware and tracked by 3 infrared sensitive cameras mounted on a rigid base. This arrangement meant that no additional synchronisation needed to take place between the IRED's and cameras. The benefit of this was that if a marker became occluded during a recording session[2], it was automatically recognised by the system once it reappeared. Actors' movements were generally captured at 60 Hz (60 frames per second).

Recording sessions were controlled via a personal computer interface with software supplied by Northern Digital Inc. The software allowed online monitoring of marker positions, control of voltage input to IRED's and definition of motion files including the length of recording time, labelling of motion files and specification of a sampling frequency. Finally, the Optotrak system also provided spatial resolution to 0.05mm accuracy at 2.5m from the cameras.

The outputs of the Optotrak system were binary files representing motion data. These files were used to make animations of the movements and to measure kinematics of the movements.

Action and Emotion Selection

Actions were selected to be everyday movements that could be easily captured from only arm movements. Arm movements were selected as potential stimuli since this reduced the information available to observers. It was desirous to do this in order to prevent ceiling effects in the results from our behavioural experiments. Additionally arm movements provide only one end effector for calculating kinematics. Knocking, drinking and lifting movements were selected as being ordinary movements that could be performed in a variety of styles.

For recognition of facial emotion, 6 basic emotions have been identified that are recognisable across different cultures (Ekman, Friesen, & Ellsworth, 1972). Of these, angry, afraid, happy and sad were included in the list of motions recorded from movements. Disgust and surprised were not included in our list as they could not be easily communicated via ordinary movements, most probably because they are fleeting emotions with characteristic movements associated to them. In order to provide a baseline condition for perception of movements without emotion, neutral movements were also collected. The range of stimuli was extended by including 5 additional types of feeling – excited, relaxed, strong, tired and weak.

Actor recruitment

Dependent on the type of movements that were being recorded, a recording session could last from 30 minutes to an hour. After a pilot session had been recorded to establish a procedure and order for recording

2 This occurred whenever something passed in front of the marker or it was obscured form the cameras in another way. A marker could only be detected by the cameras when

movements, right-handed volunteers were recruited as actors. Since movements were always recorded from the right saggital viewpoint, it was decided to use only right-handed individuals to minimise the risk of inconsistency in the motion data. In cases where the movements of a large number of actors were recorded (generally more than 5) the actors were recruited from the student population at the University of Glasgow and paid for their time. For smaller sets of data, actors were friends and colleagues. At the start of a recording session, each actor was given a short introduction to the equipment and procedure for recording movements while markers were placed. The situation resulted in slight physical discomfort and to assure the continued wellbeing of novice actors during a whole recording session, actors were encouraged to make comments and ask for breaks or to leave the experiment at any stage. Although actors made use of the offer to take breaks they never left a recording session before all the data had been collected.

Recording session

To record arm movements 6 IRED's were fixed to the actor's skin with adhesive tape and double-sided adhesive stickers. Figure 2.1 shows the marker locations on the head (close to the ear), the shoulder, elbow, wrist and the second (index finger) and fifth (pinkie) metacarpal joints on the hand. An actor stood about 2m from the Optotrak cameras and Figure 2.1 provides an example of the type of view available to the Optotrak cameras.

Actions were also selected to ensure actors would be able to move comfortably and normally while ensuring that IRED's were visible to the cameras. Once the IRED's were adjusted so that all 6 were visible to the Optotrak cameras, the actor was asked to be consistent in their movements and to not change their movements or stance significantly. This was done to ensure a minimum of missing data. Interpolation can compensate for some missing data; however, a large number of missing data points would result in an artificial movement. Missing data can be particularly problematic when there is a missing chunk of data at the same place in every recorded movement. The best way to prevent too much missing data was to instruct actors to try a make sure that markers always faced the cameras. To further prevent large amounts of missing data, online monitoring and corrections took place throughout a recording session. Before each set of movements - for instance 10 *neutral* knocking movements – an actor was asked to make 2-3 practice movements to test the visibility of markers.

IR illumination was directed from it towards the cameras and there was a clear line of sight between the two.

Wherever possible, different actors were encouraged to stand in the same place and, throughout a recording session, an actor was expected to remain consistent in their movements.

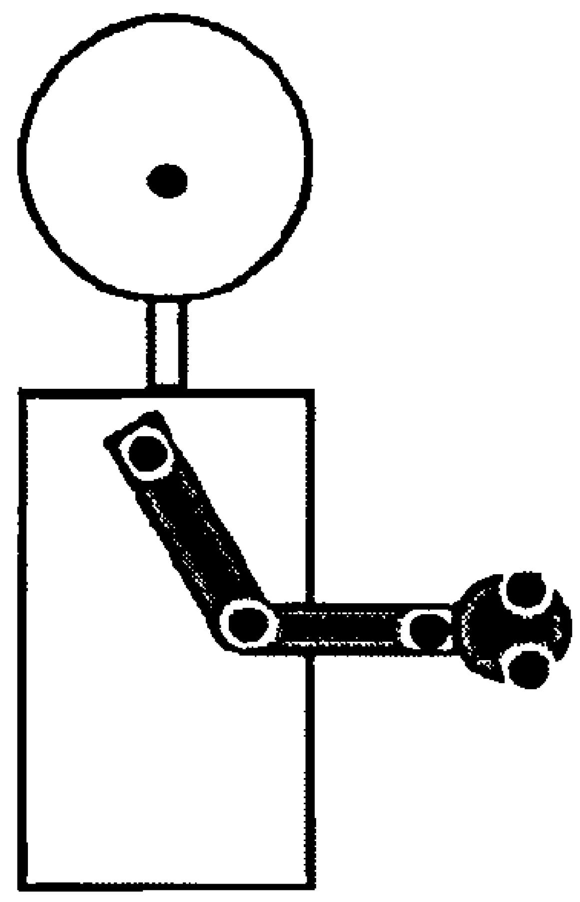

Figure 2.1 Schematic of the right saggital view of an actor with 6 IREDs attached to his right arm. Marker locations are on the head, shoulder, elbow, wrist and second and fifth metacarpal joints.

In order to ensure correspondence between movements, actors were also instructed in the kinds of movements they should make. For instance when making knocking movements, they were told to knock on a box[3] three times, at eye level and with the back if their hand towards their face. When lifting a bottle out of the way they were told pick it up from the side rather than the top and move it from a cross, marked on the table to a cross marked on a box. This procedure ensured that it would be possible to compare movements with different emotion and movements of different actors.

Before movements with emotion were recorded, the actor read a short script to set the scene. Amaya, Bruderlin and Calvert (1996) developed these scripts for recording emotion movements and a full reproduction of scripts is available in Appendix I. Below the scripts for knocking movements with *neutral*, *angry*, *happy* and *sad* emotion has been reproduced as an example.

Neutral Script
It is a sunny Saturday morning. You are well rested, you just had a *relaxed* breakfast and now you are on your way back from the corner store where you picked up a newspaper. You walk up to the front door of your house and knock three times for your spouse to come down and open it.

Angry Script

3 There were various props used for the recording sessions. One of these was a box on a filing cabinet for actors to knock on as a door.

Today you got up a bit late, so you had to rush to get ready. Then on the way to work, a cop flags you down and gives you a speeding ticket, although you were just keeping up with traffic. You finally get to work where a note is waiting for you denying your request for having Friday off. Now you are steaming and ready to tell it to your boss. You walk to his office and knock on his door three times.

Happy Script
It's Friday evening, you are sitting at the dining room table leafing through you favourite magazine. You feel great, because you finished your project at work today on time. Your boss was very pleased, he complimented you on it and said you could take Monday off. You just talked to your friend and neighbour on the phone who invited you over to watch a video. You joyously walk next door and knock on the door three times.

Sad Script
You are sitting at the dining room table trying to finish up dinner, but you are not hungry at all. You don't feel like eating, your stomach is heavy and seems to be bearing all of your body's inner activity and weight. You just move the food from one spot on your plate to another. Half an hour ago, you received a telephone call that your best friend just died in a car accident. You know that you should tell your neighbour, who also knew your friend. Deeply grieving, you walk over and knock three times.

Each recording started with the actor at rest, their hand and arm relaxed at their side in a comfortable stance. A few moments after digital recording started, the experimenter indicated that the movement should begin. This ensured that the start of the movement was captured. The actors were instructed to return their arm to rest once the movement had been completed, and remain still and with their eyes forward until the digital recording finished (normally 1-3 seconds after the end of the movement). Ideally two experimenters were present during each of the recording sessions. One to worked the Optotrak and checked that markers were not occluded from camera view, while the other watched that the actor followed instructions, started at the right time, returned to a resting position and remained in "character" throughout the recording.

Movement Processing

Once movements had been recorded, they were processed to measure properties of the actions and emotions (these will be dealt with in Chapter 3) and were made into animations for display in behavioural experiments. The following sections will deal with the conversion of Optotrak files to IVis[4] files. These steps relied heavily on

[4] IVis is customised 3D graphics visualisation software written in Open Inventor for SGI. We have used the software to display human movement data as point-lights in experiments and to make bitmaps of movement frames for later editing into movies.

programs that were written in the Matlab[5] programming language and so, where necessary, code from these programs has been adapted to illustrate the details of the movement processing.

There were 5 steps in converting movement data to animations. The first was to pre-process the movement data to convert Optotrak files to arrays of movement data for use in Matlab. The second step was interpolation of any missing data. Next the start and end-points of the movements were found and motion files cropped accordingly. Experiments were presented in custom software (IVis) and hence the cropped Matlab arrays were re-organised as IVis motion files and saved for later use.

Pre-processing of Optotrak files

Captured motion data were output to binary Optotrak files that cannot be easily viewed and manipulated. Hence the first step was to isolate the relevant data from such files. In order to do this a Matlab procedure was developed that essentially read the binary formatted Northern Digital file and rewrote it into a usable array of motion tracks, organising the data according to the changing position for each marker over time. Each frame of the captured data was represented in the new array by a row. The position points are represented as three columns, one column each for the x, y and z components of the marker. So, for a typical recording session that lasted for 6 seconds at 60 Hz and in which 6 markers were recorded, the data array would have 360 (6 seconds x 60 Hz) rows and 18 (6 markers by 3 xyz components) columns.

At times when markers disappeared from camera view, the data was declared as missing by changing the x, y and z components to an arbitrary value in the output array. This simply made it easy to later identify missing data for interpolation. Interpolation is the estimation of a missing value by taking an average of known values at neighbouring points. It is a very useful tool; however, when too much data is missing the interpolation method can render artificial movements. To combat this a limit was set on the number of data points that may be missing in any given movement. In cases where this criterion was violated, the movements were not used.

[5] Matlab is a high-level computer language for mathematical computing that deals particularly well with data that is arranged in matrices. Programs were written or adapted by the author or by other members of the Biological Motion Lab at Glasgow University.

Start and End-Points

The next stage in movement processing was to extract the useful movement data from the new array of motion tracks. In order to do this start and end-points of the movements were calculated using the tangential velocity of the wrist or one of the metacarpal joints. The hand is the end-effector of an arm movement (Mather, Radford, & West, 1992) and hence is the most dynamic component of a movement. For this reason, the markers on the hand were used by an automatic procedure for locating the start and end points of a movement. A movement was typically defined as starting when the hand began to move and ending when it returned to a resting position. It did not matter whether the position of the wrist or one of the metacarpal joints were used in these calculations, however, throughout this thesis the wrist has been exploited most frequently. In the automatic procedure the tangential velocity of the wrist was calculated and then a criterion of 5% of the peak velocity was set and used to suggest a start and end to the movement. These were then visually examined and corrected online where necessary.

The first step in finding start and end-points was to manually select one of the 3 markers on the wrist and hand. The x, y and z components of the marker were then identified from the movement array. In the simplest example the spatial displacement of these x, y and z components between successive frames, was differentiated to obtain the components of the instantaneous velocity ($\frac{dx}{dt}$, $\frac{dy}{dt}$ and $\frac{dz}{dt}$). The tangential velocity (V_{tan}) is the square root of the sum of squares of the individual components.

$$V_{tan} = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} \qquad \text{Equation 1}$$

The quantity V_{tan} gave the speed of movements over time and was used in all subsequent calculations of movement properties.

The start-point (P_{start}) was defined as 10 movement frames before the first frame at which the velocity of the hand exceeded a pre-set criterion. This criterion was 5% of the peak tangential velocity in the first half of the movement. The procedure located the first of 5 consecutive frames in which the 5% criterion was exceeded and subtracted 10 frames to produce the P_{start}. If this yielded a value less than one, P_{start} automatically became the first movement frame.

The end-point (P_{end}) was defined as 10 movement frames after the last frame at which the velocity of the hand exceeded the 5% criterion. This time the procedure located the last 5 frames in each movement at which the 5% criterion was exceeded. The last of these frames with 10 added movement frames became the P_{end}. An automatic Matlab procedure was used to find the start and end-points and the results were later inspected visually in relation to the velocity profile to insure against possible errors.

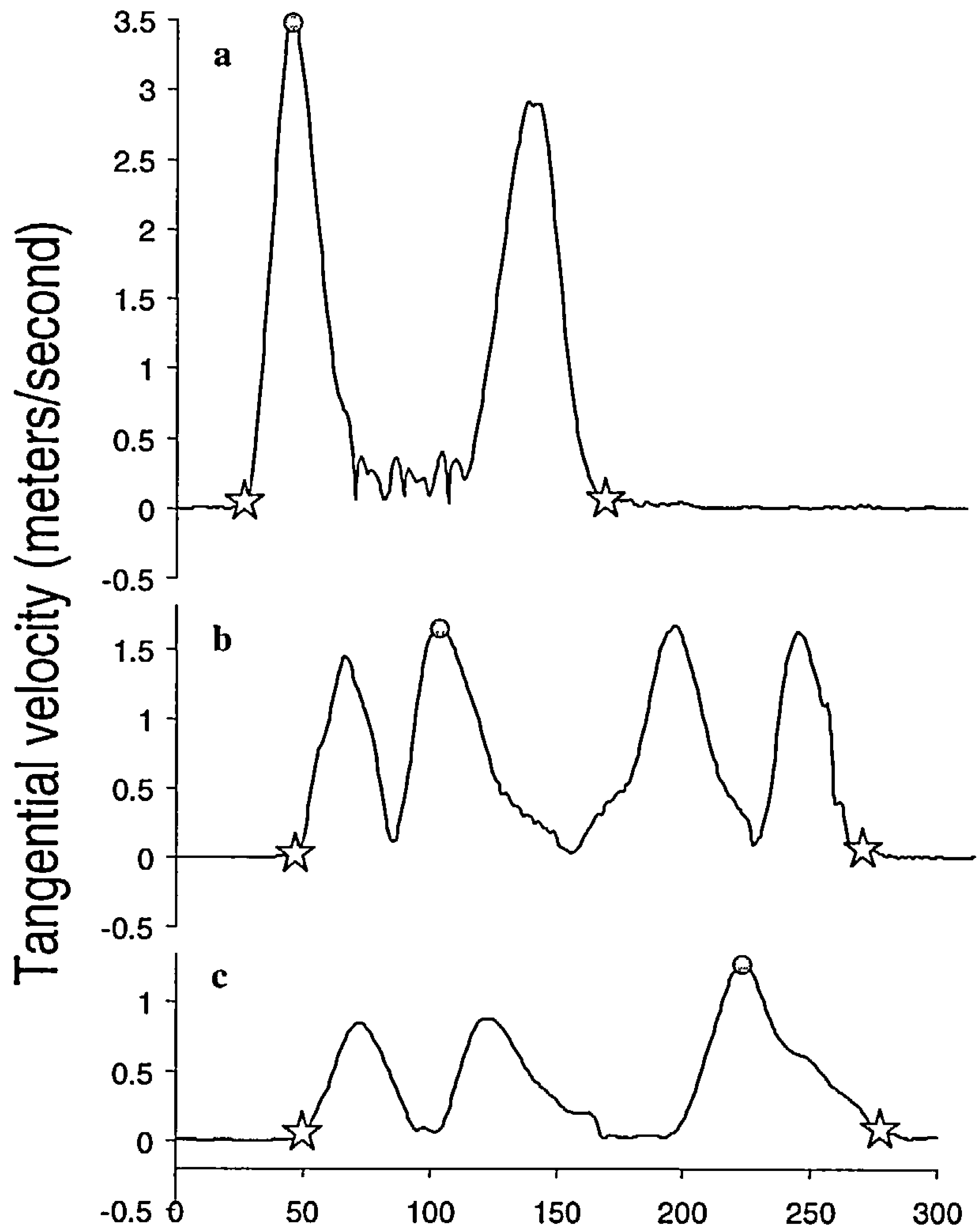

Figure 2.2 The profile of tangential velocity for a knocking (a) drinking (b) and lifting (c) movement. Stars represent start (on the left) and end-points (on the right) for each movement and circles indicate locations of maximum velocity.

In Figure 2.2 are illustrated the tangential velocity profiles of the wrist for a neutral knocking, drinking and lifting movement with arbitrary start-points indicated by stars closest to the y-axis on each graph. Stars corresponding to the end-point for each movement are located furthest from the x-axis and circles indicate the position of the peak velocity.

Once P_{start} and P_{end} of a movement had been identified, a final array of the movement data was constructed that contained only the data within the P_{start} to P_{end} time period. This array was used for all subsequent calculations as well as to make the stimulus animations.

Motion files for display in IVis

The final stage in preparing movements for experimental display was to write the x, y and z motion of each marker as motion tracks into a file format that could be read by IVis Software. IVis has been developed as specialist software to display Optotrak recorded motion data as animations that can be used for experimental and demonstration purposes. The 6 sets of motion tracks were typically displayed as yellow spheres on a black background and movements were played back at 30 Hz. The colours of the background and point-lights as well as the shape of the point-lights were defined in IVis before movement displays started. Within IVis the viewing direction, size and position of a display may also be set and this was done in a consistent manner across subjects and experiments.

Behavioural Methods

In the next sections are described the behavioural methods that were common to most of the experiments in the following chapters.

Participants

Participants for all experiments were student volunteers from the University of Glasgow that were generally paid for their participation. They were also naïve to the purpose of experiments and where possible should not have taken part in previous experiments in the Biological Motion Lab.

Displays

Stimuli were presented on a graphics computer (Silicon Graphics Octane) in a specially designed experimental package named "IVis". The six moving points were presented as yellow spheres on a black background and were oriented to the right saggital viewpoint in an upright position. This resulted in the information from the z-axis not being immediately obvious and to further reduce the effects of 3D cues from stereo and accommodation, participants were seated about 1 meter from the screen.

Experimental set-up

In most cases the specialist software, IVis, was used to run experiments. IVis combines graphical display of movements captured with an Optotrak and some basic experimental designs[6]. At the start of each experiment IVis windows were set up and the dimensions of the display specified. The IVis dimensions that were typically controlled in displays were the size, orientation, viewing direction and the position for each stimulus. In general displays took up about one third of the display window (displays were about 10 cm tall and 2-5cm wide), were positioned centrally on the screen in an upright orientation and were seen from the right saggital view. Point-lights were depicted as yellow spheres on a black background.

Collecting Judgement Data

At the start of their contribution each participant was given a set of typed instructions detailing the procedure and task of the experiment. To make sure that instructions were understood, they were summarised verbally and a short practice session took place to familiarise the participant with the IVis interface, point-light displays and the task. Typically each trial consisted of a viewing a display and making a judgement about it. After each display a small window appeared on the screen in which the participant was expected to indicate their judgement by using the mouse to select an appropriate box (in forced choice paradigms) or by moving a sliding bar (in ratings paradigms). The participant then confirmed their choice and moved on to see the next display. Judgements were recorded electronically.

[6] In cases where IVis' experimental designs were not sufficient, PsyScope was used instead. This was the case only rarely and procedures are explained in the methods sections for those experiments.

Participant Debriefing

At the end of the experiment there was a short interview to debrief the participant. This consisted of pre-determined questions and was used to reject any participants that clearly had not understood the instructions and to gain insight into deliberate strategies a participant may have used. Questions changed from experiment to experiment, however there was some consistency in all the debriefing sessions. Appendix II shows a typical set of questions asked during a debriefing interview.

Reasons for participant rejection

Participants were occasionally rejected when it became clear during an experimental session or in the debriefing session that they did not understand the experimental instructions or found the task too difficult. The reasons for rejection were 1) that a participant was not computer literate to the extent that they could not use a mouse to make judgements (these tended to be older, non-student participants). 2) That a participant became ill during the experimental session (this happened on one occasion when the participant had a severe hangover). 3) That they did not understand the instructions after the practice session had taken place and been repeated once. 4) That a participant was unable to resolve the point-light displays as a person or found the task too difficult and wanted to leave the experimental session. 5) A final reason for rejection was if a participant had taken part in an experiment using biological motion previously and during the debriefing session indicated that they had been influenced by their previous experience.

Experiment 1 - Perception of Emotion from Motion

Introduction

In the following section a behavioural study in which movements were shown to observers with the task of identifying the emotion depicted in each movement, will be described. These judgements made by observers were analysed to obtain the proportion correct responses and were then used to construct a psychological space for the perception of the movements.

Critical to the development of this thesis had been research in the computer animation literature by Amaya, Bruderlin, and Calvert (1996). In this paper Amaya et al made a detailed computational analysis of expressive human arm movements and proposed a transfrom that would change *neutral* movements into expressive ones. The basis of their procedure was the assumption that an emotion, as represented in motion, is a secondary movement that is piggybacked on the primary movement (the basic action of drinking or knocking for instance). Furthermore, the primary and secondary movement types are measurable and can be used to reconstruct an emotional movement from a *neutral* one by calculating an emotional transform. In computational terms the emotional transfer is basically the motion signal that is left after a *neutral* movement has been subtracted. To calculate the emotional transforms Amaya et al (1996) concentrated on two components in the movements that were posited as modulating such an emotional transform; the speed of a movement and its spatial amplitude or range.

Their first step was to capture the movements of two actors while they performed drinking or knocking movements with emotion. Using *sad*, *neutral* and *angry* movements, Amaya et al (1996) established that speed and spatial amplitude are different between these movements. Next the transforms of speed and spatial amplitude were calculated and applied to *neutral* movements to be compared with original, captured movements. The procedures were successful since *neutral* movements to which the *angry* or *sad* transforms had been applied, resembled captured *angry* and *sad* movements. The transforms were also successfully transferred to animated movements and to captured human movements of a different action.

From a psychological perspective the conceptualisation of emotions as secondary signals encoded in deviations from a normal signal is reminiscent of the concept of transformational and structural invariants (Shaw, McIntyre, & Mace, 1974). The emotional transform would represent a structural invariant that can be revealed by the transformational invariant or the primary motion signal in the same way as identity can be revealed through a point-light display of walking (Cutting and Kozlowski, 1977).

Amaya et al (1996) showed that movements with emotion are measurably different from each other from a purely quantitative perspective. Since variations between emotions are quite large in some cases and form a stable pattern across actions and actors[7], it stands to reason that humans would be perceptually sensitive to them and that they would utilise this variation to understand the emotional meaning in a movement. The following experiment was designed to establish whether this is the case.

Knocking and drinking movements with 10 different emotions were presented to 14 observers and the ability of observers to categorise emotions was measured as proportion correct answers. The perception of emotion was further examined within the framework of a psychological space.

Methods

Participants

Point-light displays of knocking and drinking actions with emotion were displayed to fourteen Glasgow University student volunteers. Participants were naïve to the purpose of the experiment and were paid for their participation.

Stimuli

The movement data recorded by Amaya, Bruderlin and Calvert (1996) were converted to a visual display and shown to observers on a graphics computer. Each display consisted of 6 yellow dots moving on a black background. The dots represented the 6 marker locations from the motion capture data so that the appearance was of a moving human arm and head from the right saggital viewpoint.

Two actors performed knocking and drinking movements with 10 different emotions while their movements were recorded. The emotions were afraid, angry, excited, happy, neutral, relaxed, sad, strong, tired and weak. After reading a short script that set the scene for each emotion, an actor knocked on a door three times or lifted a glass with liquid from a table and took a sip before replacing it. Each movement was repeated three times while the movement of markers on the head, shoulder elbow, wrist and hand were recorded. This procedure

[7] See Chapter three, for a quantitative analysis of the movements used in Experiment 1.

yielded a total of 120 movements (10 emotions x 2 actors x 2 actions x 3 repetitions), but due to recording difficulties data was lost for 2 of the movements[8] of one actor.

Design

Stimuli were divided into 4 blocks according to the possible combinations of actor and action (2 actors x 2 actions) and were displayed to observers in a random order. In each block there were the 30 movements (28 in the case of drinking) of a particular actor making one action. These 4 blocks were the experimental blocks. There was also a practice block of four *neutral* movements. These *neutral* movements were randomly chosen - one from each experimental block. In the Experimental blocks each movement was repeated three times so that 90 stimuli (10 emotions x 1 actors x 1 actions x 3 exemplars x 3 repetitions) were displayed in each experimental block. Movements were presented in a random order within both experimental and practice blocks and the order of blocks was randomised across observers.

Procedure

The experiment took place over two 45-minute sessions on the same day and with at least an hour's break between them. Participants were told that they would see knocking and drinking arm movements and that they had to identify the emotion being expressed in each movement. In the first session participants were given a practice session of 4 trials - one from each block of trials - to familiarise them with the stimuli, equipment and procedure, this was followed by the first two blocks of trials. The second session consisted of the two remaining trial blocks and a debriefing interview.

At the start of the experiment the participant was presented with a set of written instructions that detailed the procedure. In the instructions the emotion labels were detailed, but not defined. Non-native speakers of English were asked if any of the terms were unfamiliar to them. Asking non-native English speakers to try and define the terms tested their familiarity with the emotion labels[9]. In the instructions it was also emphasised that

[8] In this corpus there were two missing movements, both were drinking movements from the same actor (Actor 2), one tired movement and one relaxed movement.
[9] This was an informal test and all non-native speakers of English found it easy to define the terms.

there would be *neutral* movements and that participants should not use this as a means of indicating that they could not recognise the emotion. The instructions were repeated verbally before the practice session.

In the practice session observers viewed each movement while the experimenter checked that they could resolve the stimulus into an arm movement and that they could distinguish knocking from drinking movements. During the practice session observers also became familiar with the graphics layout on the computer screen and with using a computer mouse to select their response. Since the experimental design did not allow the observer to skip trials, it was stressed at this stage that a response should be made for each trial.

For each trial the participant viewed the computer display of a movement and was then presented with a dialog box that contained the names of the ten possible emotions. Their task was to choose a single emotion by clicking the mouse on the matching label. After completing all trials in the experiment there was a short debriefing interview.

Results and Discussion

Correct Identification

In the first instance data was analysed to obtain the percentage correct responses and this analysis is detailed in the following section.

Over all the trials participants answered correctly 30% of the time; ranging from 15% for movements with *strong* emotion to 50% for movements with the emotion *afraid*. The 30% overall correct identification was significantly better than the chance value of 10% [$t(13) = 20.3$, $p < .005$, two-tailed]. Figure 2.3 shows the proportion of correct identifications for each emotion averaged over movement exemplars, actions and actors. The two best-recognised emotions were *afraid* and *excited* followed by *angry*, *happy*, *neutral*, *relaxed* and *tired*. The emotions that were least well identified were *sad*, *strong* and *weak*.

In Figure 2.4 the proportion of correct identifications for each emotion has been plotted according to actions and actors in order to show the range of correct recognition for different emotions depending the different actors that performed them and the different actions that were used to convey them.

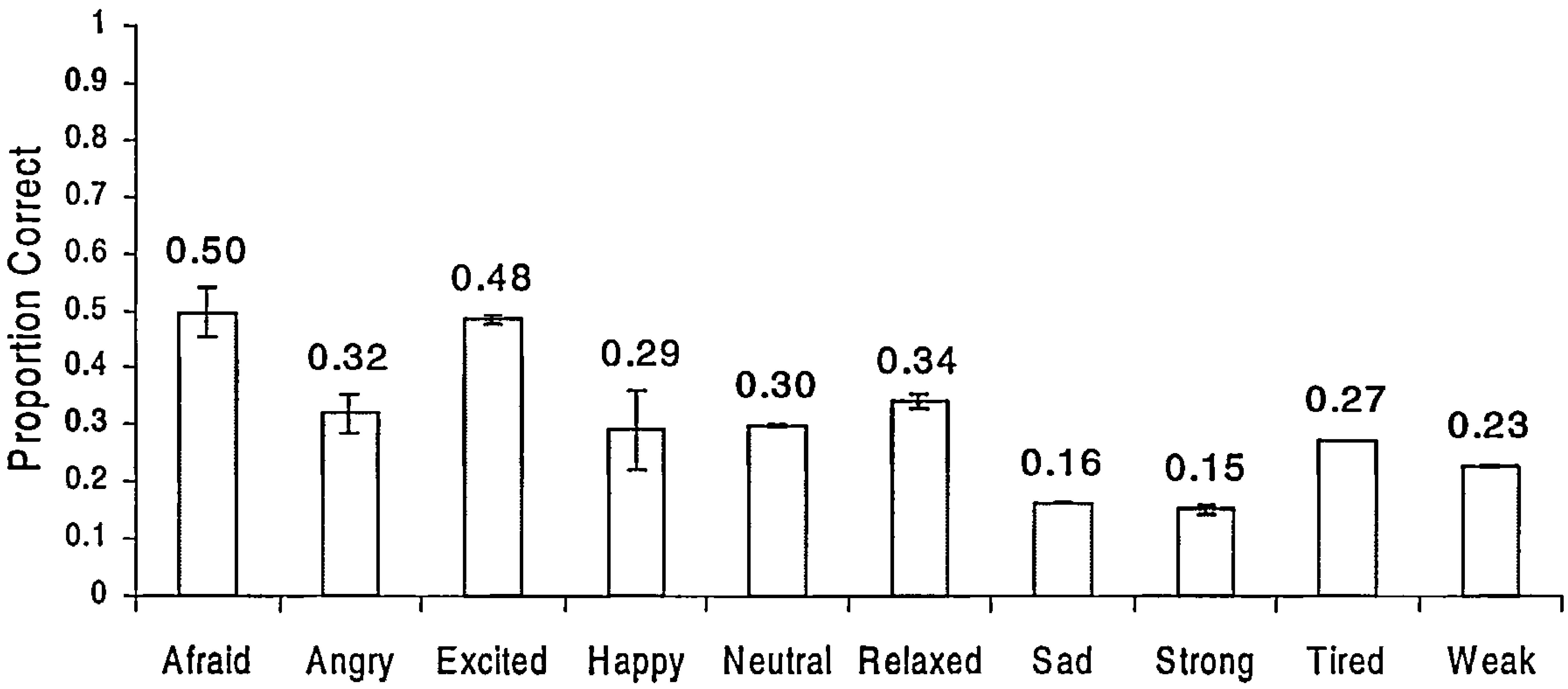

Figure 2.3 Mean proportion correct identification of emotions. For each emotion the mean is calculated by averaging over actor, action and movement exemplars and is displayed above the appropriate bar. Error bars show standard errors for each mean.

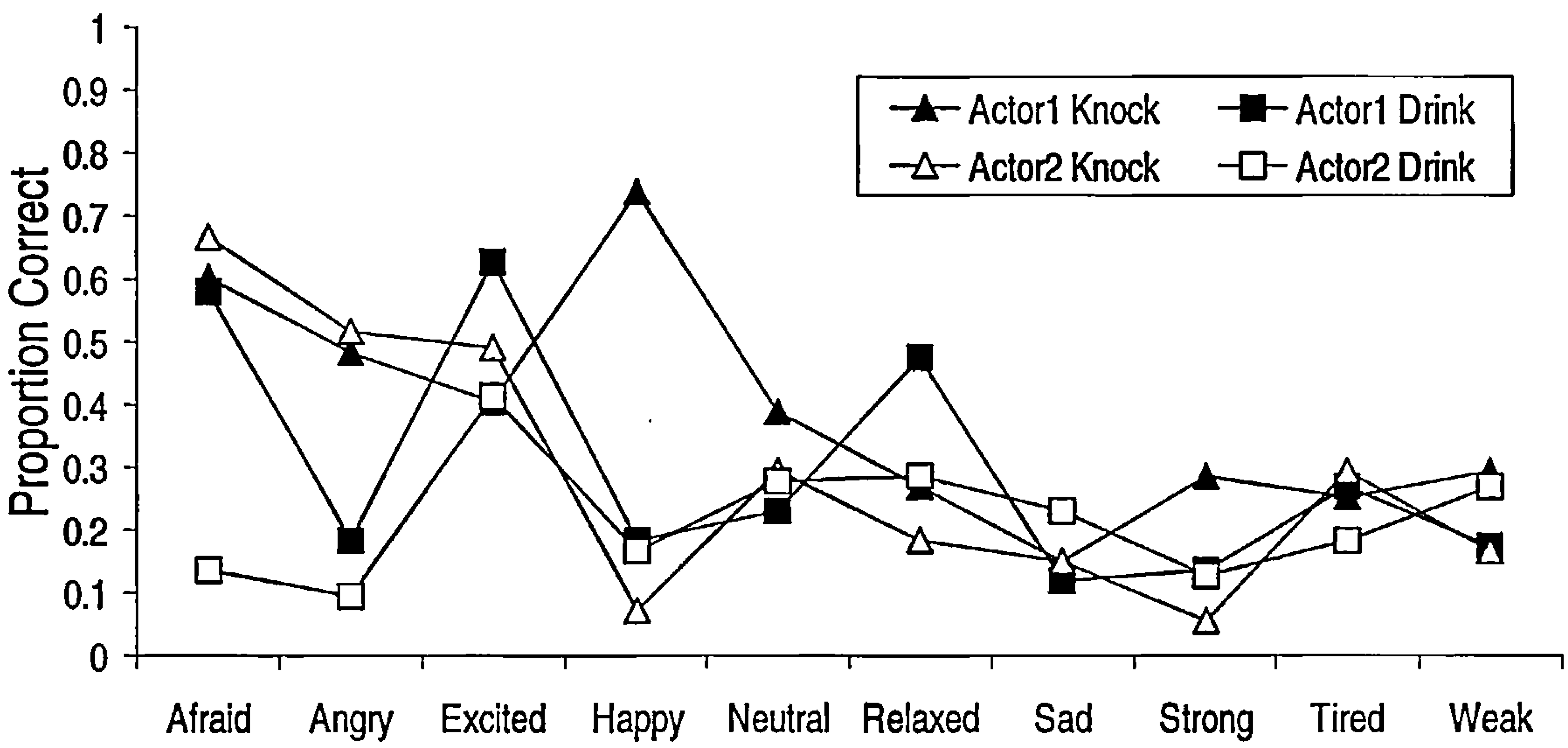

Figure 2.4 Proportion Correct identification plotted according to actor, action and affect. There seems to be little consistency in the proportion correct identification of affects for different actors and actions.

It would seem from Figure 2.4 that observers were sensitive to individual differences of actor and action in the movements. These individual differences meant that some movements were more easily identified than others were. For instance the *happy* knocking movements of Actor1 were easily recognised compared to the *happy* movements from Actor2 and when Actor1 was drinking. Similarly the *angry* movement conveyed with drinking actions seemed harder to recognise than the *angry* movements from knocking. *Tired* movements were recognised equally well for all types of movements and for movements with the emotion *afraid*, Actor 2's drinking movements conveyed emotion less well than *afraid* movements of any other type.

These observations were tested in a 2 (actor) x 2(action) x 10 (emotion) repeated measures ANOVA of the proportion correct recognition of emotions.

Results from the ANOVA showed significant main effects for all three factors [Actor: $F(1,13)=19.01$, $p<.001$; Action: $F(1,13)=27.343$, $p<.001$; Emotion: $F(9,117)=12.033$, $p<.001$]. A closer look at the means for each factor showed that the emotions acted by Actor1 (34.3%) were more accurately recognised than those of Actor2 (25.4%) and emotion was better recognised from knocking actions (33.8%) than from drinking actions (25.8%). Post Hoc Tukey tests of the honestly significant differences (HSD's) between emotions showed that *afraid* and *excited* movements were better recognised than all other movements at $p<.01$ ($p<.05$ for *angry* movements), but there was no significant difference between *afraid* and *excited*. Angry movements were recognised with greater accuracy than sad and strong movements at $p<.05$.

The ANOVA also showed significant interactions for the two-way interactions: actor x emotion [$F(9,117)=7.616, p<.001$] and action x emotion [$F(=9,117)=8.367, p<.001$], but not for actor x action [$F(1,13)= .623$, $p=.444$]. There was a significant three-way interaction between actor, action and emotion [$F(9,117)=14.269, p<.001$]. The two-way interactions reflect effects in the three-way interaction, so further analysis for the three-way interaction has been discussed below.

Three-way Interaction between Emotion, Action and Actor

The three-way interaction was more closely examined and **showed simple main effects of action** for Actor1's movements with *angry*, *excited*, *happy* and *relaxed* emotion and for Actor2's movements with *afraid* and *angry* emotion. In most cases knocking movements were better recognised than drinking movements (the exceptions were for actor1's *excited* and *relaxed*) as reflected in the simple main effect of action.

Apart from the individual differences between actions; there were also individual differences in the accuracy with which emotions were recognised dependent on the actors that performed them. **Simple main effects of actor** were found for knocking in *happy* and *strong* emotion and for drinking at *afraid*, *excited* and *sad* emotion. In most of these cases Actor1's movements were better recognised than Actor2's[10] movements. There was also a general trend for Actor1's movements to be better recognised than Actor2's movements as evidenced in the main effect of actor described above.

Introspectively, in our everyday experience, it is also the case that we seem to more easily interpret the feelings expressed by some people while the expressions of others are not as easy to recognise. Although it is not standard for individual difference of actors to be reported in papers about biological motion there are a few cases in which such data was discussed.

Sumi (Sumi, 2000) reported individual differences between actors in their walking style and similarly Cutting and colleagues found that the sex (Kozlowski & Cutting, 1977) and identity (Cutting & Kozlowski, 1977) of some actors were better recognised than others. As a future project it may be interesting to more carefully examine the differences in recognising emotions from different actors in the context of finding invariant properties in actions that may inform perception of emotion.

There were **simple main effects of emotion** at all levels of actor and action at $p<.001$ apart from drinking movements of Actor2 where the effect was significant at $p<.05$. Post hoc analysis of the HDS's between emotions revealed that in most cases the pattern of differences followed that found in the main effect of emotion so that in general *afraid* and *excited* movements were better recognised than other movements the exceptions to this are discussed below.

The knocking movements by Actor 1, were generally well recognised and the post hoc analysis of HSD's showed the following. *Afraid* movements were better recognised than *relaxed*, *sad* and *tired*. *Angry* was better recognised than *sad*. Additionally, *happy* movements were exceptionally well recognised and the HSDs between *happy* and other emotions were significant, excluding *angry* and *afraid* movements. *Happy* movements were poorly recognised in the other movement conditions and the reason for the good recognition in of Actor 1's *happy* knocks was probably due to these movements having a distinct and comical rhythm. Emotion from drinking movements for Actor1 was well recognised for *afraid* and *excited* movements compared to other emotions, but *angry* movements were poorly recognised and *relaxed* was well recognised.

For Actor 2, knocking movements *afraid, angry* and *excited* were recognised best. Drinking movements by Actor2 were poorly recognised and there were few significant HSDs between any emotions. It is worth noting, however, that *afraid* (well recognised in all other conditions) was poorly recognised, as were *angry* movements, while the best recognised were *excited* movements.

Finally, there was one consistent change in recognition accuracy between actions and this occurred for *angry* movements. From observing knocking movements, participants were much more successful at recognising *angry* than when they viewed drinking actions. A reason for this may be that for drinking actions actors were expected to control a cup filled with liquid. This made it difficult to freely act since actors had to ensure no spillage of the liquid occurred. In knocking movements the *angry* movements appeared to be very vigorous and limiting the ability of actors to make vigorous movements in the drinking action is probably what caused them to be less successful at conveying this emotion.

In summary, the results from this analysis show that on average observers could correctly recognise emotion from point-light displays of human arm movements significantly above chance. However, there were also individual differences in terms of which emotions were being communicated, which actor was performing an emotion and what action was used to convey the emotion. Particularly characteristic actions – such as a comical knock depicting *happy* – seem to have an effect over recognition accuracy. In general, however, there is variation in the way different individuals depict emotion. For this reason it would be interesting to find a consistency in the way emotion is dealt with by observers that supersedes the individual differences between actors and actions.

---

[10] The exception was for Actor2's *sad* movements.

Consistent Misidentifications

The task for participants in Experiment 1 was to choose a label from 10 possibilities that would correspond to the movement they had see. This made it possible to collect confusion data for each observer and these confusions were inspected for any consistent misidentifications.

We expected that *neutral* and *relaxed* movements would be confused regularly since both the movement and emotion of these seemed similar. This was indeed the case, with *relaxed* being labelled *neutral* 23% of the time and *neutral* being mistaken for *relaxed* 33% of the time. For *relaxed* labels of "relaxed" and "neutral" accounted for 63% of date collected for the movements, for *neutral* these labels accounted for 57% of data. Apart from "tired", no other labels were chosen for either of these emotions at levels above chance and *tired* was used to describe the movement 13% and 14% of the time respectively for *relaxed* and *neutral.*

In the data there were also other consistent misidentifications such as between *sad, tired* and *weak.* Confusions amongst these three accounted for 52% of all data collected for *sad,* 62% for *tired* and 69% for *weak* movements. Similarly *angry* and *excited* movements were often confused for each other and these confusions accounted for 44.6% and 73.8% of data respectively. Finally *happy* movements were confused for *relaxed* and *neutral,* with responses of "happy", "neutral" or "relaxed" accounting for 74.8% of all data collected for *happy* movements. *Strong* was also confused for *relaxed* and *neutral* so that these labels and "strong" accounted for 62% of data collected for *strong.*

Another feature of these confusions was that they seemed fairly consistent for all the different types of movements. For instance, for all four types of movements there were high rates of confusion between *sad, tired* and *weak,* and also between *neutral* and *relaxed. Happy* movements were confused for *neutral* and *relaxed* in all cases apart from Actor1's knock, which was very well recognised, which left little room for confusion with other emotions. The consistency in confusions suggests that there may be some common underlying pattern in the way emotions were perceived from arm movements made by different actors and with different actions.

It is difficult to see patterns in the movements when comparing 10 x 10 confusion matrices. Re-conceptualising the confusions between emotions as distances between two emotions allows the use of the multivariate statistical methods known as Multi-Dimensional Scaling (MDS). This analysis would allow us to find any underlying or hidden pattern in the recognition data. The MDS procedure also represents any structure in the data as a geometric model that can be though of as the psychological space of the representations of emotions from perception.

Psychological Space

To better understand the structure of the data, a psychological space of the emotions was constructed using the INDSCAL multidimensional scaling procedure (Kruskal and Wish, 1978).

In order to do this the 10 x 10 confusion matrix for each of the four types of movement was converted to measures of dissimilarity. The four types of movements were used rather than individual observers, since we were interested in establishing that perception of emotion from different actors and action would utilise a similar psychological space. Dissimilarity was calculated between each pair of emotions by first averaging the two cells from the confusion matrix that corresponded to the pairing. To change this average proportion of confusion between two emotions to a distance between two emotions, the value was then subtracted from 1. If the two emotions were never confused, the average confusion would be 0 and the dissimilarity 1. For instance, the dissimilarity between *angry* and *tired* movements would be calculated as follows. The proportion of times *angry* movements were identified as *tired* and the proportion of times *tired* movements were identified as *angry* would be averaged. This average value would then be subtracted from 1, yielding the dissimilarity between movements.

The new confusion matrices of dissimilarity were input to the INDSCAL multi-dimensional scaling algorithm (Kruskal and Wish, 1978). Kruskal and Wish (1978) developed the INDSCASL procedure as an extension to the original ALSCAL procedure. The main addition is that with the INDSCAL procedure the weights of different subjects or different conditions are included in the analysis (Davison, 1983) to form a Weighted Euclidean Model of the data. This model takes account of individual differences since a separate confusion matrix represents each subject or experimental condition and the resultant solution does not need rotation. Additionally the importance of each dimension in the Euclidean Model can be quantified.

The resultant solution from the INDSCAL procedure for the four movement conditions was a low-dimensional psychological space (see Figure 2.5) with $r^2 = .874$ and stress = .15 in the 2-dimensional solution and with $r^2 = .928$ and stress = .098 for a 3-dimensional solution. In the 2-dimensional, the first dimension ($D1_{Canonical}$[11]) accounted for 70.9% of the variance and the second dimension ($D2_{Canonical}$) for 16.5% of the variance. The stress, r^2 and weirdness (Kruskal and Wish, 1978) of the 2-dimensional solution for each type of movement can be seen in Table 2.1. Three of the movement types corresponded well with an average confusion matrix and showed low weirdness (by having weirdness values close to 0). Actor 2's drinking movements had a relatively high weirdness and this probably reflects the difficulty participants had in identifying these movements.

Actor2's drinking movements were recognised with 21.8% accuracy, compared with 28.9%, 29.8% and 38.7% respectively for Actor2's knocking and Actor1's drinking and knocking movements. Since the emotions were hard to recognise from Actor2's drinking movements, observers may also not have been as consistent in the way that they confused movements.

The two-dimensional structure of the psychological space appeared similar to that from a circumplex model of experienced emotion (Russell, 1980). The circumplex model of emotion was first proposed by Russell (1980), however other 2-dimensional models of emotion have been proposed that share similar features with this model (Watson and Tellegen, 1985; Larsen and Diener, 1992; Thayer, 1989; for a summary see Yik, Russell and Barrett, 1999). Characteristically these models of emotion have 2 bipolar dimensions that account for most of the variation in participants' responses. Typically, although they may be named differently, the two dimensions represent valence and energy/arousal dimensions of emotion.

[11] In notation for dimensions of the psychological spaces, they are referred to as canonical in order to distinguish them from dimensions for the psychological space for scrambled movements in Chapter 5. No distinction is made between dimensions for the two- and three-dimensional spaces in this analysis since there was correspondence between the first two dimensions for both.

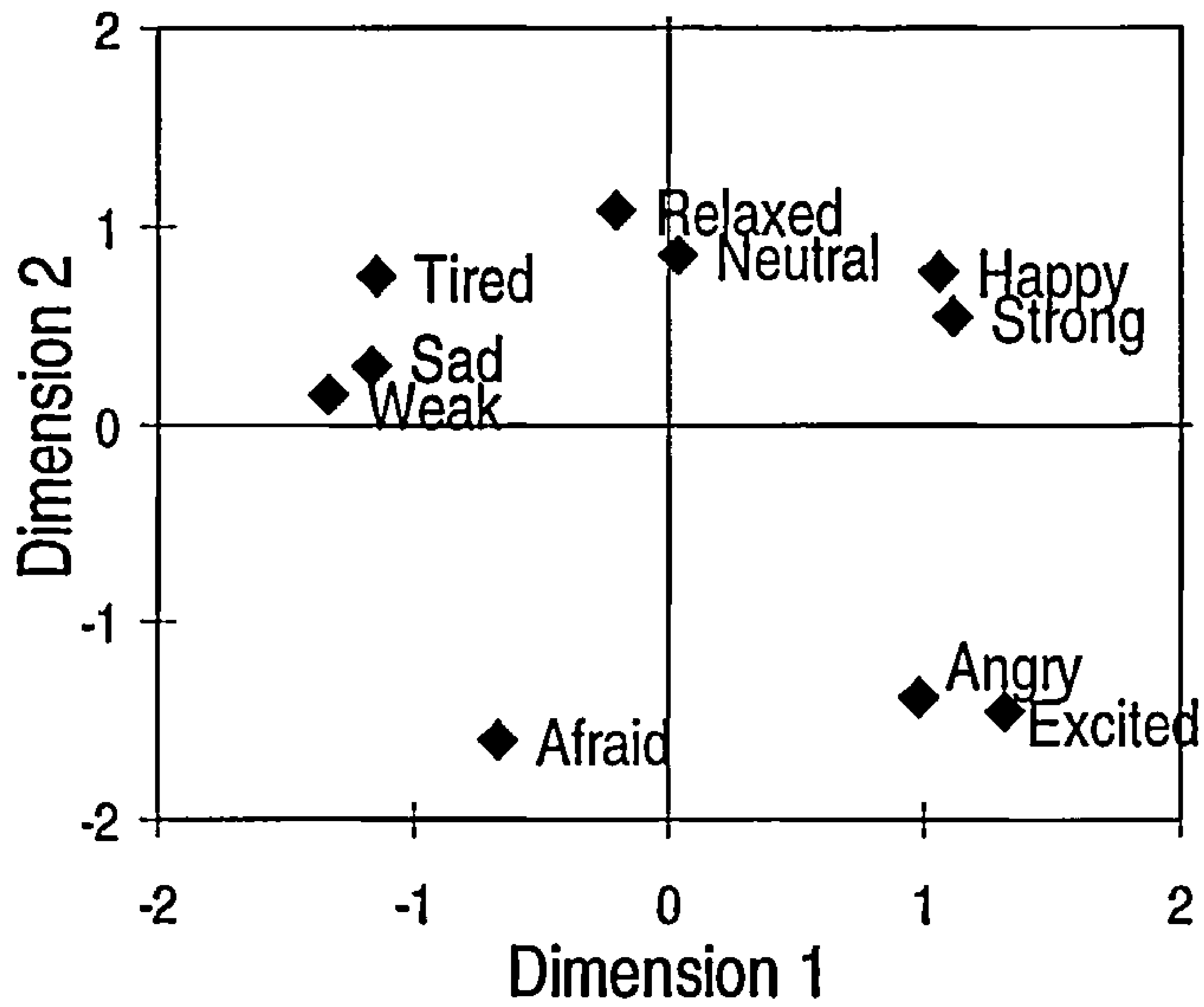

Figure 2.5 Psychological space for Affects

Table 2.1 Stress, r^2 and weirdness for each type of movement in the Euclidean Distance model for the INDSCAL Multi-dimensional scaling procedure.

Matrix		Stress	RSQ	Weirdness
Actor1	Knock	0.146	0.875	0.016
	Drink	0.127	0.904	0.214
Actor 2	Knock	0.152	0.863	0.140
	Drink	0.172	0.855	0.469

The 2-dimensional structure of our psychological space for the perception of emotion from human movements does bear some relationship to the circumplex model. $D1_{Canonical}$ accounted for 70.9% of the data and seems to represent the second dimension from experienced emotion – arousal. Mode dynamic movements such as *angry* and *excited* are at one extreme of this dimension while less dynamic movements such as *tired* and *sad* were arranged at the other extreme. The $D2_{Canonical}$ accounted for 16.5% of the data and seemed weakly related to valence since highly negative emotions such as *angry* and *afraid* were at one extreme of this, while at the other extreme *happy* and *relaxed* were located. *Sad*, *tired* and *weak* are problematic since they do not fall in the negative sphere of the psychological space[12]. *Excited* is also problematic since the script supplied to actors indicated that this should be a positive emotion.

[12] But see the next chapter for a rotation of the psychological space that addresses this issue.

It stands to reason that arousal would be of primary importance in emotion perceived in the context of actions, since a particularly good indication of physiological arousal of an actor would be the speed at which they move. All the observers indicated that they used the speed of movements in their strategies for doing the task and the correlation between movements speed and judgements have been examined in more detail in Chapter 3. On the other hand, valence would be harder to judge and may be represented as an attraction-avoidance dimension. This does seem fit with the 2-dimensional psychological space since *sad*, *tired* and *weak* may encourage an observer to approach the individual displaying them.

In Experiment 1, the variance accounted for by the Euclidean Model that forms the psychological space, and the stress of the fit to this model showed some improvement when a 3-dimensional model was considered [stress = .098; r^2 = .928]. This psychological space has been reproduced in Figure 2.6. It is clear that there is little change in the dimensions that are also included in the 2-dimensional solution. This was tested formally by a correlation between the first two dimensions of the psychological spaces. The first dimension from the 2D and 3D space were correlated with r^2=.965 and the second dimension r^2=.958, indicating that for $D1_{Canonical}$ and $D2_{Canonical}$ there is close correspondence between the two psychological spaces.

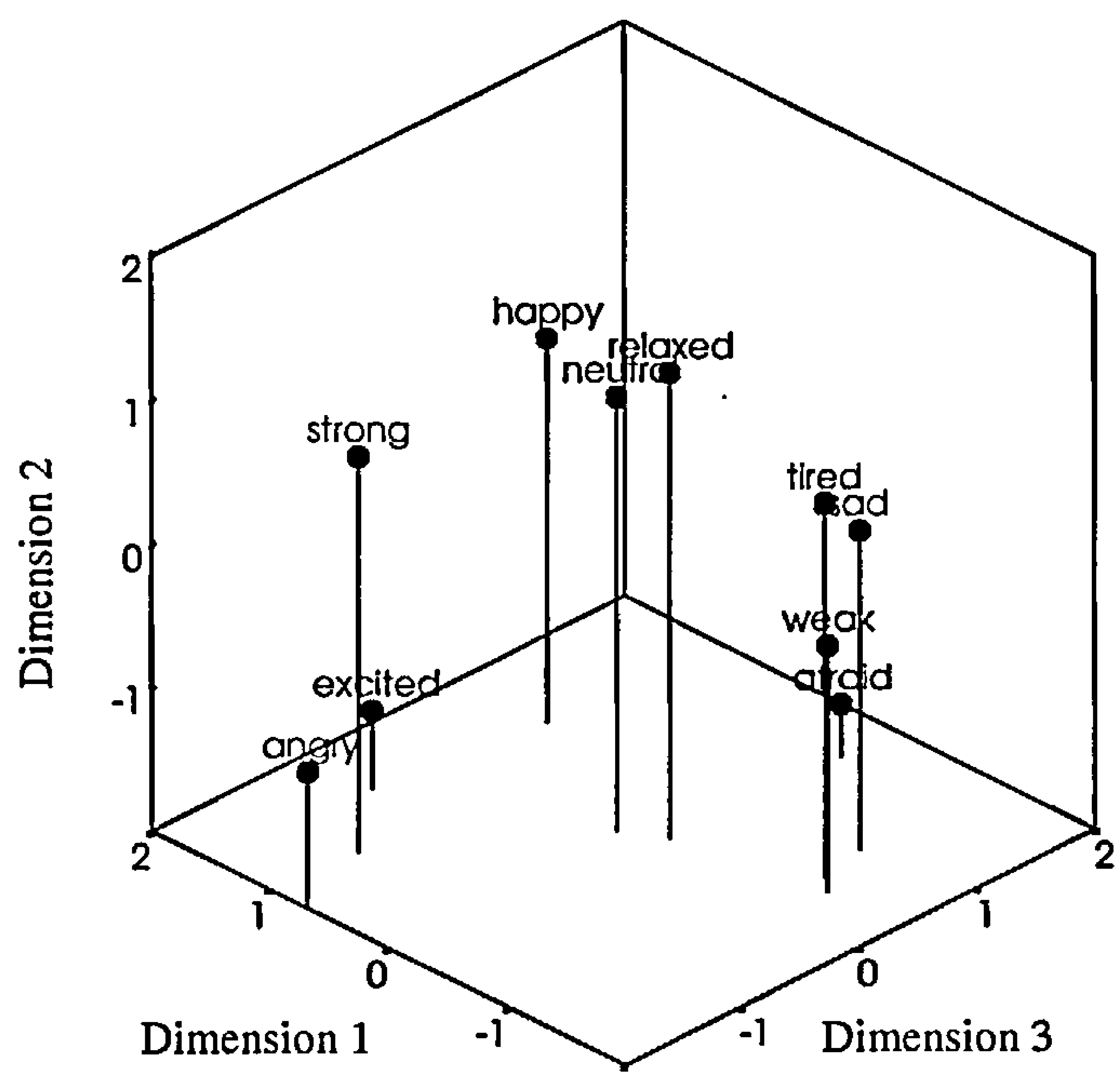

Figure 2.6 Three-Dimensional Psychological space

What of the third dimension? In modelling representations of affective facial expressions, a 3-dimensional solution also sometimes accounts for the data (Tagiuri, 1969). In the research reviewed by Tagiuri, representation of facial expressions was also 3-dimensional with two major dimensions (pleasant-unpleasant and attention-rejection) along with a third dimension (activation) that was always less important than the first two and was related to the second.

In our data the first dimension was more important with $D1_{Canonical}$ accounting for 66%. $D2_{Canonical}$ accounted for 16% while $D3_{Canonical}$ accounted for 11% of the data. If, as with facial expressions, it was also the case in our data that the third dimension was related to the arousal dimension ($D1_{Canonical}$), there should be a correspondence between $D1_{Canonical}$ and $D3_{Canonical}$, but not between these and the valence dimension ($D2_{Canonical}$). In order to test this, a correlation was calculated for each pairing of the three dimensions and it was found that $D1_{Canonical}$ and $D3_{Canonical}$ corresponded weakly [r^2=.28], however when $D1_{Canonical}$ and $D3_{Canonical}$ were compared with $D2_{Canonical}$ there was much less correspondence with r^2=.03 and r^2=.002 respectively. These results suggest that although the first and third dimensions were more similar to each other than to the second dimension, there were still fairly different.

Apart from the third dimension being accounted for as a subsidiary of the arousal dimension, it may also be the case that this is an artefact that emerges from the way in which participants dealt with knocking compared with drinking movements, or another combination. One way in which to assess whether the configuration is due to an artefact, is to look at the subject weirdness index. The weirdness index is a measure of how well the data from each confusion matrix used in the analysis, is accounted for by the Euclidean distance space. The range of possible values lies between 0 and 1, with an index of 1 representing a good fit between data and model. In cases where the weirdness index is high, the movement are interpreted as using only some of the dimensions.

Weirdness for Actor1, Knocking is small (.138), compared with the other types of movements (Actor1 Drinking .4160; Actor2 Knocking .3037; Actor2 Drinking .3127). The main difference in the recognition data between Actor1's knocks compared to other movements is that *happy* movements were well recognised since they were a comical knock. It is unlikely, however, that the third dimension was due only to the particular *happy* knocks by Actor1, since in the two-dimensional Actor2's knocking movements also had a low weirdness index.

The result from both the above investigations to what the third dimension may represent were inconclusive. Instead it may be a better suggestion to try and find a property in the stimulus that would account for this dimension. In Chapter 3 an analysis of the kinematics markers of movements will be described and correlated to the psychological spaces from the current chapter. Such an analysis will allow us to test whether $D1_{Canonical}$ and $D3_{Canonical}$ were related to each other and also whether they represent activation in the movements.

In conclusion it is important to stress that a two dimensional Multi-dimensional scaling solution accounts for most of the variance in the data. The two dimensions are also related to dimensions identified in similar research that used facial expressions and experienced emotion as stimuli. The critical differences here is that the importance of the dimensions are reversed so that valence is the second dimension and arousal is the first. A third dimension improves the fit of the model and allows it to account for more of the variance in the data. It has been hypothesised that this dimension is related to the arousal dimension since this is a finding from research on experienced emotion. There was a small trend for this to be so, but further research is necessary to confirm this finding.

Summary of Research and General Discussion

Human arm movements with emotion were captured and made into animations that were shown to observers with the task of judging the emotion that was being depicted. The results were analysed in two ways, firstly as the proportion of correct recognition and secondly according to the confusions in the data. From the proportion correct recognition, it was clear that although participants could recognise emotions at levels above chance, recognition was by no means perfect. This analysis suffered from complications since different actors depicted emotions differently and the emotions conveyed in different actions were also judged differently. This problem is not unique to the movement expressions used in the current research, but has also been noted in research with facial expressions (Tagiuri, 1969).

Despite the individual differences in communicating emotion, there was some consistent confusion in the data from observers suggesting that a common psychological space may underlie recognition of emotion from all four types of movements. In order to test this, the confusion data were transformed in to distances between emotions and input to the multi-dimensional scaling INDSCAL algorithm. The resultant psychological space had a low number of dimensions, with most of the variance in the data being accounted for by two dimensions. This psychological space resembled that found in similar research that has used facial expressions and experienced emotion as stimuli, but with the importance of dimensions reversed. The most important dimension was arousal and this seemed reasonable considering the conscious experience of participants who all indicated they had used speed to gauge affect. The second dimension weakly resembled valence, but may be better thought of as approach/avoidance. This is because highly negative emotions such as *angry* and *afraid* were at one extreme of this dimension while *sad*, *weak* and *tired* along with positive emotions such as *relaxed* and *happy* were at the other end of the second dimension.

A third dimension also emerged, but this accounted for little of the data and seemed to have some relationship to arousal. It may represent intensity since *angry* and *strong* movements were at one extreme of this dimension while *afraid*, *sad*, *tired* and *weak* were at the other extreme.

The naming of these dimensions has been based on findings from previous research, however, human movements are a very different type of stimulus from faces and experiences. A way in which this difference has asserted itself has been in the ordering of dimensions. In all previous research that we have been able to find, the first and most important dimension has always been valence, while the second dimension is generally arousal. In the case of human movements, arousal seems to be most important, while valence took a secondary role. This does seem to be reasonable since a powerful signal in the data was the speed at which movements occurred. During the debriefing session observers also indicated that they were sensitive to the speed of movements and used it in their strategy to categorise emotions.

Using human movements to convey emotions resulted in one major change in the psychological space of perceived emotion compared with other stimulus types. It may also, therefore, be the case that the other dimensions are different from those suggested before. For instance, it has already been hypothesised that this dimension could represent approach/avoidance since all of the emotions that may be regarded as negative were not concentrated at one extreme of the dimension, in opposition to positive emotions as would be expected if this was a dimension of valence.

That the psychological space for emotion from human movements is different than the psychological space for other stimuli deserves to be examined more closely. In particular we believe that the representation space for emotion resembles underlying dimensions in stimulus properties. Both the debriefing sessions with participants and the data has suggested that the varying speed of movements may be such an underlying stimulus dimension. For this reason the relationship between the judgement data from observers and the speed of movements have been described in more detail in Chapter 3. In Chapters 3 and 4 we will show the important role that speed plays in the perception of emotion from motion and in Chapter 5 we will suggest possible stimulus properties that may account for Dimension 2.

Conclusions

From Experiment 1 it can be concluded that recognition of emotion from human movements follows a pattern that has been found in previous research in which perception of facial emotion and a person's own emotions were the stimuli. In particular, a low dimensional psychological space accounts for nearly all the variance in the data, suggesting that there are only a few dimensions that humans use to categorise emotion. That these dimensions may be different from that those used other potential sources (such as faces) suggests that different information in different modalities all contribute to a perceptual event. For this reason it is interesting to find similarities in representations that could be used to bind the different stimulus parts into a single event.

In the next chapter, the relationship between speed and the dimensions from the psychological space will be tested.

Chapter 3 - Perception of Emotion from Motion II

In Chapter 2 it was shown that human observers have some basic level of competence in recognising emotion from highly impoverished stimuli such as point-light displays of arm movements. Multi-dimensional Scaling analysis of observer's judgement data showed that the data could be re-conceptualised according to the confusions that were made between emotions. From this a low-dimensional psychological space was constructed that accounted for the data. This psychological space was different from that found when faces or experience were use as stimuli, however, this may represent the different type of information available from movement. For instance, dynamic information such as the movement speed is available from arm movements, compared with the mostly configural and spatial information from faces. Speed represents a good indicator of the actor's state of arousal and is, therefore, a good candidate for a stimulus property that would account for the first dimension of the representation space.

In this chapter the relationship between movement speed and perception of emotion is examined more closely. The aims of the research was to quantify the kinematics properties of expressive movements to show that there were consistent variations among emotions and that there is a relationship between this variation and the behavioural findings from Experiment 1. In the current chapter, firstly the methods for measurement of kinematics properties for the movements will be described. Following this is a description of Experiment 2.1 in which the kinematics properties for movements used in Experiment 1 have been analysed to determine statistical differences between movements with different emotions. Finally in Experiment 2.2 the kinematics properties have been correlated with the psychological space constructed from judgements made by participants in Experiment 1.

Experiment 2 - Emotion and Speed

Introduction

Amaya et al (1996) described differences in the properties of movements they recorded, but did not expand beyond neutral, sad and angry movements. In this section a similar analysis was undertaken to analyse the differences in speed between movements. The number of emotions was also expanded to incorporate the whole corpus of movements that were captured by Amaya et al (1996).

Firstly the kinds of properties of speed that are encapsulated in the kinematics of a movement will be described. Following this we will demonstrate that the value of a given kinematics marker measured form the wrist varies across movements, dependent on which emotion an actor is depicting in their movement and largely independent of the action being performed or the actor performing the movements. All the kinematics properties in Table 3.1 were compared across emotion, however the discussion will be centred on velocity as a representative case.

Secondly the psychological spaces from Experiment 1 were related to physical properties of the movements. There were two predictions for the results. A) Categorisation would be closely related to physical properties in the movements. B) In cases where the speed of movements differed with greater magnitude, observers would distinguish movements with greater accuracy. In cases where differences in speed between movements were smaller, there would be more confusion between movements.

Measurement of Kinematics Properties

In this section the way in which kinematics properties of movements were measured from the wrist will be detailed. Two types of movement kinematics have been described: waveform kinematics and kinematics properties. Waveform kinematics refers to instantaneous calculations of movement kinematics and they have typically been used as a tool for movement processing and in the calculation of kinematics properties. The array of instantaneous tangential velocities described in Chapter 2, is an example of waveform kinematics and can be visualised as the velocity profile of a movement. Kinematics properties are averages, peaks, minima, sums, differences and indices that represent landmarks on the waveform or that are calculated from a waveform. Kinematics properties have been used to test differences between movements and to test possible relationships between observers' judgements and stimulus properties. The movement duration is an example of a kinematics property derived from the tangential velocity waveform and is encapsulated for each movement as a single value that denotes the time taken to complete the action.

Table 3.1 Waveform kinematics and the kinematics properties that are calculated from them.

Waveform Kinematics	Kinematics Properties
Tangential Velocity	Duration Maximum Velocity Average Velocity Distance Moved
1st Derivative of Tangential Velocity - Acceleration	Maximum Acceleration Maximum Deceleration
2st Derivative of Tangential Velocity - Jerk	Jerk Index

In the following paragraphs the calculations for movement speed and its derivatives have been described. Table 3.1 summarises the different kinematics properties that were measured from waveform kinematics of speed. It is structured to show the kinematics properties that were measured from different kinds of waveform kinematics – velocity, acceleration and jerk.

Kinematics Properties for Duration and Distance

Two basic movement properties are the **duration** of a movement and the **distance moved** or path length of the hand. The duration (T) of a movement can be defined either as the number of frames in a movement or as the time taken to complete a movement. In either case it is calculated using the start and end-points (P_{start} and P_{end}) defined in Chapter 2. In terms of frames, duration (T_f) is defined as the difference between P_{start} and P_{end}. In terms of time, duration (T_t) is simply the quotient of T_f and the sample frequency (*sf*).

$$T_f = P_{end} - P_{start} \quad \text{Equation 2}$$

$$T_t = \frac{T_f}{sf} \quad \text{Equation 3}$$

As a complement to the duration of a movement, the distance moved is also informative. In the case of the hand or wrist, it is the path length that is measured to give a crude idea about the spatial properties in arm movements. Path length (D_{PL}) is defined as the integral of V_{tan} (see Equation1, pp 38).

$$D_{PL} = \int_{T_{start}}^{T_{end}} (V_{tan})\,dt \quad \text{Equation 4}$$

Kinematics Properties for Velocity, Acceleration and Jerk

Kinematics properties for the velocity are **peak and average velocity.** These properties are calculated using the tangential velocity waveform (V_{tan}) and are defined as the maximum and mean V_{tan} respectively. To attain the acceleration the V_{tan} is differentiated once. From the acceleration waveform the largest value is read as the **maximum acceleration** and the minimum value is the **maximum deceleration.**

Apart from velocity and acceleration there is a third time derivative of position, the jerk, which can be used to provide a measure of the smoothness of a movement. To calculate instantaneous jerk the spatial displacement of x, y and z components of the wrist between successive frames is differentiated three times with the sample frequency. This results in the three components of jerk ($\frac{d^3x}{dt^3}$, $\frac{d^3y}{dt^3}$ and $\frac{d^3z}{dt^3}$) that can be used to attain the **jerk index** (JI) kinematics property. Adapted from the Cartesian jerk of Flash and Hogan (Flash & Hogan, 1985) the jerk index (Kagarer et al, 1998; Teulings et al, 1997) is the integrated squared jerk.

$$JI = \frac{1}{T_t} \int_{t_{start}}^{t_{end}} \left(\sqrt{\frac{d^3x}{dt^3} + \frac{d^3y}{dt^3} + \frac{d^3z}{dt^3}} \right) dt \quad \text{Equation 5}$$

Experiment 2.1 - Quantifying Speed Differences in Movements

Introduction

After the kinematics properties were measured, statistical data analysis techniques were used to test for significant differences between movements with different emotion.

Methods

Movement Collection

The movements were the same as those used by Amaya et al (1996) and in Experiment 1; however, movements with emotion other than neutral, sad and angry were included in the analysis.

Procedure

As described in the general methods, kinematics properties were measured for each of the 118 movements using Matlab and recorded in a text file. As a replacement for missing values for missing movements from Actor2, the mean was calculated from the 2 remaining exemplars for the weak movements and the relaxed drinking movements of Actor 2.

Results and Discussion

In order to analyse differences in emotions, each type of kinematics property was compared in a separate 2 x 2 x 10 ANOVA with mixed design by treating the 3 movement exemplars as subjects. The 2 actors were treated as a between groups variable, the 2 actions and 10 emotions were repeated measures variables. Average velocity presents a good example of the differences between movements and has been described more carefully in the result section that follows. Results from the ANOVAs for movement duration, peak velocity, distance moved, peak acceleration and Jerk Index have been presented in Appendix III, but in general the pattern of results obtained with average velocity, has been replicated in the other kinematics.

Average Velocity

There was a significant main effect of emotion [$F(9, 36) = 229.133$, $p<.001$] - the variable of most interest in the current research – as well as significant main effects for the movements of different actors [$F(1, 4) = 382.851$, $p<.001$] and for movements of different actions [$F(1, 4) = 1023.429$, $p<.001$]. The individual differences of actor and action were not inherently interesting, since an actor and action represent the primary motion patterns described by Amaya et al (1996), however, they did result in complex interactions with emotion. As seen in Chapter 2, the differences between actors and actions were also reflected in the behavioural data from Experiment 1.

The interactions between factors in the data will be discussed further to show that despite individual differences between actors and actions, there was still a largely consistent pattern to the differences in velocity between emotions. The interactions were: a **three-way interaction** between *actor, action and emotion* [$F(9, 36) = 5.427$, $p<.001$], two **way interactions** between *emotion and actor* [$F(9, 36) = 28.525$, $p<.001$] and *between emotion and action* [$F(9, 36) = 38.409$, $p<.001$] and finally a two way interaction between *actor and action* [$F(1, 4) = 152.053$, $p<.001$].

The **three-way interaction** between actor, emotion and action is illustrated in Figure 3.1 where the mean average velocity for each emotion is plotted according the actions and actors. Action, actor and emotion have been dealt with separately in the sections below.

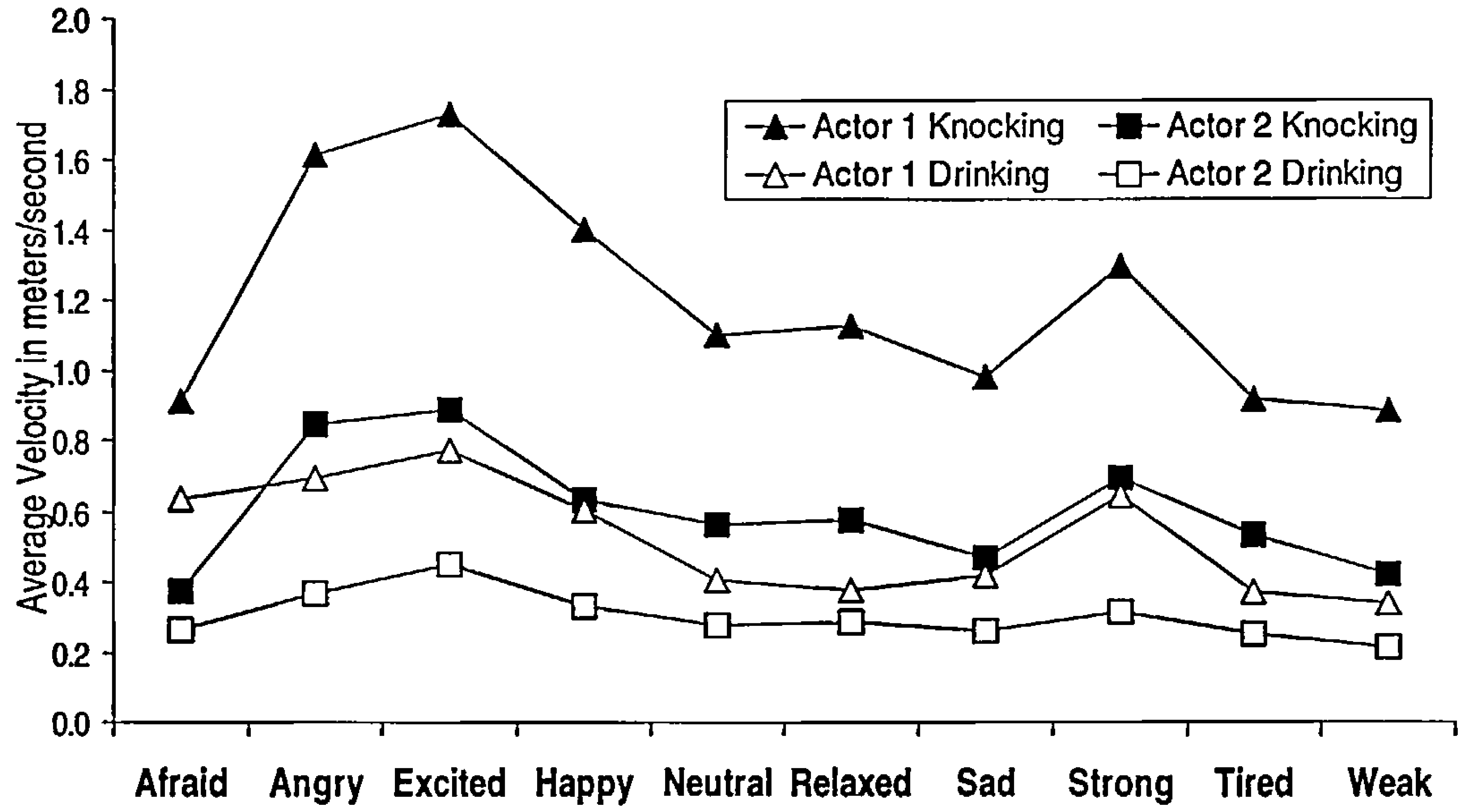

Figure 3.1 Average velocity of emotions for different actions and actors.

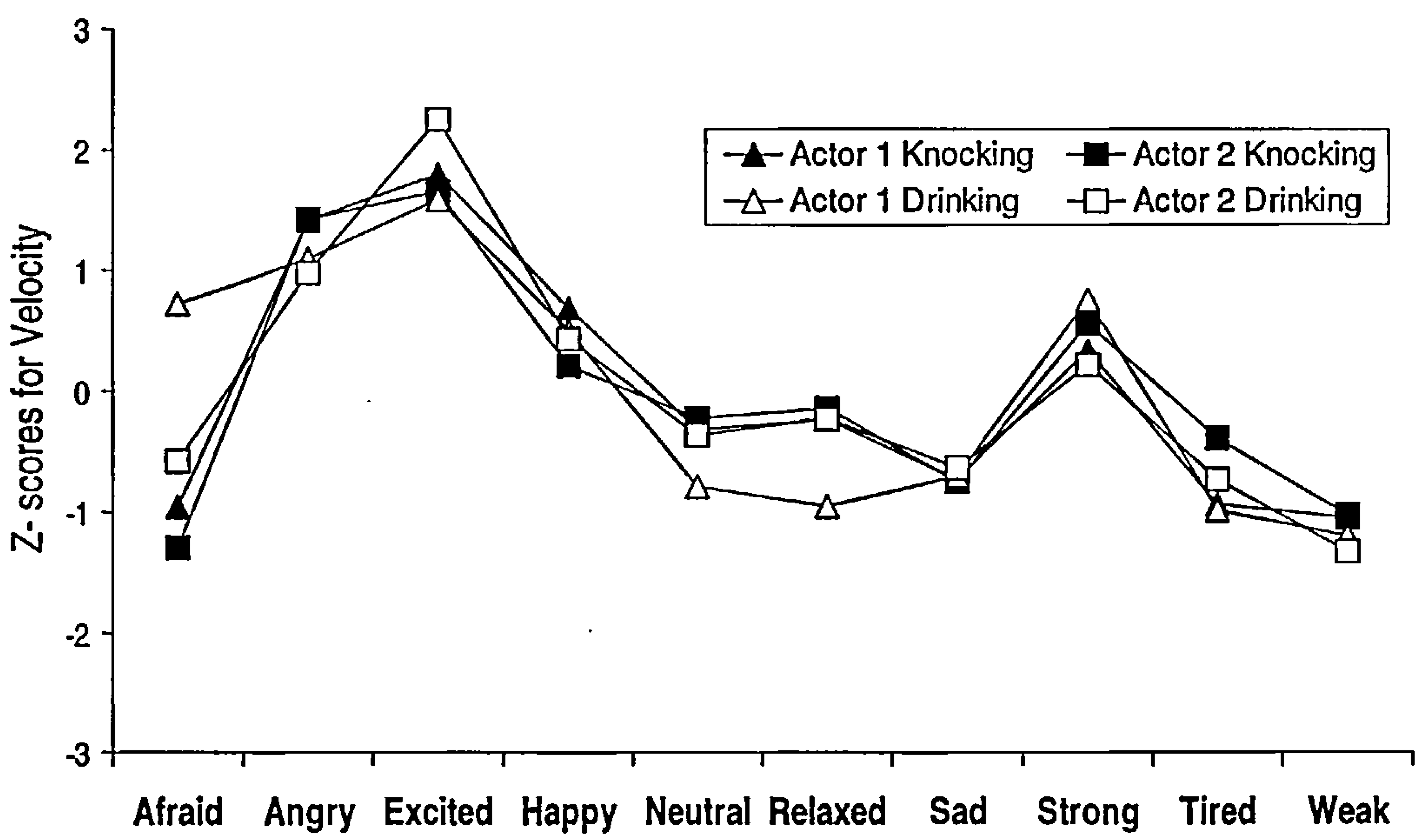

Figure 3.2 Differences between affects plotted as z-scores. Velocity was normalised for actor and action throug linear transformation that preserved the pattern of results but allowed comparison between movements.

Actions

Simple main effects were found at most levels of actor and emotion for action (in all significant cases $p<.05$) in the three way interaction of actor x action x emotion. The only exception was for the *afraid* movements of Actor2 where there was no significant difference between drinking and knocking actions. Knocking movements with *afraid* emotion were still a little faster (mean Average Velocity = .64 meters per second) than drinking movements with *afraid* (mean = .45 m/s). In all other cases knocking movements (mean = .9 m/s) were also faster than drinking movements (velocity = .41 m/s) and this has been reflected in the main effect of action as reported above and illustrated in Figure 3.3a. It was not possible to directly compare drinking movements of Actor1 with knocking movements of Actor2, however the average velocity for these two seem very similar.

In general the differences between drinking and knocking movements were larger for Actor1 than for Actor2 and this was reflected in a two-way interaction between the actor and action graphed in Figure 3.3b. This was probably because the movements of Actor2 were much less dynamic than those of Actor1, resulting in a smaller range of velocity for the different actions. Finally there was a two-way interaction between action and emotion and this has been graphed in Figure 3.3c. There were simple main effects of action at all levels of emotion at $p<.001$. The interaction most probably reflects the greater range of velocity for emotion in knocking movements than in drinking movements. The greater range of velocity resulted in more of the differences between emotions being significant for knocking movements than for drinking movements. The differences between emotions will be further discussed in a later section.

For the purposes of this research it is not problematic that there were differences in the kinematics of different actions. This is the case because a robust pattern in the differences of the secondary motion signal (emotions) despite different primary motion signals (actions and actors) would be evidence of a potential invariant in expressive movements. The difference between actions in this case was probably a result of the actors having to control a glass with liquid during the drinking movements. This caused drinking movements to be slowed down. On the other hand, knocking movements were faster since the object actors interacted with – a door – did not have to be handled as carefully.

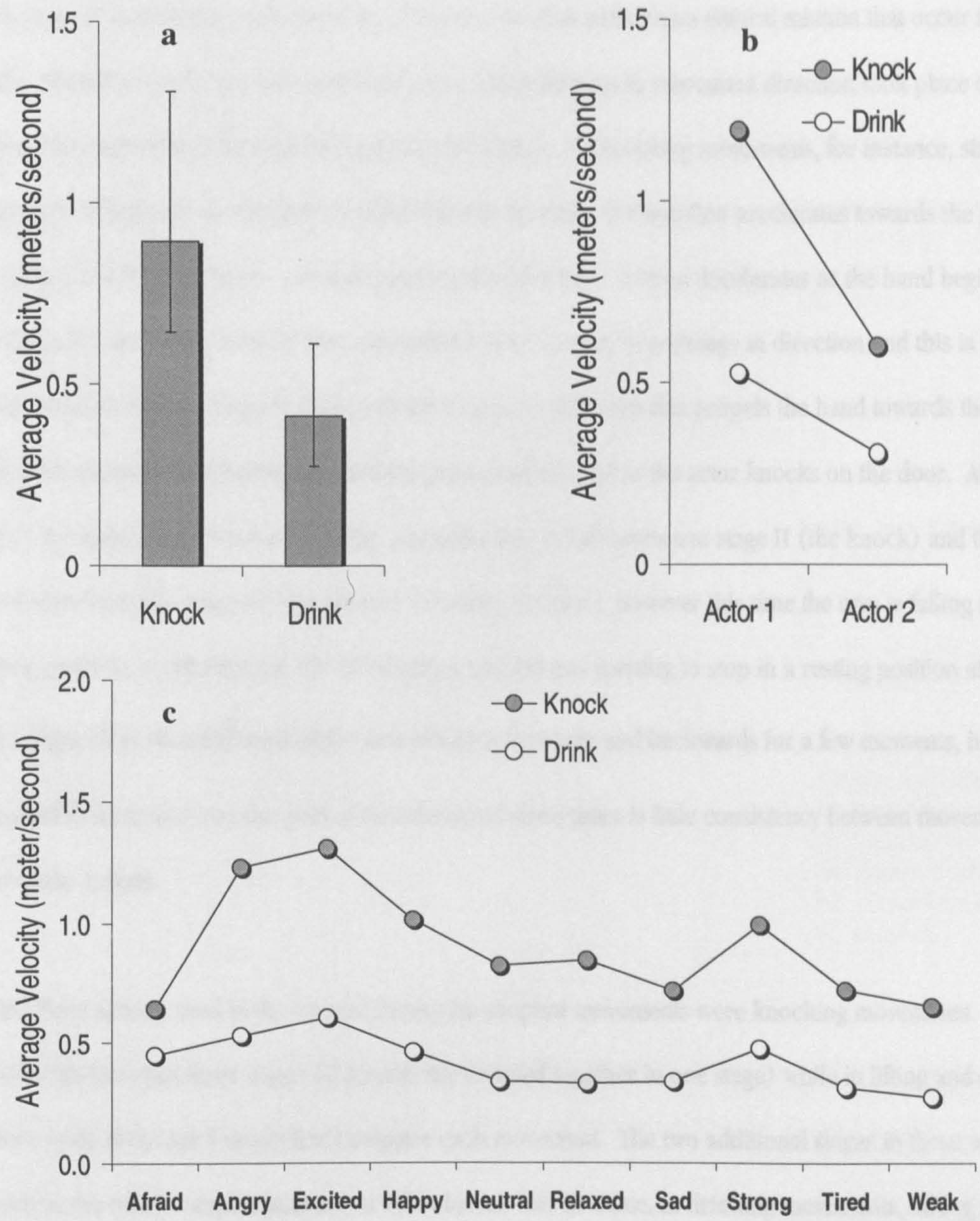

Figure 3.3 Significant main effect and interactions with action **a)** main effect of action, error bars depict the range; **b)** interaction between action and actor and **c)** interaction between action and affect.

Parts of Actions

Because of the different ways in which objects were handled in drinking and knocking actions, there was also diversity in the number and type of stages (Hill & Pollick, 2000) that comprised different actions as depicted in Figure 3.4. A description of the different movement stages is another way in which to account for possible consistent differences in kinematics between actions.

Each stage in a movement represents the portion of movement between natural minima that occur in the motion track. Such minima on our data occurred either when changes in movement direction took place or when the movement stopped and then started again for any reason. In knocking movements, for instance, stage I (the approach) represents the hand being lifted towards the door, the arm first accelerates towards the position at which the knock will happen – normally eye or shoulder level. It then decelerates as the hand begins to change direction for the knock, velocity becomes virtually zero as there is a change in direction and this is when stage I of the movement ends. Stage II starts with the change in direction that propels the hand towards the door. Stage II is characterised by rapidly changing peaks and troughs as the actor knocks on the door. After the final knock the hand changes direction again, signalling the end of movement stage II (the knock) and the start of movement stage III. Stage III (the retreat) is similar to stage I, however this time the arm is falling back to the resting position, accelerating in the initial stages and the decelerating to stop in a resting position at the actor's side. Stage III is often followed by the arm swinging forwards and backwards for a few moments, however, P_{end} is normally set to discount this part of the movement since there is little consistency between movements for this part of the actions.

Of the three actions used in the current thesis, the simplest movements were knocking movements. Knocking movements have just three stages (if knocks are counted together in one stage) while in lifting and drinking actions both, there are 5 stages that comprise each movement. The two additional stages in these movements represent the more complex interaction with objects. For instance, in drinking movements, taking a sip represents a movement stage in itself. Lifting an object also requires more careful motor control to prevent clumsiness, resulting in a more complex motion signal.

One of the consequences of such delicate interaction with objects was that the drinking and lifting movements were less dynamic than knocking movements. Additionally since the maximum velocity of knocking movements was less constrained, it was possible for actors to incorporate a greater range of speed in their movements. This greater range allowed observers of the movements in Experiment 1 to better distinguish and categorise movements.

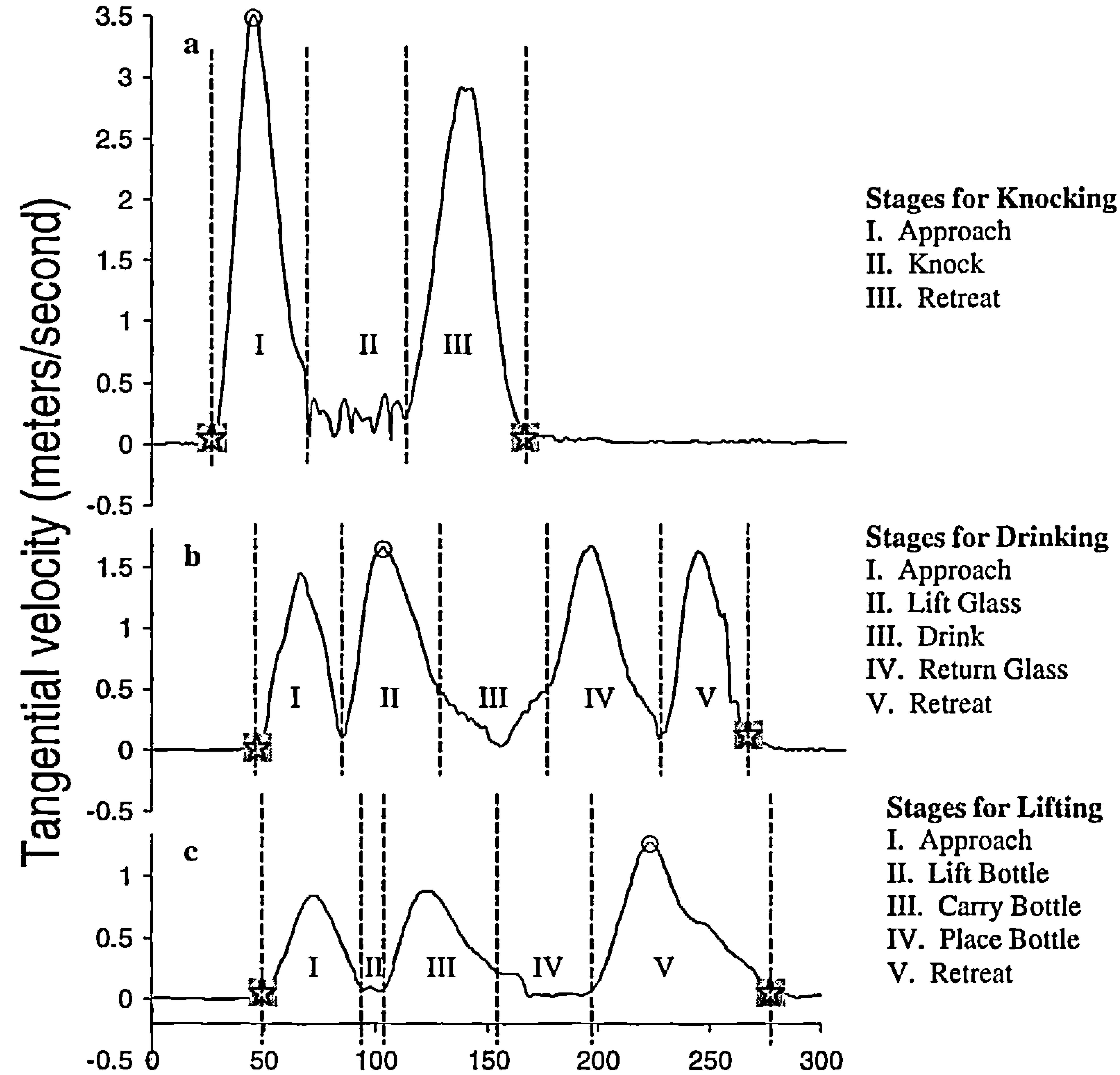

Figure 3.4 Movement stages for a) knocking b) drinking and c) lifting movements. Stars represent start (on the left) and end-points (on the right) for each movement and circles indicate locations of maximum velocity. Frame numbers are shown on the x-axis and dashed lines define the borders in between movement stages. Each stage is labelled with a Roman Numeral.

Actor

In the same manner as there was variety in the kinematics for different actions, there was variety in the movement signals of different actors. Although they did not analyse movement kinematics, Cutting and Kozlowski (1977) found that in a recognition task observers could recognise different actors from biological motion displays. This indicated that in the distal stimuli that depicts different actors, there was enough variation for perceptual categorisation of identity. It is not entirely clear how these differences are encoded, however, the movement characteristics that depict different actors is regarded by us as part of the primary motion signal for the purposes of analysing differences between emotions.

For the three-way interaction between actor, action and emotion, simple main effects were found for actor at every level of emotion for knocking movements (in all cases $p<.005$) and in each case Actor1 moved faster than Actor 2. For drinking movements the differences between Actor1 and Actor2 were not as pronounced, but nevertheless there were simple main effects for half the movements. *Angry*, *afraid*, *excited* and *strong* all showed simple main effects significant at $p<.005$; *happy* actions were significantly different for Actors 1 and 2 with a simple main effect significant at $p<.05$. *Neutral*, *relaxed*, *sad*, *tired* and *weak* movements showed no simple main effects for the difference between Actors 1 and 2. The trend for *neutral*, *relaxed*, *sad*, *tired* and *strong* was for the movements of Actor1 to be faster than those of Actor2. This finding was reflected in the main effect of actor.

There was a two-way interaction between actor and emotion, however, as with actions, this interaction most probably reflected the greater range of speed for Actor1 (mean = .86 m/s) compared with Actor2 (mean = .45 m/s) and the fact that Actor1's movements were always faster than those of Actor2.

Summary of Effects for Actions and Actors

Knocking movements were always faster than drinking movements and had a greater range of speeds between different emotions. Similarly, the movements of Actor1 were faster than those of Actro2 and had a greater range of values between emotions. There was an interaction between action and actor since the movements of Actor2 were slower than the movements of Actor1. This resulted in a smaller difference between knocking and drinking actions of Actor2 than for the movements of Actor1.

Emotion

The differences between actions and between actors seem to have been consistent. However it is still an open question as to whether there was similar consistency in the pattern of velocities for different emotions or whether the pattern of differences between emotions would also change depending on the type of movement. In the following discussion this issue has been more carefully examined.

For the three-way interaction there were significant simple main effects of emotion at each level of actor and action and have been summarised in Table 3.2 below. For each pair of emotions the Honestly Significant Difference (HSD) was calculated using a Tukey test and these results are tabulated in Table 3.3 (pp. 74). From this post hoc analysis, there seems to be a fair number of consistencies in the data. For instance angry and exited movements were in general consistently faster than other movements and excited movements were faster than angry ones. Happy and strong movements were the next fasted movements, however, there was little agreement about which of the two movement were faster. Neutral and relaxed movements were never significantly different from each other and the speed of tired, sad and weak movements were often similar.

Table 3.2 Simple Main Effects for emotion at each level of actor and action.

Actor	Action	df	F	p
Actor1	Knock	9,39	191.288	<.001
	Drink	9,39	52.281	<.001
Actor2	Knock	9,39	59.644	<.001
	Drink	9,39	9.573	<.001

Examining the results in this way was, however, is not very informative for discovering whether there was a consistent pattern across different movement types in the way expressions were related to each other. This was mostly due to the expressions of different actions and different actors not being comparable because of the individual differences discussed above. These individual differences resulted in a greater range of velocities for a dynamic actor or action and in a smaller range for less dynamic movements.

The direct result of these individual differences can be seen in Table 3.3 where there were significant differences between expression in almost all cases for knocking movements of Actor1 but in only half the cases for drinking movements of Actor2.

Table 3.3 Results of post hoc Tukey tests of differences between affects at each level of actor and action for movement data. Results are tabulated as levels of significance according to action and actor.

		Actor 1										Actor 2									
		Afraid	Angry	Excited	Happy	Neutral	Relaxed	Sad	Strong	Tired	Weak	Afraid	Angry	Excited	Happy	Neutral	Relaxed	Sad	Strong	Tired	Weak
Knocking Movemens	Afraid																				
	Angry	***										***									
	Excited	***	***									***	ns								
	Happy	***	***	***								***	***	***							
	Neutral	***	***	***	***							***	***	***	**						
	Relaxed	***	***	***	***	ns						***	***	***	*	ns					
	Sad	*	***	***	***	***	***					***	***	***	***	***	***				
	Strong	***	***	***	***	***	***	***				***	***	***	*	***	***	***			
	Tired	ns	***	***	***	***	***	*	***			***	***	***	***	ns	ns	*	***		
	Weak	ns	***	***	***	***	***	***	***	ns		ns	***	***	***	***	***	ns	***	***	
Drinking Movements	Afraid																				
	Angry	*										***									
	Excited	***	**									***	***								
	Happy	ns	***	***								*	ns	***							
	Neutral	***	***	***	***							ns	***	***	ns						
	Relaxed	***	***	***	***	ns						ns	**	***	ns	ns					
	Sad	***	***	***	***	ns	ns					ns	***	***	*	ns	ns				
	Strong	ns	ns	***	ns	***	***	***				ns	ns	***	ns	ns	ns	ns			
	Tired	***	***	***	***	ns	ns	ns	***			ns	***	***	**	ns	ns	ns	*		
	Weak	***	***	***	***	*	ns	**	***	ns		ns	***	***	***	*	*	ns	***	ns	

Levels of α: *** α=.001, ** α=.005, * α=.05
ns = non-significant

Z-scores of Kinematics

In order to make movements comparable with each other, the velocity data were transformed into standard z-scores. This is a linear transformation that normalises the data according to the mean and standard deviation of a whole population. The velocity of each movement was transformed by using the mean and standard deviation of all the actor's movements for a particular action. The resultant scores are easily comparable across actor and action and are plotted in Figure 3.2 (pp. 68).

This method of visualising the data shows that despite differences resulting from consistent variation between actions and actors, a clear pattern of changes between emotions emerged. Movements seem to fall into four groups. The fastest were *angry* and *excited* movements followed by *happy* and *strong* movements. *Neutral* and *relaxed* formed the third grouping followed by *afraid*, *sad*, *tired* and *weak*. The only anomalous movements are *afraid*, *relaxed* and *neutral* when Actor1 was making a drinking action and the *tired* movement for knocking actions by Actor 2.

We tested these observations in an ANOVA of the same design used for regular kinematics (described on pp. 66). Results showed a main effect of emotion [$F(9, 36) = 267.397$, $p<.001$]but not main effect of actor [$F(1, 4) = 0$, ns] or action [$F(1, 4) = 0$, ns]. There were also two way interactions for actor by emotion [$F(9, 36) = 10.073$, $p<.001$] and action by emotion [$F(9, 36) = 11.981$, $p<.001$], but not for the two-way interaction between actor and action [$F(1, 4) = 0$, ns]. Finally, there was a significant three-way interaction between actor, action and emotion [$F(9, 36) = 4.553$, $p<.001$].

A closer examination of the interactions indicated that there were only a few cases in which the pattern of speeds were different for actors or actions. The three-way interaction can be attributed to the *afraid* drinking movement of Actor1 being faster than all the other movements with *afraid* emotion. This was also reflected in the two-way interactions. For the interaction between emotion and action, there were simple main effects for *afraid*, *angry* and *relaxed* with knocking movements having higher values than drinking movements in the last two cases. For the interaction between emotion and actor, there were simple main effects for *afraid*, *relaxed* and *tired* with the movements of actor 1 having lower values in the last two cases.

In general, however, the patterns of z-scores for the 4 different movement types are remarkably similar. This reflects the fact that kinematics for different emotions vary consistently, even when the primary motion signals change. The question of whether this pattern is available and of use to observers in a categorisation task will be addressed in Experiment 2.2.

Summary and Conclusions

In the previous experiment the movements of two actors were recorded as they performed knocking and drinking movements with emotion. These movements were processed and their kinematics properties for speed were measured. A comparison was made between the kinematics properties for different emotions to show that despite differences in actor and action, there was a stable pattern of kinematics for emotion. Average velocity was converted to z-scores – a linear transformation that retained the shape of the data while making all movements comparable – and the results showed smooth and consistent differences between emotions in which angry and excited movements were the fastest. Sad, weak and tired movements were the slowest and two additional groups of movements were happy and strong and neutral and relaxed. Although finding an invariant for emotional expressions has not formed a part of the current thesis, the analysis of movement kinematics functions well as a pilot study to show that such an invariant may be revealed with future research.

Being able to show consistent differences between emotions has been valuable, however, there still remains the question of whether the human visual and cognitive system uses these differences in a meaningful manner that may inform categorisation of emotion from movement. The rest of this chapter is devoted to establishing whether the categorisation of human movements is comparable to the kinematics differences in movements. The following comparison between movement kinematics and behavioural data from Experiment 1 is the first in a series of experiments that examine the relationship between perception and physical properties in a movement.

Experiment 2.2 - Comparison of Perception to Movement Kinematics

The movement kinematics were examined to see whether any physical properties of the movement were related to either of the two dimensions defining the psychological space from Experiment 1.

One of the striking features in the movement data was that the kinematics properties we measured consistently and smoothly differed between emotions. For instance, *sad* movements were always slower than *neutral* movements and both these were slower than *angry* movements. This seems to correspond to the activation dimension from the psychological space where *angry* and *sad* movements were at opposite ends of the dimension with *neutral* somewhere in the middle. Hence there seems to be a relationship between the psychological data and the kinematics of speed in movements.

To test whether movement speed informed perception of emotion, the kinematics properties were correlated to Dimension 1 and Dimension 2 and Dimension 3 co-ordinates of the 10 emotions in the psychological spaces from Experiment 1.

Results and Discussion

Two-dimensional Psychological Space

Results of the correlation between the 2-dimensional psychological space from Experiment 1 (Figure 2.5, pp. 54) and kinematics properties from Experiment 2.1 are presented in Table 3.4, and Figure 3.5 shows an example of this relationship.

Table 3.4 Correlation between kinematics properties and psychological space. All values are r^2.

	Duration	Peak Velocity	Average Velocity	Distance Moved	Peak Acceleration	Peak Acceleration	Jerk Index
Dimension 1	0.710	0.827	0.839	0.824	0.691	0.626	0.694
Dimension 2	0.417	0.278	0.235	0.002	0.458	0.323	0.351

The $D1_{Canonical}$ (activation) co-ordinates of emotions correlated with the kinematics properties in such a way that energetic movements such as *angry* and *excited* were positively correlated with shorter duration and greater magnitudes of average velocity, peak velocity, acceleration, deceleration and jerk. For $D2_{Canonical}$ we found that, to a lesser extent there was a tendency of longer duration and smaller magnitude of the other kinematics properties to be correlated with positive emotion. The r^2 values presented in Table 3.4 indicate that the best kinematics marker for a correlation with $D1_{Canonical}$ was average velocity, closely followed by peak velocity. Overall, however, this dimension was correlated well with all the kinematics properties. $D2_{Canonical}$ on the other hand, was only weakly correlated with kinematics properties.

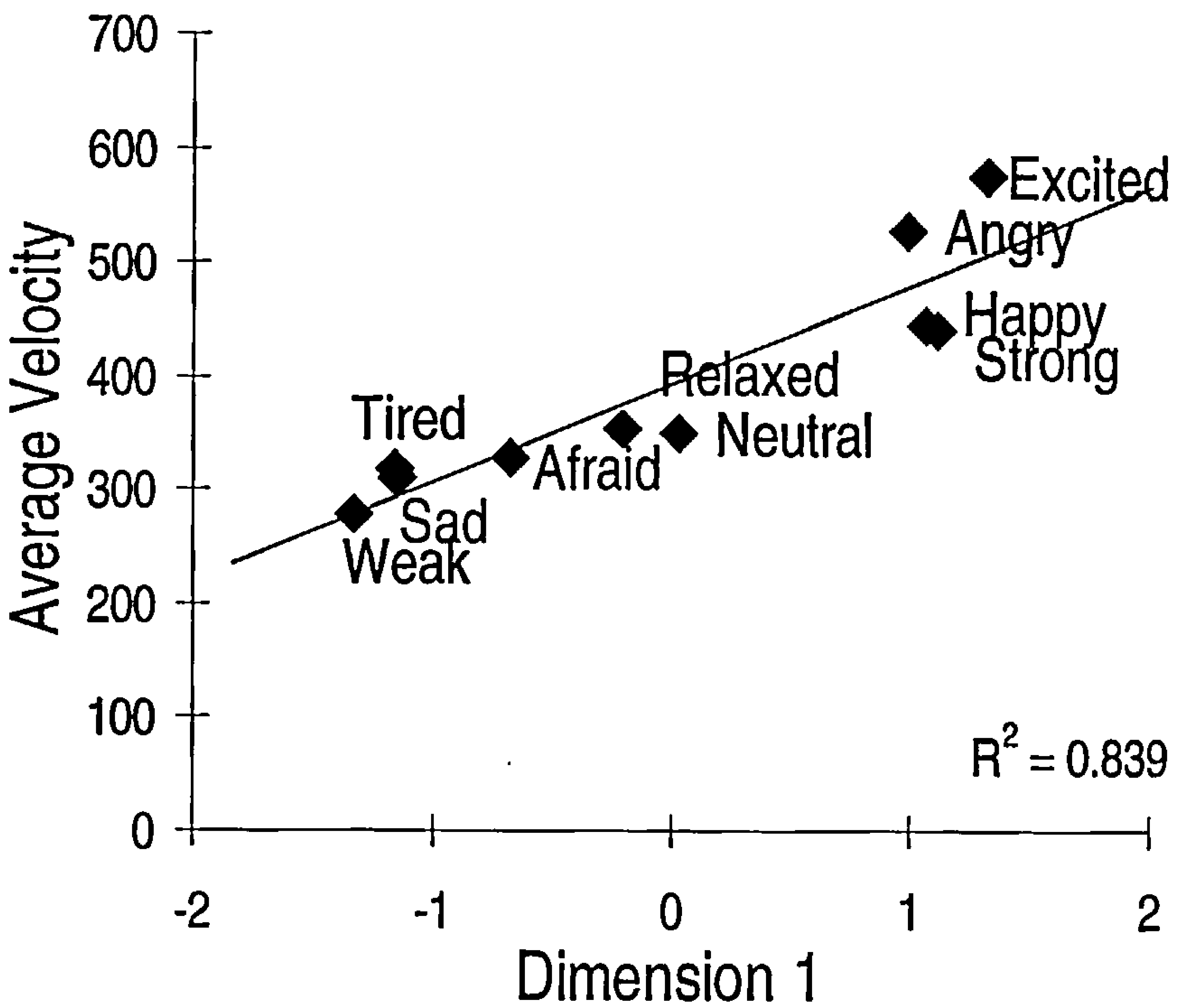

Figure 3.5 Correlation between Dimension 1 of the 2-dimensional psychological pace and movement velocity.

We examined the relationship between speed and the psychological space further by rotating the psychological space to find the orientation that maximised the r^2 values of the correlation with the six kinematics properties. It was found that a 27° counter-clockwise rotation resulted in the highest correlation with the kinematics properties with an average r^2 value of 0.88 for $D1_{Canonical}$ and 0.03 for Dimension $2_{Canonical}$. This rotated psychological space has been reproduced in Figure 3.6. From these results it is clear that while the original psychological space (Figure 2.5, pp.54) was roughly oriented so that energy in Dimension 1 is correlated with the speed of the movement, rotation of the space can improve the correlation. Rotating the psychological space also had the added effect of making the distinction between positive and negative emotions clearer so that *sad* emotion are now on the same plane as *angry* emotion.

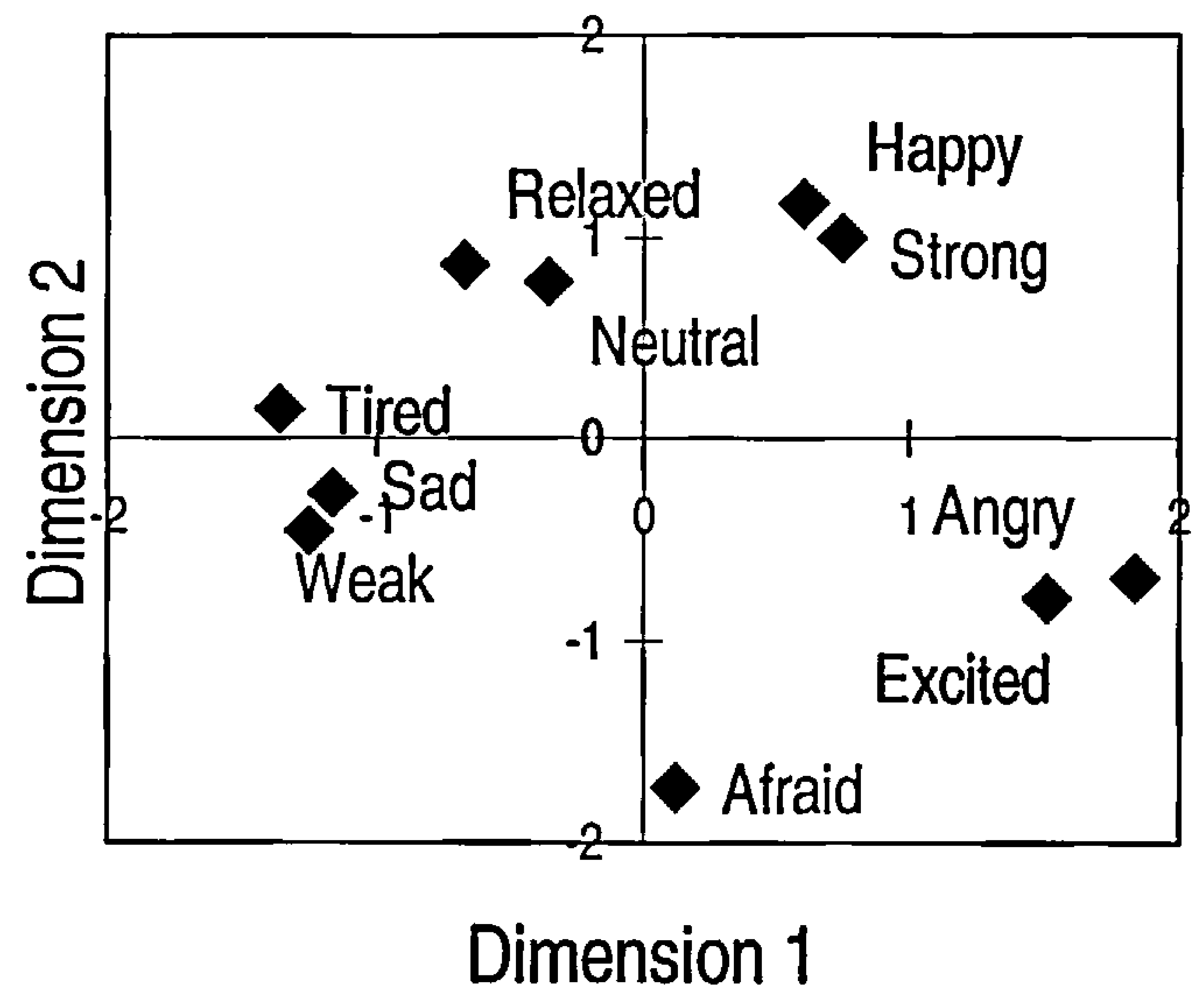

Figure 3.6 Psychological Space rotated for maximum correlation between Dimension 1 and kinematics properties.

Three-dimensional Psychological Space

In constructing the psychological space in Chapter 2, a third dimension emerged that accounted for about 11% of the data. It was hypothesised that in line with previous research, this dimension would be related to $D1_{Canonical}$. If this were the case, then $D3_{Canonical}$ would also be correlated to kinematics properties. To test this, the dimensional co-ordinates of the three-dimensional psychological space were correlated with the kinematics properties. The r^2 values for this correlation is shown in Table 3.5.

Table 3.5 Correlation between kinematics properties and 3-dimensional psychological space. All values are r^2 and best correlations for D1 andD3 have been highlighted.

	Duration	*Peak Velocity*	*Average Velocity*	*Distance Moved*	*Peak Acceleration*	*Max Deceleration*	*Jerk Index*
D1	0.708	0.813	0.790	0.822	0.658	0.522	0.614
D2	0.305	0.174	0.139	0.011	0.318	0.192	0.229
D3	0.358	0.424	0.504	0.253	0.471	0.690	0.620

From Table 3.5 it is clear that $D3_{Canonical}$ correlated with kinematics properties to some degree. In particular there is a good correlation of $D3_{Canonical}$ with maximum deceleration and with jerk index (these correlations have been shaded in Table 3.5 along with the best correlations for $D1_{Canonical}$). Maximum deceleration and jerk index were both calculated from higher order derivatives of velocity that give some measure of dynamics aspects of a movement. For instance, the jerk index represents the smoothness of a movement while the peak deceleration occurred in the movements at points when an opposing force was encountered by the hand, such as when the hand made contact with the door or when a glass made contact with the table surface.

Most importantly, these results show a clear relationship between $D1_{Canonical}$ and $D3_{Canonical}$ in that both are correlated with speed kinematics while in comparison $D2_{Canonical}$ showed no relationship with kinematics. It seems clear that both $D1_{Canonical}$ and $D3_{Canonical}$ represent a dimension of arousal that is perceptually encapsulated in the speed at which movements take place and in the higher order derivatives of speed.

Proportion Correct

A final comparison can be made between the proportion correct recognition and the kinematics properties. From the analysis of movements it became clear that drinking actions were slower than knocking actions and that the movements of Actor2 were also slower than the movements of Actor1. In the same way as there were differences in the movements, observers also found some movement types harder to categorise than others. For instance, the knocking movements of Actor1 were the fastest and were also best recognised by observers. In kinematics properties there was not much difference between Actor1 drinking and Actor2 knocking, a finding that was also reflected in the recognition data, finally the drinking movements of Actor2 were the slowest and were also the worst recognised.

Runeson (1994) suggested a reason for gender discriminations to be better from faster movements than from slower ones. He observed that more energetic movements made gender more discernible since greater acceleration also leads to an increase in reactive forces. These reactive forces reveal a person's anatomical makeup allowing better discrimination. It is possible that vigorous movements would also reveal more about the range of transient person characteristics - such as emotion - since they allow a greater range of kinematics to be exploited that in turn would allow for finer discriminations of movements by an observer.

It is not possible to directly compare velocity and the perceptual judgements since they are comprised of different units. However, by transforming the data into standard z-scores, this comparison becomes reasonable. Z-scores were calculated for average velocity and for the proportion correct recognition of emotions and graphed as shown in Figure 3.7. From Figure 3.7 it is clear that there is a close correspondence between movement speed and competence of observers to correctly categorise emotion. This data supports comments by Runeson and Frykholm (Runeson, 1994; Runeson & Frykholm, 1983) that more dynamic movements enhances perception. From the analysis of movement kinematics (Experiment 2.1), it was clear that a wider range of kinematics are apparent in highly dynamic displays. Since kinematics offer a major cue to categorisation, the wider range of kinematics values that is perceptually available enhanced discrimination.

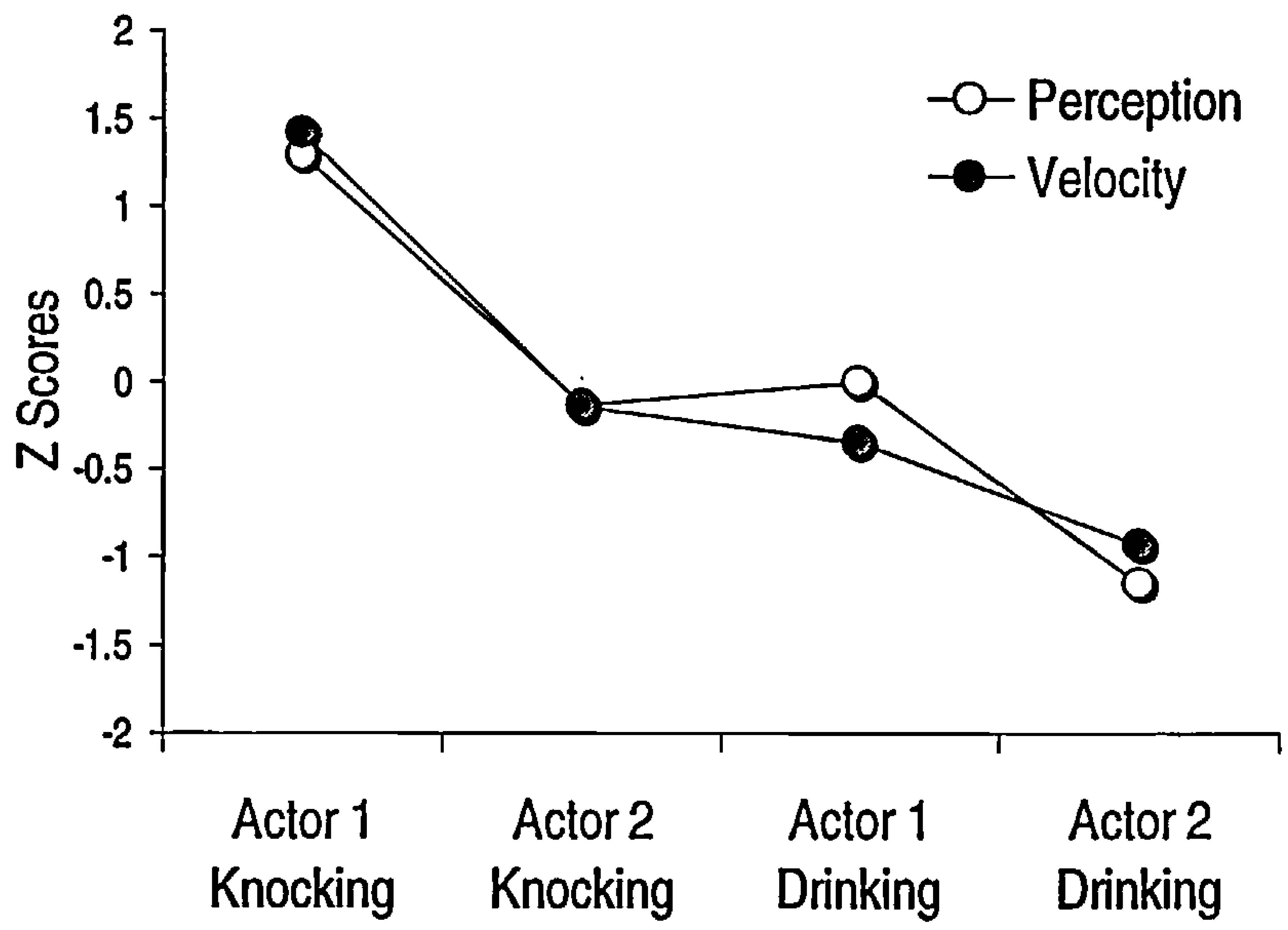

Figure 3.7 Z-scores for average velocity and proportion correct recognition

Discussion

From the analysis of movements described in Experiment 2.1 it was clear that different emotions resulted in movements with different kinematics. We then showed in Experiment 2.2 that there was a relationship between the kinematics properties of movements and perceptual judgements made by observers. This relationship was shown with the psychological space from Experiment 1, indicating that as proposed earlier, the first dimension along which movements are categorised was speed.

The finding that kinematics are correlated with judgements, is reminiscent of the Kinematics Specification of Dynamics (the KSD-principle) of Runeson and Frykholm (Runeson, 1994; Runeson & Frykholm, 1983). With the KSD-principle they suggest that information about the dynamics of an event may be uniquely specified in the kinematics available to perception. For human movements such dynamic information would also include any causal properties that would play a role in generating movement (Runeson, 1994). Causal properties would include both the primary (actions, gender, identity) and secondary (intention to deceive, mood, emotion) motion signals. In a range of experiments it has already been shown that causal properties are perceived and used by observers even when they are hidden - such as the weight of a box being lifted (Runeson, 1994). In our case emotion was the hidden property that could be recognised from point-light displays.

What does speed show?

The question still remains; what relevance would kinematics properties have for recognition of emotion? As discussed with reference to Runeson and Frykholm (Runeson, 1994; Runeson & Frykholm, 1983) dynamics of movements are specified through kinematics. For perception of emotion it is probably the case that speed provides an indication of physiological arousal. Physiological arousal in turn indicates whether a person is active or inactive, aroused or sleepy or intensity of a feeling. These last three are the various forms of the activation dimension in the representation of emotion. It is therefore our belief that there is a direct mapping between the perceived velocity of a person's movements and the perception of the person's emotion.

Summary and Conclusions

In Experiment 2.1 we showed that there was correspondence between different movement types in the way that kinematics properties of emotions were relate to each other. This suggested a possible invariant of emotion in motion that observer may be sensitive to when observing movements with emotion. In Experiment 2.2 we tested the hypothesis that observers of the movements utilised the differences between kinematics for emotions in a strategy for categorising the emotion from movements. In order to do this, we correlated the kinematics with the psychological space that was constructed from responded in Experiment 1. The result of this correlation suggested that human observers had been sensitive to the pattern of kinematics in emotions and that they used such information in a strategy to classify emotions. In conclusion the major finding from Experiments 1 and 2 is that speed is a stimulus property observable from point-light displays of emotion and that this information plays an important role the categorisation of dynamic emotions.

From Experiments 1 and 2 it has not been possible to prove a causal relationship between movement speed and observers' judgements. This is a topic that will be dealt with in the next section when we try to modulate perception through speed.

Chapter 4– Speed Modulated Movement Categorisation

In Chapter 3 the digital capture of human movements was explicated as enabling the precise measurement of physical properties in human movements. This process made it possible in Experiment 2 to compare perceptual judgements of human observers with kinematics properties of the movements and allowed us to make inferences about the role movement speed played in the categorisation of emotion. It is also possible, however, to turn this process around and manipulate a movement's signal by enhancing or suppressing specific qualities. Alternatively it may be desirous to add aspects from another movement either to the primary or secondary motion signal to create a larger repertoire of movements (Amaya, Bruderlin and Calvert, 1996).

The aim of the following experiment was to further investigate the role of speed in the recognition of emotion from human movements by showing a causal relationship between speed and categorisation. New *sad*, *neutral* and *angry* lifting and knocking movements were recorded from 3 women and 3 men. Through time warping (Bruderlin & Williams, 1995) the duration of movements were manipulated to make time morphs. The original and morphed movements were displayed as point-light stimuli to 10 participants who categorised them and judged the intensity of emotion. Results showed that manipulating speed in movements could modulate perceived intensity of emotion. Categorisation not mediated by velocity was still observed for *angry* movements, but not for *sad* and *neutral* movements.

Experiment 3 - Speed Modulated Movement Categorisation

Introduction

Image processing techniques have allowed cognitive and vision scientists to test specific hypotheses about face, object and scene categorisation (Morrison & Schyns, 2001) and about facial expression (Calder, Rowland, Young, Nimmo-Smith, Keane, & Perrett, 2000b; Calder, Young, Rowland, & Perrett, 1997; Young et al., 1997). More recently similar techniques have been applied to dynamic stimuli (Frank E. Pollick et al., submitted; Giese, 2002; Troje, 2002; Hill & Johnston, 2001; Hill & Pollick, 2000).

The underlying principles for manipulating motion signals have been adopted from image and signal-processing techniques with which still images are normally processed and manipulated (Bruderlin and Williams, 1995). For instance, Hill & Pollick (2000) applied signal-processing techniques to human movements to exaggerate temporal properties of actors' movements so that identity might be enhanced. In the current experiment movement signal-processing techniques have been applied to arm movements with emotion.

Preliminary research (Walker, 2000) has shown that *sad* and *angry* movements were distinguishable from each other and in the psychological space from Experiment 1 we showed that they fall at opposite poles on the first dimension ($D1_{Canonical}$). $D1_{Canonical}$ has been associated with perceived arousal, and in particular the speed of movements (Experiment 2). It is possible that, for *angry* and *sad* movements, the diagnostic features for recognition and categorisation lies in the speed of *angry* and *sad* movements. It therefore stands to reason that if the speed of *sad* and *angry* movements were modulated in some way the perception of emotion would also be modulated.

Subsequently the main aims of Experiment 3 were to further examine the effect of speed on the categorisation of *sad* and *angry* movements and to modulate categorisation through manipulating movement speed. To achieve this, temporal characteristics of the movements were exploited to ***morph between movements of different emotion*** and to ***exaggerate temporal properties of emotions.***

Although the aim of this research was to manipulate movement speed in a controlled and precise way, a direct manipulation of speed was neither required nor technically straightforward. The technical difficulties arose since the combinations of distances and time that would result in a specified velocity are infinite in number. In principle, it is possible to solve this problem, but in practice one of the components - velocity, distance or time - would always only be an approximation of the desired value[13]. We were primarily interested in the role temporal features played in the perception and categorisation of movements rather than the spatial structure. For these reasons it was sensible to use movement duration to mediate the temporal manipulations since

[13] The desired value for morphing is the exact midpoint between two values since this is the scaling factor to be used in the morphing procedure.

duration also correlated closely with dimension 1 in Experiment 2. Manipulating movement duration would also ensure that only temporal properties of the movements were changed while preserving spatial properties.

Temporal manipulations were accomplished through time warping (Bruderlin & Williams, 1995) with which temporal morphs of movements were made by using an interpolation algorithm. Time warping for motion signals is a procedure that has been adapted from dynamic time warping in signal processing for speech recognition. In animation, time warping is used to change a motion signal from its original temporal configuration to a new one. The new temporal configuration is typically that of a second motion signal. For instance, in the example by Bruderlin and Williams (1995), a drunken stagger can be transformed so that it takes place at the pace of a military march. In this way the temporal characteristics of a movement has been changed without affecting spatial characteristics such as form[14].

In the Experiment 3 the original temporal configuration of a movement was dictated by the emotion being depicted by an actor. The new temporal configuration was obtained from a movement by the same actor, in which a different emotion was portrayed. Using this new temporal configuration, the original movement (movement 1) was interpolated so that it had the new temporal configuration. In this case movement duration was used to dictate the temporal configurations so that after interpolation, movement 1 had the same duration as the second movement (movement 2). The chief benefit of using time warping was that the temporal configurations were not limited to those of the two movements. For instance the duration for interpolation of movement 1 could be either the duration of movement 2 or could fall anywhere in the range of durations between movement 1 and movement 2. Finally the duration of either movement could be exaggerated to formulate a new temporal configuration. This smoothly changing temporal configuration of movements shares similar characteristics to the spatial morphing that has been used for research in face perception and so the time warped movements have been named temporal morphs.

Original and temporal morphs were made into animations and shown to observers with the task of categorising them as *sad*, *angry* and *neutral* emotion. We predicted that in the case where temporal movement properties carried all the diagnostic information for categorisation, regardless of the original emotion depicted in a

[14] For other examples, see Giese (Giese, 2002) and Troje (Troje, 2002).

movement, movement speed (as mediated through time warped duration) would modulate categorisation of temporal morphs. If it were the case that diagnostic features of the movements were not only in the temporal aspects of movements, perfect modulation of categorisation would not occur due to a manipulation of duration alone.

Movement Collection

From Experiment 1 differences in velocity were clear between *angry* and *sad* movements, with *neutral* movements falling roughly in the middle between the two. A corpus of knocking and lifting movements was collected for 28 actors[15], using the procedures described in Chapter 2. Lifting movements replaced drinking movements, since they allowed more scope for kinematics differences because actors' movements would not be slowed by the control of a liquid-filled cup. Actors were instructed to make 10 *sad*, *angry* and *neutral* movements for each action and the motion paths of 6 IRED's – fixed to their head and arm - were recorded at a sampling frequency of 60 Hz.

Movement Selection

Three female and three male actors were pre-selected from the corpus based on preliminary data collected for an honours student project (Walker, 2000). In this project the *sad* and *angry* movements of 16 randomly selected actors were time-warped and displayed to 12 participants in a similar task to the present study. On a 100-point scale observers were asked to categorise movements as *angry* and *sad* and to rate the intensity of emotion. To select the actors for the current experiment, judgements about the original, pre-warped *sad* and *angry* movements were compared.

Actors were selected for the experiment in cases where the ratings of *angry* movements were judged as very angry and *sad* movements were judged as very sad. A second constraint was that few of the lifting and knocking movements be rated as *neutral*. This procedure ensured that *angry* and *sad* movements were distinguishable from each other at a baseline level. Two final constraints were that there would be an equal number of male

[15] These data were collected for this thesis, but also for a variety of other research projects. Because of this the corpus of movements was much larger than required.

and female actors and a minimum of missing data in the motion tracks. Since Walker (2000) had not included *neutral* movements, these were randomly selected for each subject. On a subject-by-subject basis, the duration of *neutral* movements was compared to the durations of the other emotions. This was done to confirm that *neutral* duration was more or less exactly half of the difference between duration for *angry* and duration for *sad* movements – this issue will be further discussed in the next section.

Morphing Procedure

The morphing procedure was effected through a number of steps:

- Firstly movements were processed to obtain their duration and other kinematics markers for speed and the differences between emotions were measured.
- A step-size of half the difference between *angry* and *sad* movements was calculated for each actor.
- The step-size was used to calculate the duration for new movements and
- An interpolation algorithm was employed to resample the movements with new duration.
- Finally, original and morphed movements were animated as point-light displays for presentation in the judgement task.

Movement Processing and Measurement of Kinematics Markers

For each movement a start and end point (P_{start} and P_{end}) was found as described in Chapter 1 (pp. 37).

Waveforms of the tangential velocity and acceleration were obtained from the second metacarpal joint (index finger) and used to calculate kinematics markers of the duration, average velocity, peak velocity and peak acceleration for each movement.

Table 4.1 Correlation between movement Duration and Velocity. Values are Pearson's r.

	Duration
Peak Velocity	-0.77
Average Velocity	-0.75

We have been primarily interested in the temporal differences between movements as a physical property that would be used for categorisation. It was reasonable to use movement duration to mediate the temporal manipulations of movements since duration correlated closely with the psychological space in Experiment 1 (see Experiment 2.2). For the *angry*, *neutral* and *sad* movements in the current experiment, there was also a smooth change in duration that correlated with and velocity for the current movements (see Table 4.1 and Figure 4.1).

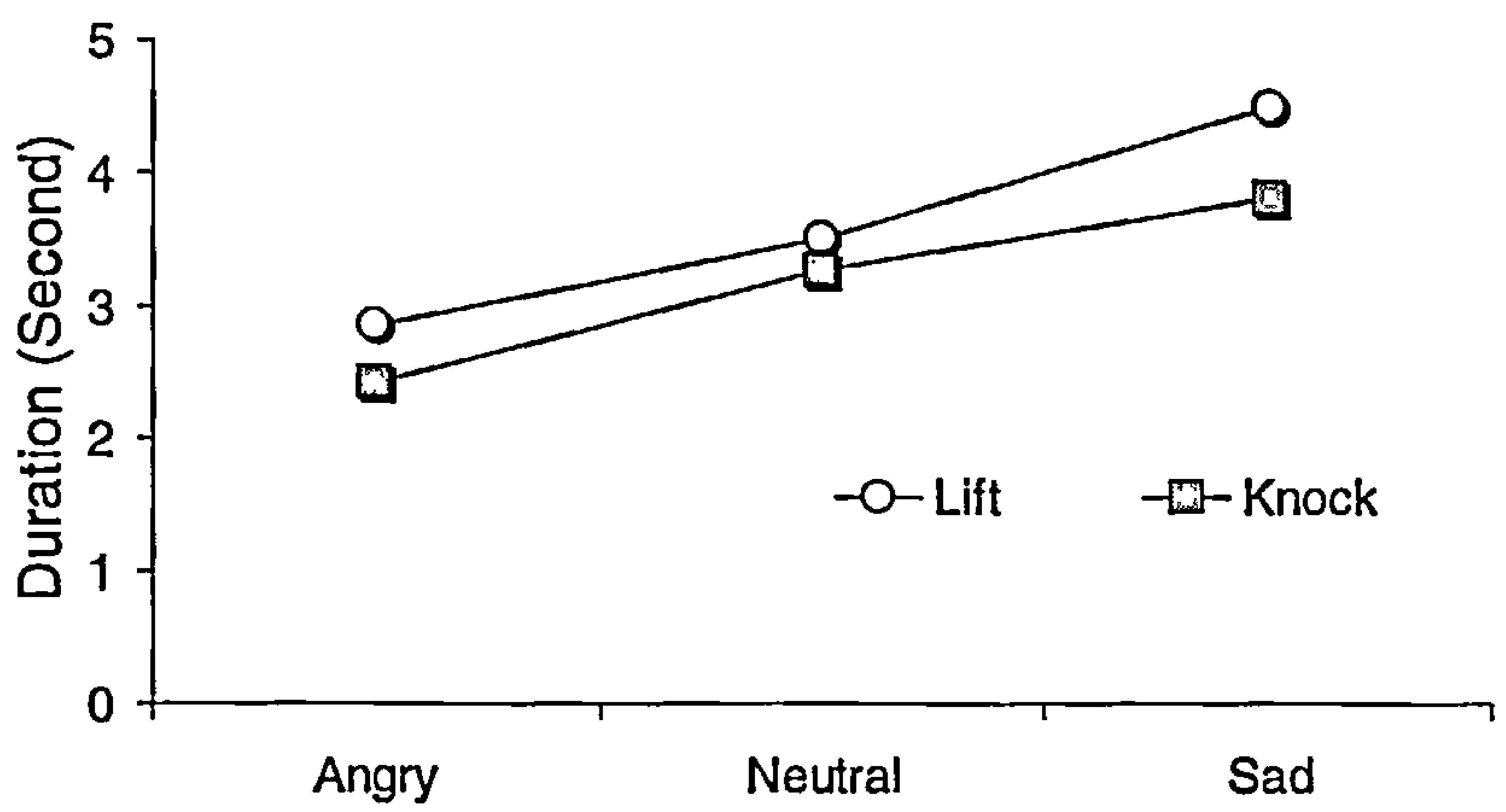

Figure 4.1 Mean Duration of movements with different emotion and action. Means are the average of natural movements for 6 actors.

Calculation of New Temporal Configurations

The next step in producing temporal morphs was to calculate the temporal configurations into which original movements would be morphed.

Firstly the start and end-points of each movement was used to calculate the natural duration of the movement from the velocity profile of the second metacarpal joint. A movement's natural duration was defined as the time taken to complete and action an was calculated with Equation 3 (pp. 64). The duration (T_t) was calculated for each of the 18 original movements.

Secondly the *sad* and *angry* movements were used to calculate a unique temporal step-size (T_{step}) for each actor. The T_{step} can be described as half the difference between *sad* duration (T_{sad}) and *angry* duration (T_{angry}) for a given actor and is depicted by the following formula:

$$T_{step} = \frac{T_{Sad} - T_{Angry}}{2}$$ **Equation 6**

Finally, three new temporal configurations were calculated for each pair of *angry* and *sad* movements. The first was a slow duration (T_{slow}) one T_{step} longer than the *sad* duration. The second was a fast duration (T_{fast}) one T_{step} shorter than the *angry* duration and the final was in the middle (T_{middle}) between *angry* and *sad*, defined as the sum of the T_{step} and the T_{angry}.

$$T_{slow} = T_{step} + T_{sad}$$ Equation 7

$$T_{fast} = T_{angry} - T_{step}$$ Equation 8

$$T_{middle} = T_{angry} + T_{step}$$ Equation 9

The duration values T_{slow}, T_{fast} and T_{middle} were used in conjunction with T_{sad} and T_{angry} to construct 12 movements with new temporal configurations for each actor. This was accomplished through the use of an interpolation algorithm (Hill & Pollick, 2000) and will be described in the next section.

Construction of Time Morphs – Movements with New Duration

The basic idea in constructing a movement with faster or slower temporal configuration than that of the original movement, is that fewer or more frames from the original movement are sampled, but that the frames are played at the original sample frequency.

For a movement that was time warped to a slower temporal configuration (for instance when an *angry* movement was morphed to the speed of a *sad* movement), frames were inserted to the number of frames in the new configuration. In the example case frames were added to the *angry* motion signal until it was the same duration as a *sad* movement, performed by the same actor. By using an interpolation algorithm the marker positions were then filled in for inserted frames and movements smoothed without changing their spatial configuration. A result of the interpolation is that the characteristic path of the wrist, as can be seen in a velocity profile for a movement, remains the same despite changes in duration.

In the case where a movement was time warped to a faster temporal configuration – for instance when a *sad* movement were time warped to the temporal configuration of an *angry* movement - frames were incrementally deleted from the motion signal until the correct number of frames remained. The interpolation algorithm was again used to smooth out the motion path of each marker. These temporal manipulations preserved the spatial information such as the distance travelled by the wrist and the shape of the arm, but changed temporal characteristics such as the duration, velocity and acceleration. An example of the above manipulations of *angry* and *sad* movements has been plotted in Figure 4.2.

In Figure 4.2a, the tangential wrist velocity profiles of a *sad* and *angry* movement for one actor is shown as they change over the course of the movement. Apart from differences in peak velocity and duration[16] that are obvious in the figure, other characteristics in the movements also differ. For instance, the knocking phase[17] of the *angry* movement was characterised by a faster completion and higher frequency signal than that of the *sad*

[16] On these graphs peak velocity is read as the maximum value on the waveform and duration can be read as the final value on the x-axis at which a value occurs in the y-axis.
[17] The knocking phase of a movement has a relatively high frequency signal that occurs between the first and final peaks in the velocity profile of a knocking movement.

movement. At the start of the knocking phase for this particular *angry* movements there was also a second peak in velocity, this was absent in the *sad* movement.

In Figure 4.2b a temporal morph of each movement is shown at the T_{middle} temporal configuration. Note that although duration was the same for *sad* and *angry* movements and peak velocity was more or less equivalent, the second peak in the *angry* movement was still evident in the knocking phase of the movement, but absent in the *sad* movement. Additionally this phase of the movement is still completed in a shorter time and has a signal with higher frequency.

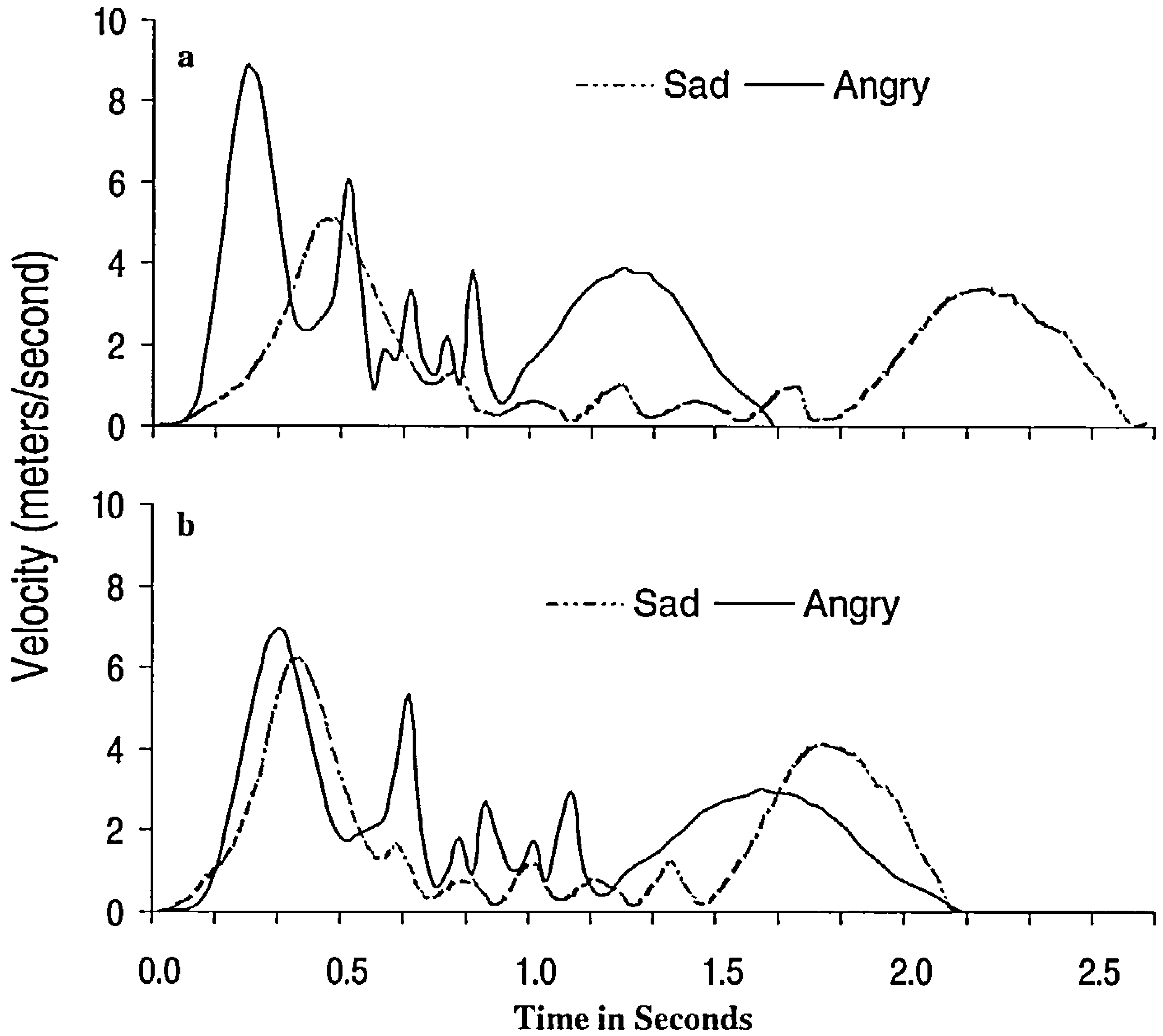

Figure 4.2 Velocity profiles of an original sad and angry movement and one time morph for each. Original movements have are plotted in a. In b temporal morphs of both movements with middle duration are plotted.

For each actor the *sad* and *angry* movements were time warped with 4 new temporal configurations that differed according to which emotion had been depicted in the natural movement.

In Table 4.2 the notation for natural movements and time morphs has been summarised. As a guide to the notation fast, angry, middle, sad and slow in subscript denotes the five possible temporal configurations. For instance T_{angry} denotes the temporal configuration of an *angry* movement. Angry, Neutral and Sad in normal script denote the emotion of an original movement that was morphed with one of the temporal configurations. They can also be regarded as representing the spatial configuration of a movement. A "natural" temporal configure denotes a movement that has not been morphed.

Table 4.2 Notation for natural movements and temporal morphs.

Original Affect	Temporal Configuration				
	T_{fast}	T_{angry}	T_{middle}	T_{sad}	T_{slow}
Angry	$Angry_{fast}$	$Angry_{natural}$	$Angry_{middle}$	$Angry_{sad}$	$Angry_{slow}$
Neutral	$Neutral_{fast}$	$Neutral_{angry}$	$Neutral_{natural}$	$Neutral_{sad}$	$Neutral_{slow}$
Sad	Sad_{fast}	Sad_{angry}	Sad_{middle}	$Sad_{natural}$	Sad_{slow}

Angry movements ($Angry_{natural}$) were speeded up with the temporal configuration T_{fast} (to become $Angry_{fast}$) and slowed down with the temporal configurations T_{middle} (to become $Angry_{middle}$), T_{sad} (to become $Angry_{sad}$) and T_{slow} (to become $Angry_{slow}$). Sad movements ($Sad_{natural}$) were slowed down with the temporal configuration T_{slow} (to become Sad_{slow}) and speeded with the temporal configurations T_{middle} (to become Sad_{middle}), T_{angry} (to become Sad_{angry}) and T_{fast} (to become Sad_{fast}). Neutral movements ($Neutral_{natural}$) were speeded with the temporal configurations T_{angry} (to become $Neutral_{angry}$) and T_{fast} (to become $Neutral_{fast}$) and slowed with the temporal configurations T_{sad} (to become $Neutral_{sad}$) and T_{slow} (to become $Neutral_{slow}$).

It was decided that as an equivalent to $Angry_{middle}$ and Sad_{middle}, *neutral* movements should appear with only their natural temporal configuration. This was because the addition of *neutral* movements with temporal configuration T_{middle} would introduce an asymmetry in the experimental design in which movements with spatial configurations of *sad* and *angry* would appear 5 times for each actor, while *neutral* movements would appear 6 times.

In light of this, it was important to establish that there were no significant differences in the duration of *neutral* movements ($T_{neutral}$) and T_{middle} duration values for actors. A repeated measure 2(action) x 2(duration) ANOVA was performed with T_{middle} and $T_{neutral}$ as dependent variables and actors as subjects. The results of the ANOVA showed no main effect of duration values [F (1,5)=.055, p=.824] and no interaction between action and duration [F(1,5)= 2.271, p= .192]. These results indicated there was little difference between $T_{neutral}$ and T_{middle}.

Experiment Methods

Participants

Ten University of Glasgow student volunteers were recruited to take part in the experiment. All observers were naïve to the purpose of the study and were paid for their participation.

Stimuli

Temporal morphs and original movements were animated as described in Chapter 2 to yield 15 stimuli for each actor. Three of these stimuli were the natural movements: *angry*, *sad* and *neutral* and the other 12 movements were temporal morphs. The interpolation preserved spatial properties of the movements such as the path length, but caused changes in spatio-temporal properties, such as the average and peak velocities.

Stimuli were presented on a graphics computer as 6 yellow dots moving on a black background from the right saggital viewpoint. Displays were 10 cm tall and 2-5 cm wide and were presented about 1 m from the observers.

Design

Presentations were arranged in 2 experimental blocks according to the action being depicted in a movement. Half of the participants saw the knocking movements first and the lifting movements second. Each block consisted of 90 stimuli (6 actors x 15 Movements) that were shown twice. This yielded 180 randomly ordered presentations in each experimental block. There were also two practice blocks, one each for knocking and lifting movements. Within each practice block the natural *sad*, *neutral* and *angry* movements of 2 actors were shown. Movements of the next best male and female actor (chosen according to the criteria described on pp. 88) were chosen for presentation in the practice.

The practice sessions took place immediately before the experimental block in which the same action was depicted. For instance a participant would see the practice session for knocking movements followed by the experimental session in which knocking movements were shown.

Procedure

At the start of the experiment each participant was given a set of written instructions that detailed the procedure for the experiment.

Observers were told that they would see human knocking and lifting movements and that their task would be to categorise the movements as angry, neutral or sad, then to rate the intensity of the emotion. It was explained that on a 100-point scale a score between 1 and 49 would indicate that the movement was perceived as angry, 1 being intense anger and 49, mild anger. A score of 51-100 would indicate that the movement was perceived as sad, 51 being mildly sad and 100, intense sadness. Finally a score of 50 would indicate that the movement was neutral. The scale has been replicated in Figure 4.3 as it was shown to participants in the instructions. Throughout the experimental session participants had free access to the figure and other experimental instructions.

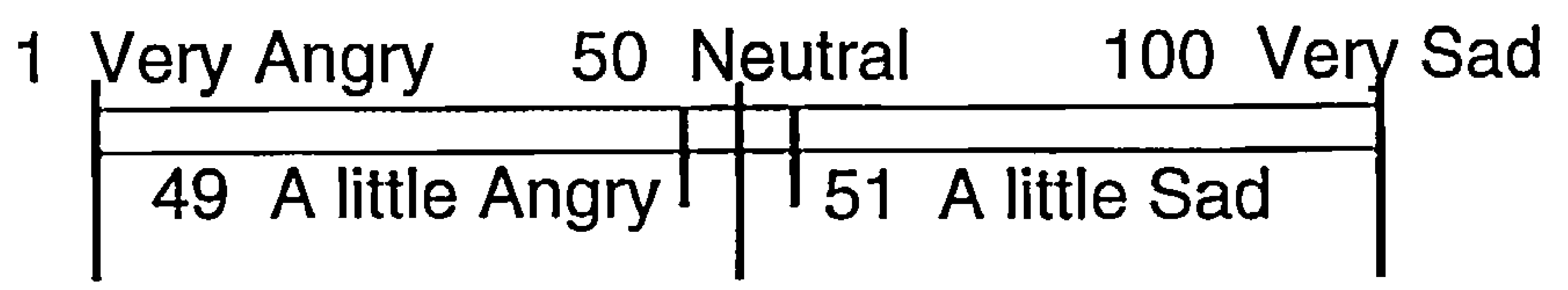

Figure 4.3 Replication of the rating scale for Experiment 5 as it appeared in the instruction to participants.

Finally observers were instructed to make a guess when they were uncertain of the answer. In particular they were told that there would be *neutral* movements and that they should not use "neutral" to indicate that they did not know the answer.

For each trial, the participant viewed a computer display of a movement followed by a dialog box in which they had to make their response. The dialogue box contained the 100-point scale that the participant could manipulate by moving a sliding bar.

Results and Discussion

It was expected that the original *sad* and *angry* movements would be distinguished from each other at a high rate of accuracy. Additionally participants would rate a movement that had been speeded, as *angry* and a slowed movement as *sad*. Slower movements would be rated as more intensely *sad* while faster movements would be more rated intensely *angry*. Additional expectations were that there would be statistical differences in the speed of *angry* and *sad* movements and that there would be a correlation between kinematics markers of original movements and temporal morphs with ratings of emotion intensity.

Comparison between Kinematics Markers of Angry, Neutral and Sad

Figure 4.1 (pp. 90) showed the differences in duration between *angry*, *sad* and *neutral* movements for knocking and lifting actions averaged over the 6 actors. From this figure it was clear that *angry* movements had the shortest duration and consequently would have the highest speeds. *Sad* movement had the longest duration and lowest speeds and the duration and velocity of *neutral* movements were more or less half way between *angry* and *sad*.

These differences were confirmed in a 2(action) x 3(emotion) repeated measures (ANOVA) with different actors as subjects. Results from the ANOVA indicated that there is a main effect of action [$F(1,5)=15.454$, $p<.05$], with knocking movements being faster (3.19 s) than lifting movements (3.63 s). More interestingly there was also a significant main effect of emotion [$F(2,10)=53.459$, $p<.001$] and post hoc Tukey tests revealed honestly significant differences (HSD's) in all pair-wise comparisons among means at $p< .01$ (comparisons were: *angry* and *neutral*, *angry* and *sad*, *neutral* and *sad*). There was no interaction between action and emotion.

We also checked for differences between actors and found none. A 2 x 6 mixed design ANOVA was calculated using the three movements that had been recorded for each actor as subjects, actors as the between groups variable and action as the within groups variable. The results showed a main effect of action [$F(1, 12)= 20.456$, $p<.001$] with lifting movements having a longer duration than knocking movements. There was no main effect of actor [$F(5, 12)= 0.443$, ns] and no interaction between emotion and action [$F(5, 12)= 1.32$, ns].

These results were as expected and indicated that kinematics markers for the natural *angry*, *sad* and *neutral* human movements were reliably different from each other.

Correct Recognition of Natural Movements

The above results showed that there were reliable differences in the movement kinematics, however whether these differences were within a range that would allow the discrimination of *sad*, *angry* and *neutral* movements in a perceptual task, was still uncertain.

A reanalysis of the preliminary data from Walker (2000) showed that the *angry* and *sad* movements used in that experiment could be correctly recognised about 81% of the time in a rating task. *Angry* and *sad* movements were confused for the other emotion 9% of the time and were categorised as *neutral* on 10% of trials. This analysis coupled with results from Experiment 2 was encouraging in that it presented a good case for *angry* and *sad* movements to be discriminated from each other.

The data for natural movements was separated from other data and converted to proportion correct recognition. To convert the data to proportion correct, all intensity ratings of less that 50 were recorded as a categorisation of *angry*, ratings of 50 were recorded as *neutral* and responses of greater than 50 were recorded as *sad* categorisations.

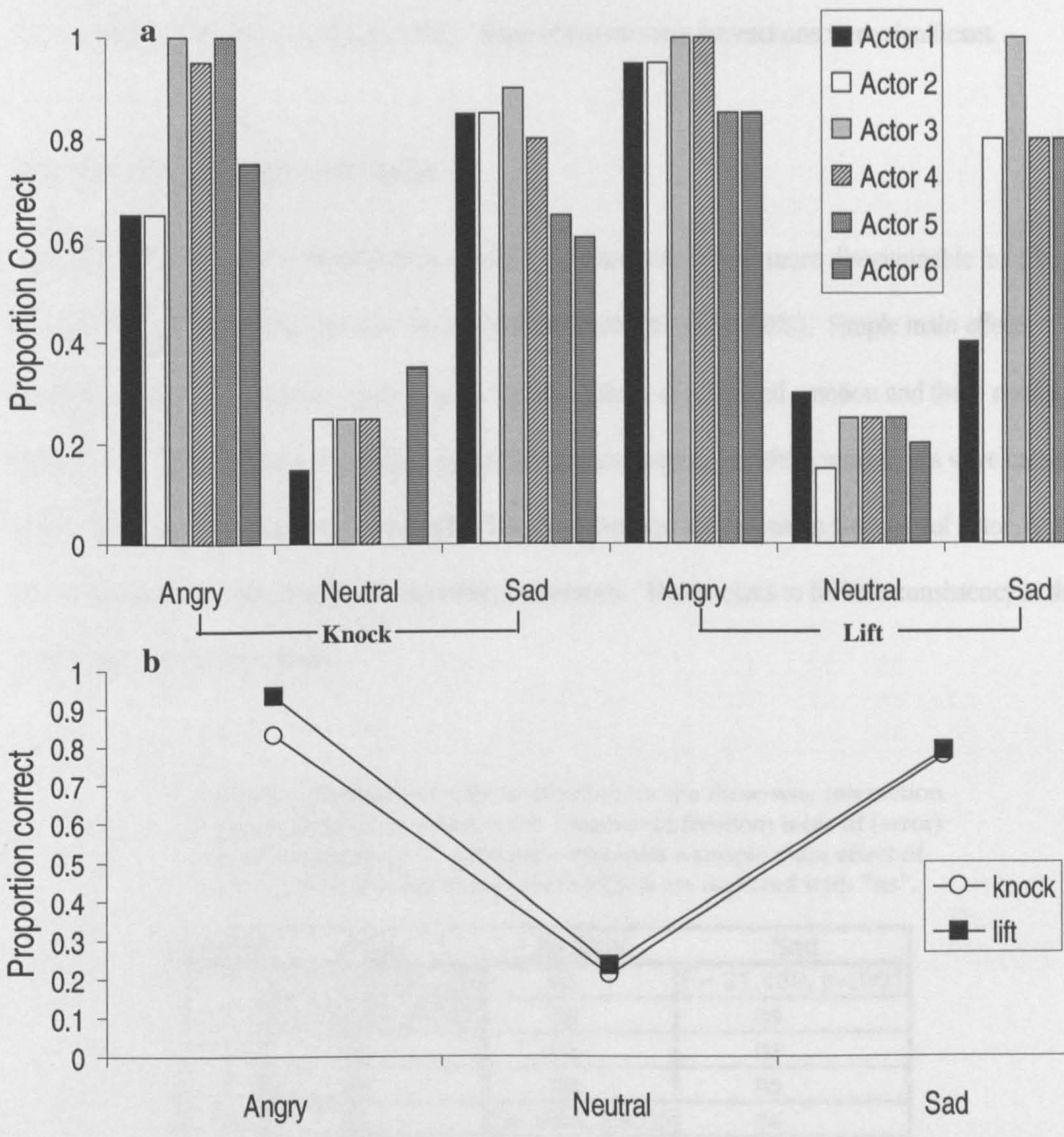

Figure 4.4 Mean proportion of correct responses to natural movements with affect (angry, neutral and sad) plotted a) across actor and action and b) across action.

Correct responses to *angry*, *neutral* and *sad* movements are plotted in Figure 4.4 to show accuracy according to actors and actions. From Figure 4.4b it is clear that *angry* and *sad* movements were categorised with much greater accuracy than *neutral* movements. Figure 4.4a show that for different actors, observers responded with different accuracy, however, for all actors and actions, *angry* and *sad* movements are recognised with greater accuracy than *neutral* movements.

To test these observations a repeated measures 2(action) x 3(emotion) x 6(actor) ANOVA was performed. The results showed significant main effects for action [$F(1, 9)= 6.178, p<.05$], emotion [$F(2, 18)= 38.98, p<.001$] and actor [$F(5, 45)= 5.361, p<.001$] and a significant three-way interaction between actor, action and emotion [$F(10, 90)= 4.303, p<.001$]. None of the two-way interactions were significant.

Correct Recognition of Action

The main effect of action indicated that, overall, emotions were slightly more discriminable for lifting movements (accuracy 65%) than for knocking movements (accuracy 60%). Simple main effects of action for the three-way interaction were significant for 5 combinations of actor and emotion and these results are tabulated in Table 4.3. In 4 out of the 5 cases the emotion depicted in lifting movements were categorised with higher accuracy than knocking movements. The exception was for the *sad* movements of Actor 1, where the emotion was better recognised from knocking movements. There seems to be little consistency in the pattern of simple main effects for actions.

Table 4.3 Simple main effects of action for the three-way interaction between action, affect and actor. Degrees of freedom were df (error) = 9 and df (action) = 1. Each cell represents a simple main effect of action, non-significant simple main effects are depicted with "ns".

Actor	Angry	Neutral	Sad
1	F= 13.853, p<.05	ns	F= 31.169, p<.001
2	F= 13.853, p<.05	ns	ns
3	ns	ns	ns
4	ns	ns	ns
5	ns	F= 9.62, p<.05	ns
6	ns	ns	F= 18.855, p<.001

Correct Recognition of Actors

For the Main effect of actor, post hoc Tukey tests of the HSD's of differences between actors showed that for Actor3 there was better recognition than for Actors 1, 2 and 5 (significant at $p<.05$). The movements of Actor4 were also recognised with greater accuracy than those of Actor1 (significant at $p<.05$).

Table 4.4 Simple main effects of actor for the three-way interaction between action, affect and actor. Degrees of freedom were df (error) = 45 and df (action) = 5. Each cell represents a simple main effect of action, non-significant simple main effects are depicted with "ns".

Actor	Knock	Lift
Angry	$F= 6.037$, $p<.001$	ns
Neutral	$F= 3.036$, $p<.05$	ns
Sad	$F= 3.106$, $p<.05$	$F= 9.354$, $p<.001$

Table 4.5 Post hoc Tukey tests of actor for the three-way interaction between action, affect and actor. Results have been plotted as levels of significance according to action and affect.

Affect	Action / Actor	Knock 1	Knock 2	Knock 3	Knock 4	Knock 5	Knock 6	Lift 1	Lift 2	Lift 3	Lift 4	Lift 5	Lift 6
Angry	1												
	2	ns						ns					
	3	*	*					ns	ns				
	4	*	*	ns				ns	ns	ns			
	5	*	*	ns	ns			ns	ns	ns	ns		
	6	ns	ns	*	*	*		ns	ns	ns	ns	ns	
Neutral	1												
	2	ns						ns					
	3	ns	ns					ns	ns				
	4	ns	ns	ns				ns	ns	ns			
	5	ns	*	*	*			ns	ns	ns	ns		
	6	*	ns	ns	ns	*		ns	ns	ns	ns	ns	
Sad	1												
	2	ns						*					
	3	ns	ns					**	*				
	4	ns	ns	ns				*	ns	*			
	5	*	*	*	ns			*	ns	*	ns		
	6	*	*	*	*	ns		*	ns	ns	ns	ns	

Levels of α: ** $\alpha=.01$, * $\alpha=.05$
ns = non-significant

For the three-way interaction between action, actor and emotion, there were simple main effects of actor for knocking movements at each level of emotion. For lifting movements there was a simple main effect of actor only for movements with *sad* emotion. These simple main effects have been tabulated in Table 4.4. Levels of significance for post hoc Tukey test have been tabulated in Table 4.5. From these tables it is clear that the differences between actors were best recognised from knocking movements, but also from lifting movements with *sad* emotion.

Despite the significant effects of actors and actions, there was a pattern in the results that suggests that *angry* and *sad* movements were easily recognised, but *neutral* movements were harder to recognise. These observations are analysed further in the next section.

Main Effect of Emotion – Correct Recognition of Natural Movements

As described above (pp. 98-99) there was a main effect of emotion [$F(2, 18) = 38.98$, $p<.001$]. Post hoc analysis of this effect[18] showed that *neutral* movements (accuracy 22%[19]) were recognised with less accuracy than *angry* (accuracy 88%) and *sad* (accuracy 78%) movements, but that there was no significant difference between the percentage correct recognition of *angry* compared to *sad* movements. T-tests were performed to test whether *angry*, *sad* and *neutral* movements were recognised at levels greater that chance (chance = 33.33 %). Accurate responses to *angry* and *sad* movements were significantly greater than chance [*angry* $t=24.75$, $df=9$, $p<.001$ (1-tailed); *sad* $t=13.272$, $df=10$, $p<.001$ (1-tailed)]. For *neutral* movements, recognition was not significantly better or worse than chance [$t=1.21$, $df=9$, ns, (2-tailed)].

The confusions between emotions were also examined in more detail and seemed to indicate that *neutral* movements were confused for *sad* movements, while recognition of *sad* and *angry* movements were at ceiling level and hence confusions were at or below the level of chance. Figure 4.5 is a plot of the confusions between movements.

[18] For each pair of affects the Honestly Significant Difference (HSD) was calculated using a Tukey test.
[19] Percentages have been rounded to the nearest full percent in most cases.

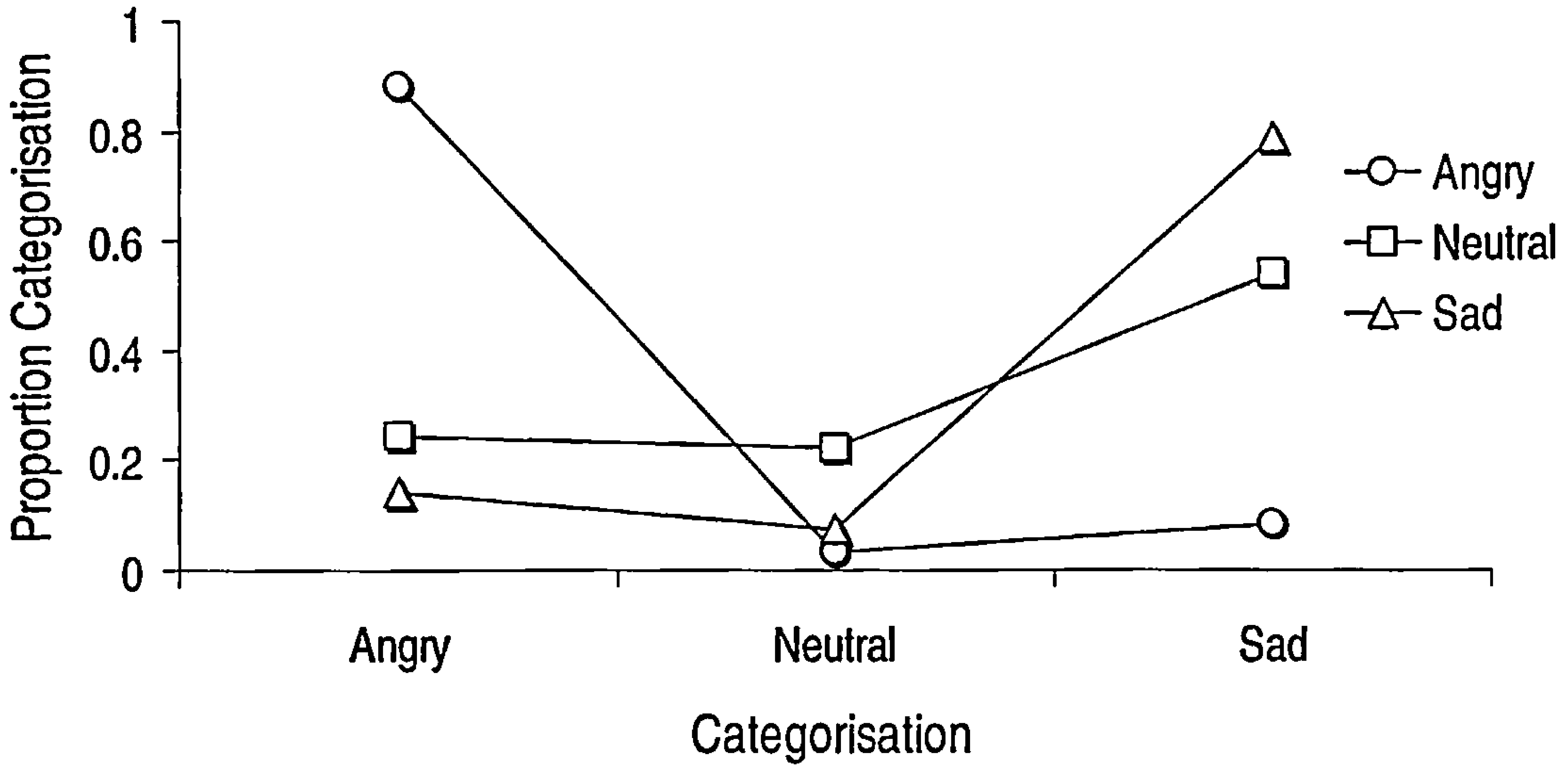

Figure 4.5 Proportion of categorisations for angry, neutral and sad movements. On the horizontal axis the three choices that were available to observers are depicted.

In general there was a tendency for observers to categorise movements as either *sad* or *angry* (averaged over all trials there were 42% *angry* and 47% *sad* responses) rather than as *neutral* (11%). This bias towards categorising movements as emotive was possibly due to the rating scale used in the experiment (see Figure 4.3). On the scale there was a much larger range of responses available for *angry* and *sad* judgements than for *neutral*. It is possible that participants were not very careful in manipulating the scale with which they indicated movement categories.

To establish whether this could be the case, rating data was re-coded so that neutral answers were depicted by ratings between 45 and 55, angry categorisations were depicted by ratings of less that 45 and sad categorisations were depicted by ratings greater than 55. With this new classification, each emotion label was used on roughly a third of all trials (angry 35%; neutral 29%; sad 36%). It was decided, however, not to use these re-coded scores, since observers had been instructed to be accurate in their use of the scale. Additionally there was no indication during debriefing sessions that participants had trouble to understand any of the instructions or that they found it hard to use the rating scale.

An alternative reason for the bias towards using emotion labels could be that participants found it hard to distinguish *neutral* movements from one or both of the emotions. This seems to have been the case since *neutral* movements were confused as *sad* 54% of the time, this was a significantly better than chance [t=2.95, df=9, p<.01, one-tailed]. *Neutral* movements were confused for *angry* 24% of the time, which was just worse than chance [t=, df=9, p=.042, one-tailed]. When the ratings were re-coded as described above to expand the range of ratings categorised as *neutral*, the accurate identification of *neutral* movements rose to 43%. The confusion with *sad* still remained with *neutral* movements being categorised as *sad* 40% of the time, but the confusions with *angry* was only 16%. These results suggest that *neutral* movements looked similar to *sad* movements.

Sad movements were, however, not confused for *neutral* with equal frequency. In fact, *sad* movements were categorised as *neutral* movement on 8% of trials compared with 14% confusion as *angry* movements. *Angry* movements were the best-recognised (accuracy 88%) and confusions with *sad* and *neutral* emotions were minimal (3% and 8% respectively).

The finding that *neutral* movements were rated as *sad* may have been due to the instructions given to participants. In the instructions, it was emphasised that there were *neutral* movements and that a "neutral" categorisation should not be used in place of "I don't know", this could have biased observers against using the neutral category.

Simple Main Effects of Emotion

As described earlier there was a three-way interaction between action, emotion and actor emotion [$F(10, 90) = 4.303$, $p<.001$]. Simple main effects of emotion for the three-way interaction are shown in Table 4.6.

Table 4.6 Simple main effects of emotion for the three-way interaction between action, affect and actor. Degrees of freedom were df (error) = 9 and df (action) = 1. Each cell represents a simple main effect of action, non-significant simple main effects are depicted with "ns".

Actor	Knock	Lift
1	ns	ns
2	ns	F= 4.605, p<.05
3	F= 4.223, p<.05	F= 4.775, p<.05
4	ns	F= 3.841, p<.05
5	F= 6.558, p<.05	ns
6	ns	F= 4.223, p<.05

In half the cases there were simple main effects of emotion, however, the trend in all cases were for *angry* and *sad* movements to be recognised with greater accuracy than *neutral* movements. Post hoc Tukey tests of the HSD showed that for 10 out of the 12 combinations of action and actor, *angry* and *sad* movements were recognised correctly with similar frequency while *neutral* movements were recognised correctly with significantly less accuracy. The exceptions were for the lifting movements of Actor1 and knocking movements of actor 6. In these cases *angry* and *neutral* movements were significantly different, however, *angry-sad* and *sad-neutral* comparisons were not. All significant Tukey tests were significant at $\alpha=0.05$.

These results show that despite noise from actor and action variables, overall *sad* and *angry* could be accurately distinguished from each other. *Neutral* movements were not as well recognised, but this was possibly due to the way in which observers used the rating scale.

Intensity Ratings

Intensity ratings were analysed with the hypothesis that judgements of the movements would be modulated with a modulation of temporal configurations for movements. Ratings of intensity were analysed in two ways. Firstly in an analysis of variance – described and secondly they were correlated with movement kinematics.

Analysis of Variance for Intensity Ratings

To analyse the ratings with reference to all variables in the data, a 2(action) x 6(actor) x 3(emotion) x 5(duration) repeated measures ANOVA was performed. The results, however, were rather confusing because of the large number of complex interactions. There was no main effect of action and this variable was involved in the fewest interactions, so it was decided to disregard it and calculate a less complex ANOVA.

A 3(emotion) x 5(duration) x 6(actor) repeated measures ANOVA was performed by using the intensity rating for each movement, averaged over actions and repetitions of movements. The results are shown in Table 4.7 and the ratings data has been graphed in Figure 4.6 to show ratings for actors at each level of emotion and duration. Ratings of 50 were taken to mean that movements were categorised as depicting *neutral* emotion. Values above 50 indicated a categorisation of *sad* and the greater the value, the more intense was the judged *sad*ness. Values below 50 indicated that observers thought the movement depicted *angry* emotion and lower values indicated greater intensity of emotion than larger values.

Table 4.7 Results for an ANOVA of affect x duration x actor for intensity ratings. Non-significant results are depicted as

Source of Variation	
Affect	F(18, 2)=90.953, p<.001
Duration	F(36, 4)=99.264, p<.001
Actor	F(45, 5)=11.77, p<.001
Affect x Duration	F(72, 8)=1.315, ns
Affect x Actor	F(90, 10)= 22.1, p<.001
Duration x Actor	F(180,20)= 3.91, p<.001
3-way Interaction	F(360,40)= 1.7, p<.01

Results from the ANOVA showed main effects for emotion, duration and actor. There were also significant two-way interactions between emotion and actor and between actor and duration, but no interaction between emotion and duration. Finally there was a significant three-way interaction as depicted in Table 4.7.

With reference to Figure 4.6, the 3-way interaction indicates that there was a general trend for movements with longer duration (sad and slow temporal configurations) to be rated with larger values – indicating ratings of sadness. Movements with middle temporal configuration were rated at around 50 and movements with short duration were rated at lower values on the scale. This trend extended over different emotions, however there were individual differences in the ratings for actors.

In Chapter 3, an analysis of the movements of different actors showed that there are some individual differences in the kinematics of different actors as they perform the same movements. For the movements used in the current experiment, however, it was not possible to show statistically significant differences between actors for the duration of movements. Additionally the movements of different actors were not presented in different experimental blocks, or identified to observers in any other way.

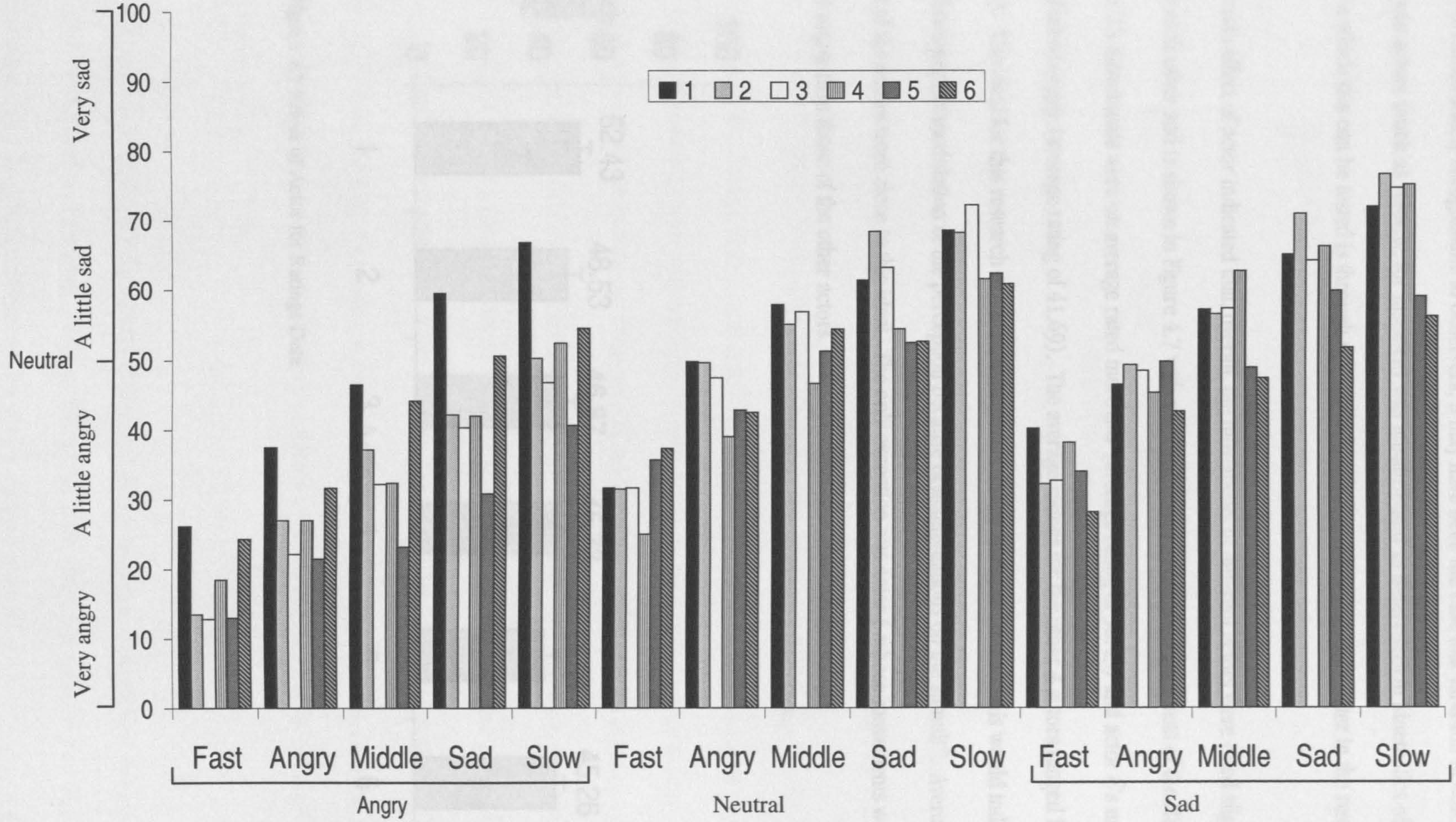

Figure 4.6 Rating plotted according to affect, duration and actor. The ratings scale used by observers have been reproduced on the x-axis data points have been averaged over action and movement repetitions.

As a consequence, most observers either did not think about actors as they did the task (4 out of 10 observers) or thought that a single actor had made that all movements (3 out of 10 observers). Since the different actors were not immediately recognisable to observers, it may have been the case that variation in other kinematics between actors (such as velocity, for instance) was misattributed as differences in kinematics of emotions. One way in which this can be tested is through regression analysis and this is covered later in the result section.

The main effect of actor indicated that, overall, the movements of different actors were rated slightly differently from each other and is shown in Figure 4.7 with standard errors. Post hoc Tukey tests of the HSD's showed that Actor 2's movements were on average rated more *sad* (average rating of 52.43) and actor 6's movements were rated more *angry* (average rating of 41.69). The average ratings for the other 4 actors ranged from 45.26 to 48.53. The ideal for this research would have been an average rating of 50 since this would indicate that there had been perfect modulation of the perception because of manipulations on the stimuli[20]. Average ratings for most of the actors were close to this ideal. The only exception was actor 6 whose movements were judged to be more *angry* than those of the other actors.

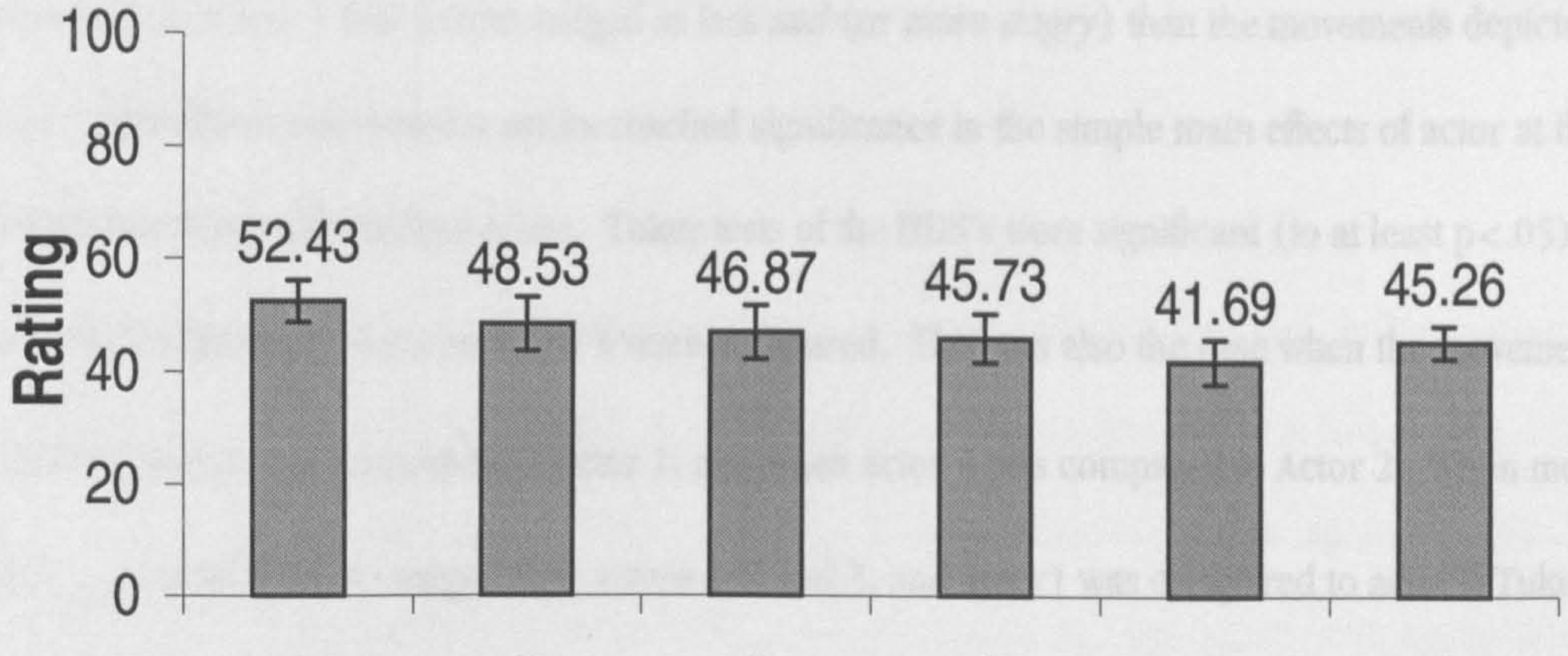

Figure 4.7 Effect of Actor for Ratings Data

[20] It could also mean that observers used only values close to 50 when rating the movements, but this was not the case as it was clear from Figure 4.6 that a range of values were used by observers to rate the movements.

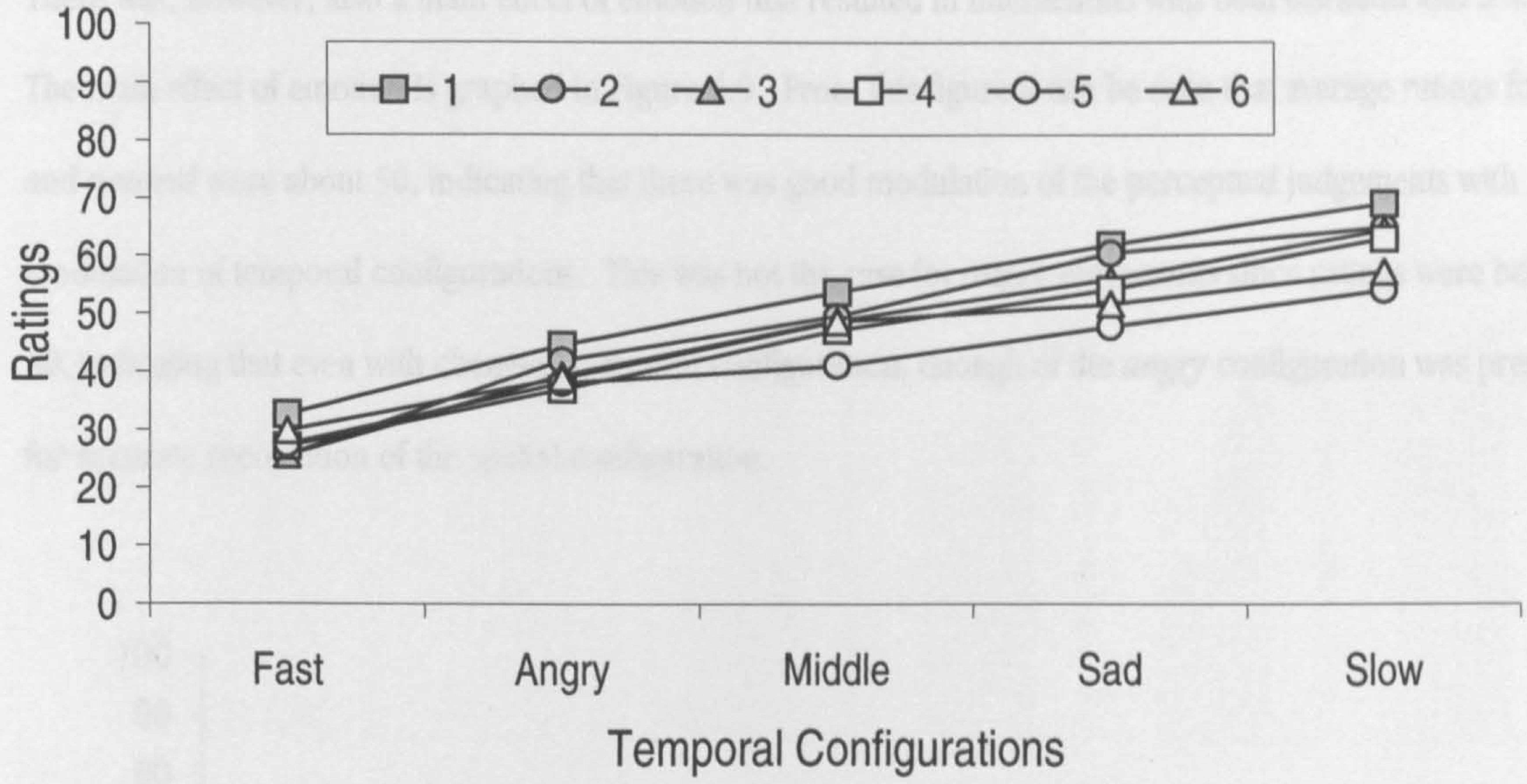

Figure 4.8 Interaction between Actor and Temporal Configuration for Ratings Data

Results of the ANOVA showed an interaction between actor and movement duration (graphed in Figure 4.8). For the most part this interaction shows little that is new, apart from to that for movements with T_{fast} and T_{angry} actors' movements were not rated significantly differently. For movements with T_{middle}, T_{sad} and T_{slow} the movements of actors 5 and 6 were judged as less *sad* (or more *angry*) than the movements depicted by other actors. The differences between actors reached significance in the simple main effects of actor at the middle, sad and slow temporal configurations. Tukey tests of the HDS's were significant (to at least $p<.05$) when the movements with T_{middle} of actors 1 and 5 were compared. This was also the case when the movements with T_{sad} of actors 5 and 6 were compared to Actor 1, and when actor 5 was compared to Actor 2. When movements with T_{slow} of actor 5 were compared to actors 1, 2 and 3, and Actor1 was compared to actor 6 Tukey tests were significant. Overall, however, there was a trend for judgements to be modulated with changing temporal configuration and there were simple main effects of duration at every level of actor for the interaction between actor and duration.

The finding of a modulation of judgements with a modulation in temporal configuration was borne out in the main effect of temporal configuration. Post hoc Tukey tests of the HSD's between different temporal configurations showed significant differences for all comparisons at $p<.01$.

There was, however, also a main effect of emotion that resulted in interactions with both duration and actor. The main effect of emotion is graphed in Figure 4.9. From this figure it can be seen that average ratings for *sad* and *neutral* were about 50, indicating that there was good modulation of the perceptual judgements with modulation of temporal configurations. This was not the case for *angry* movements since ratings were below 50, indicating that even with changing temporal configuration, enough of the *angry* configuration was preserved for accurate recognition of the spatial configuration.

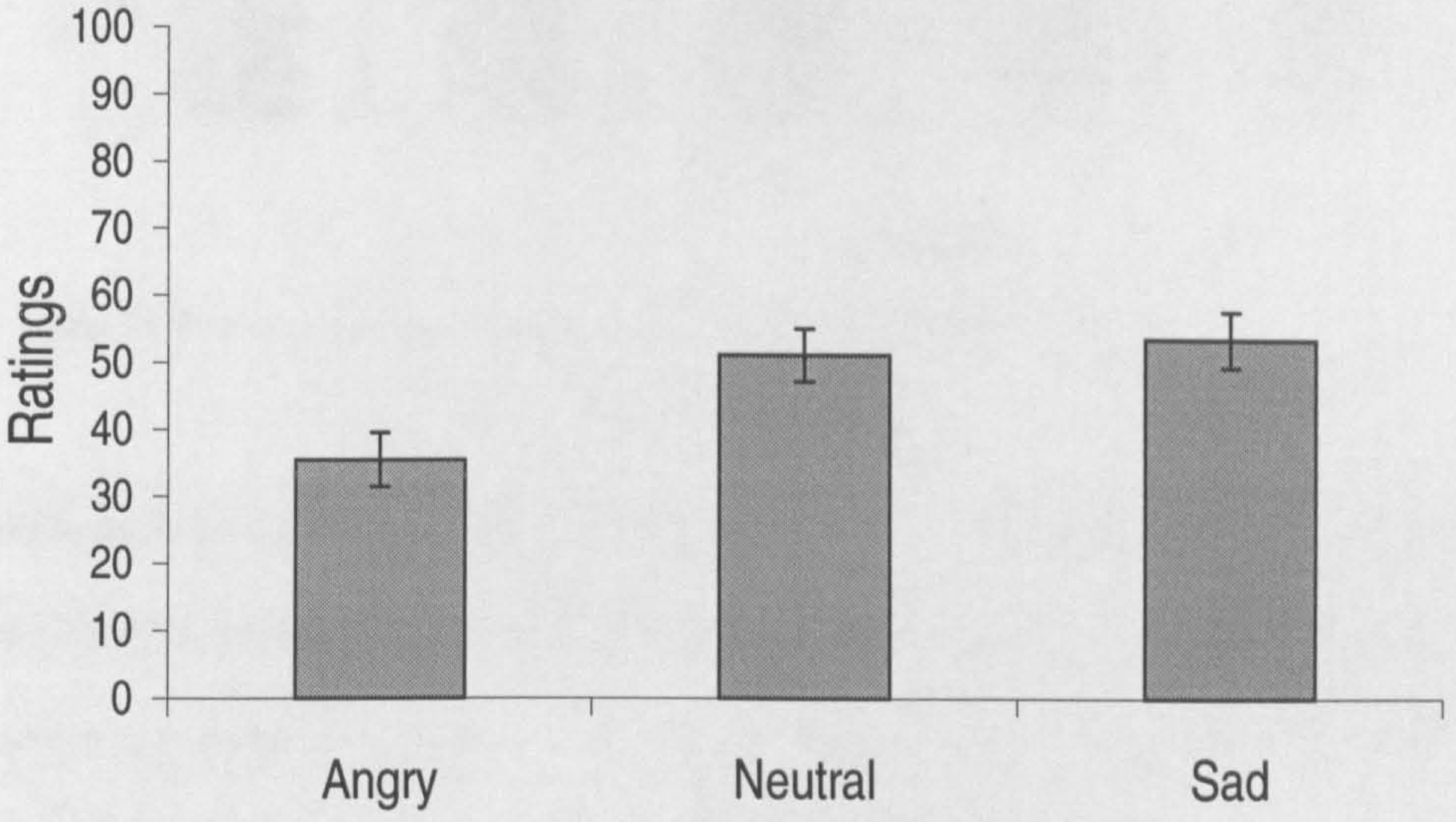

Figure 4.9 Main Effect of Emotion for Intensity Ratings

The AVOVA also showed an interaction between actor and emotion that is illustrated in Figure 4.10. For this interaction there were simple main effects of actor at all levels of emotion and post hoc analysis for the HDS's between actors showed that for movements with *angry* emotion, there were more significant differences between actors that for movements with *neutral* and *sad* emotion. For movements with *angry* emotion, there were mainly HSD's when Actors 1 and 5 compared to other actors. Ratings for Actors 3 and 6 were also different from each other. For movements with *neutral* emotion there were significant differences when Actor4 was compared to Actors 1, 2 and 3. There was also a trend for ratings of *angry* movements to have lower ratings than those for *neutral* movements, and for ratings of *sad* movements to be greater than either of these.

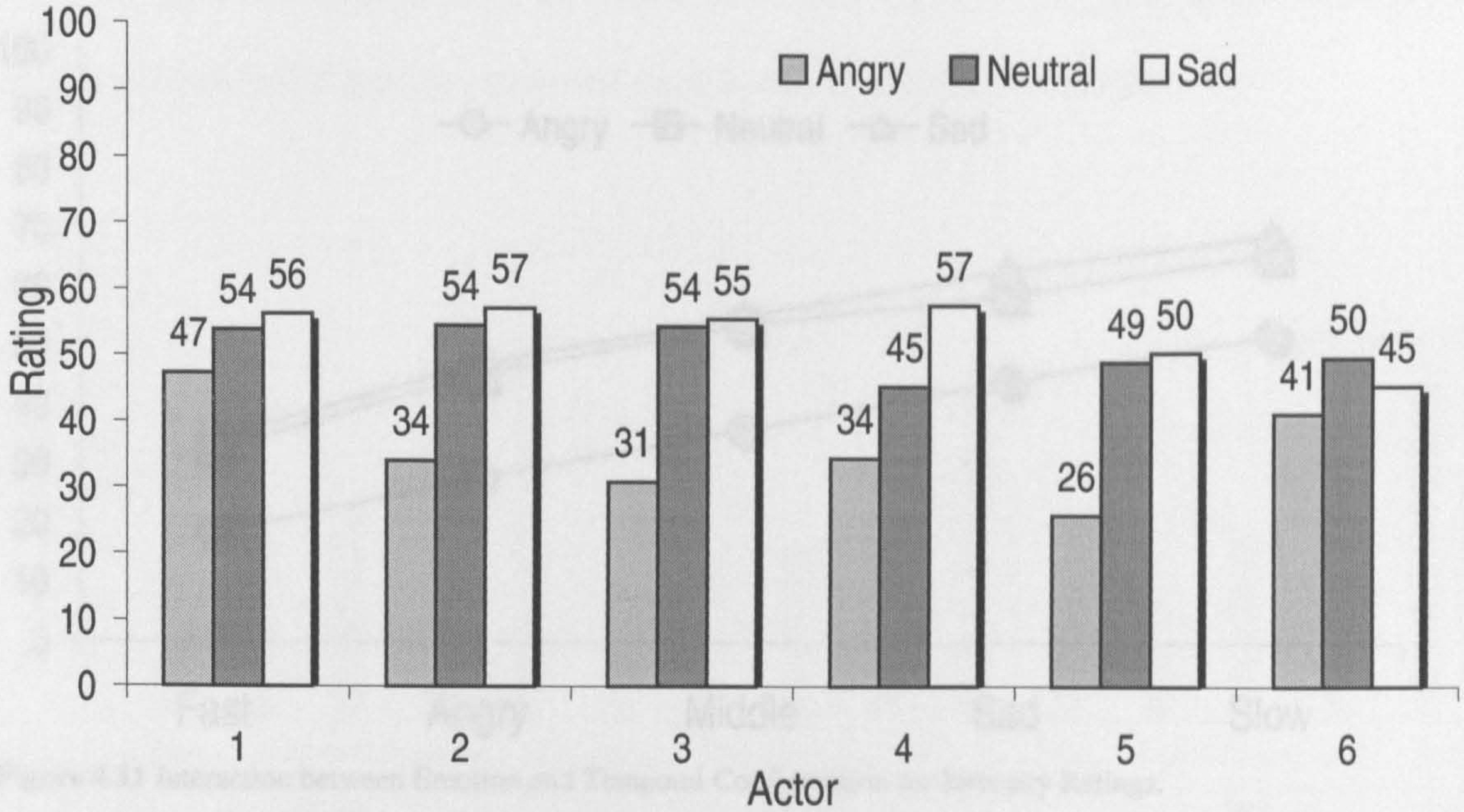

Figure 4.10 Interaction between Emotion and Actor for Ratings Data

Simple main effects of emotion were significant for 4 out of the 6 actors (Actors 2,3,4 and 5). Post hoc Tukey tests for these actors indicated that *angry* movements were rated with significantly lower values than *neutral* and *sad* movements on all 4 cases. However, only for Actor4 were the sad movements rated significantly differently than *neutral* movements. Post hoc tests of the main effect of emotion of emotion also showed that *angry* movements were rated with values lower than those of *sad* and *neutral* movements were and that there was no significant difference between *sad* and *neutral* movements.

Finally there was an interaction between emotion and temporal configuration that echoed those above to show that ratings for *angry* movements were lower than for *sad* and *neutral*. This interaction has been plotted in Figure 4.11. From Figure 4.11 it is clear that despite the effect of emotion, there was still modulation of judgements for *angry* movements so that faster movements were rated more *angry* than slower movements.

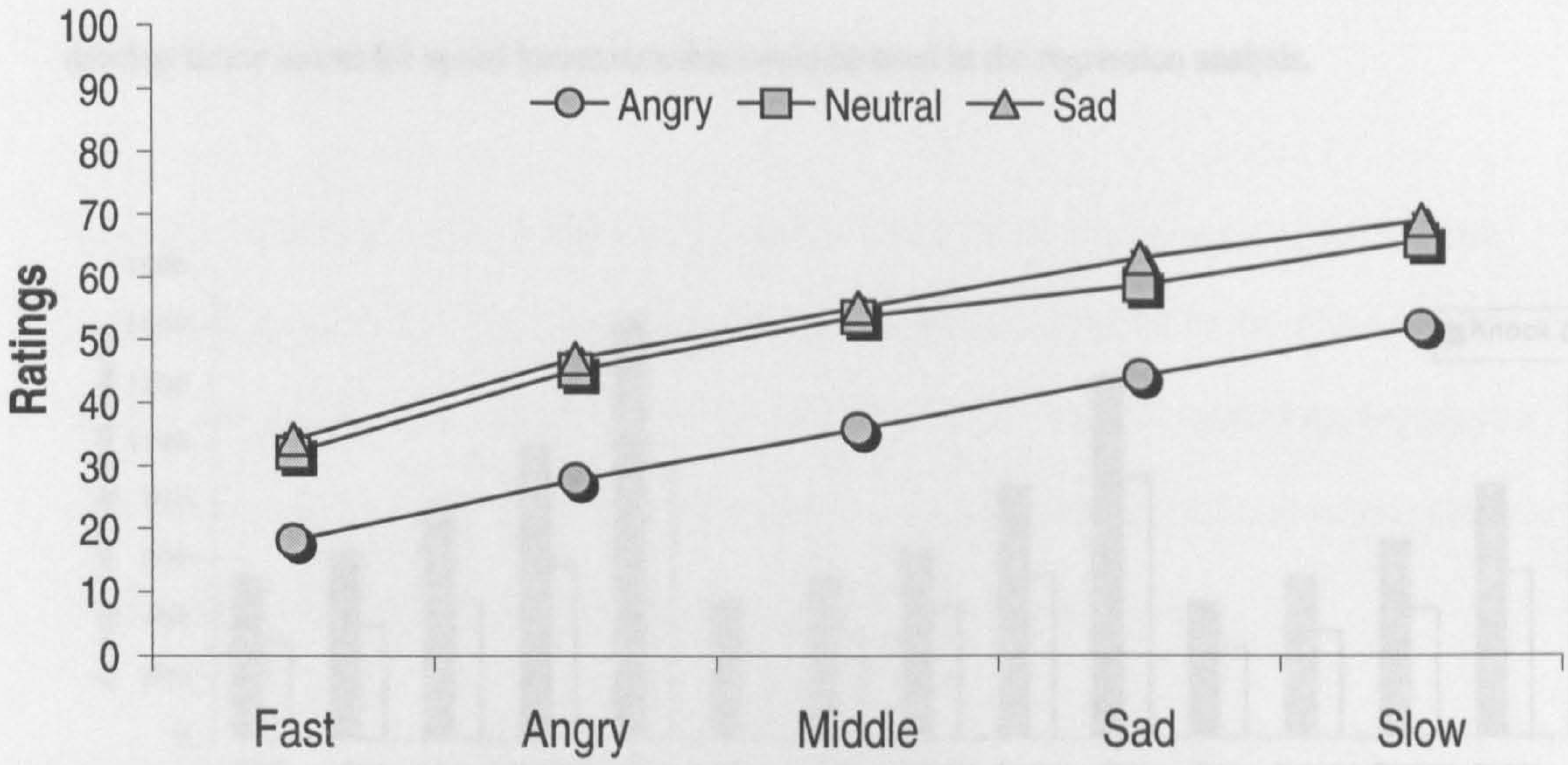

Figure 4.11 Interaction between Emotion and Temporal Configuration for Intensity Ratings.

In summary, the intensity ratings collected from participants showed that overall there were an effect of temporal configuration. However, movements with *angry* emotion were rated to be *angry*, despite changes in duration. A reason for this last finding may be that changes in movements' duration did not result in perfect modulation of speed. This was tested in a Regression analysis that is described in the next section.

Regression with Intensity Ratings and Kinematics

Results from the above ANOVA have shown that there was a modulation of participants' ratings of emotion as temporal configurations were modulated. In this section this relationship between perception and velocity will be further explored. Figure 4.12 illustrates the relationship between speed and judgements.

From Figure 4.12 it is clear that as movement velocity increased, judgements about the emotion also change so that they are recognised as *angry* with increasing intensity.

To examine the relationship between judgements made by observers and movement speed, kinematics markers were used to calculate the regression with ratings of intensity. In this case it was not sensible to use multiple regression to parse out the variance that different kinematics markers would account for since kinematics markers for speed are highly correlated with each other. Instead, Principle Components Analysis was used to develop factor scores for speed kinematics that could be used in the regression analysis.

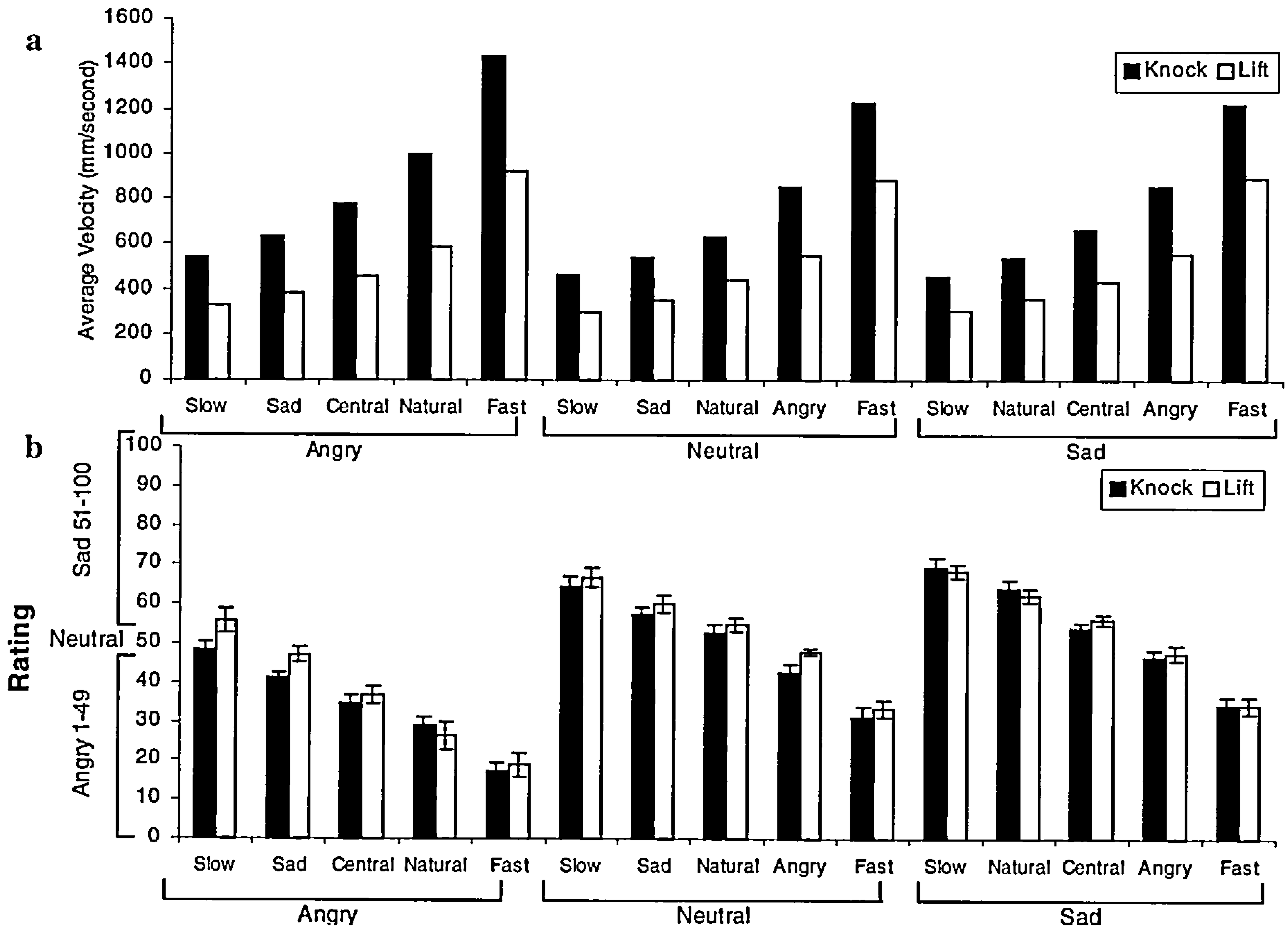

Figure 4.12 Comparison between a) average velocity and b) ratings data.

All kinematics markers were used in factor analysis to produce factor scores with the Anserson-Rubin adjustment. Since there was little effect in the ANOVA for different actions, kinematics markers were averaged over action to produce 90 values that represented movements from 6 actors with 3 different emotion and 5 different duration. For each of these a factor score was produced by using the kinematics markers of duration (in time), velocity (peak and average velocity) and acceleration (peak acceleration). Factor scores are composite values that represent a single factor resulting from factor analysis.

In this case Principal Components Analysis was used to identify factors in the speed kinematics of movements with different emotion and duration. It is not surprising that one factor accounted for most of the variance in the data (83.5%). The composite factor scores for this factor then represents the speed kinematics for the movement data.

These speed scores were used in a simple regression calculation to fit a linear model to the data from observers' judgements about the movements. Speed kinematics accounted for 57.2% of the variance in the intensity ratings. [r^2=.572 (adjusted), F(88,1)=120.187, p<.001] and the kinematics markers made a significant contribution to the model [β (standardized coefficient)=-.760, t=-10.963, p<.001]. This is a reasonable result, however, it did not show perfect modulation of the judgements.

In light of finding a significant effect of emotion in the previous ANOVA, it seemed reasonable to calculate the correlation with intensity ratings for different emotions. The correlation between intensity ratings and duration are shown in Figure 4.13 and for each emotion a trend-line has been sketched with the corresponding r^2 value.

From this figure it is clear that judgements about *sad* movements are almost perfectly modulated by manipulation of movement duration. This is also the case for *neutral* moments, although the modulation is not as good. Judgements about *angry* movements, however, seem to be based only partially on the duration of movements. For correlation with kinematics markers, duration was the best correlate with intensity ratings in the case of each emotion. The results of the correlation with each kinematics marker are shown in Table 4.8.

Table 4.8 Correlation between kinematics markers and intensity ratings. All values are Pearson's r.

	Angry	Neutral	Sad
Duration	0.810	0.912	0.959
Peak Velocity	-0.671	-0.793	-0.790
Average Velocity	-0.622	-0.788	-0.821

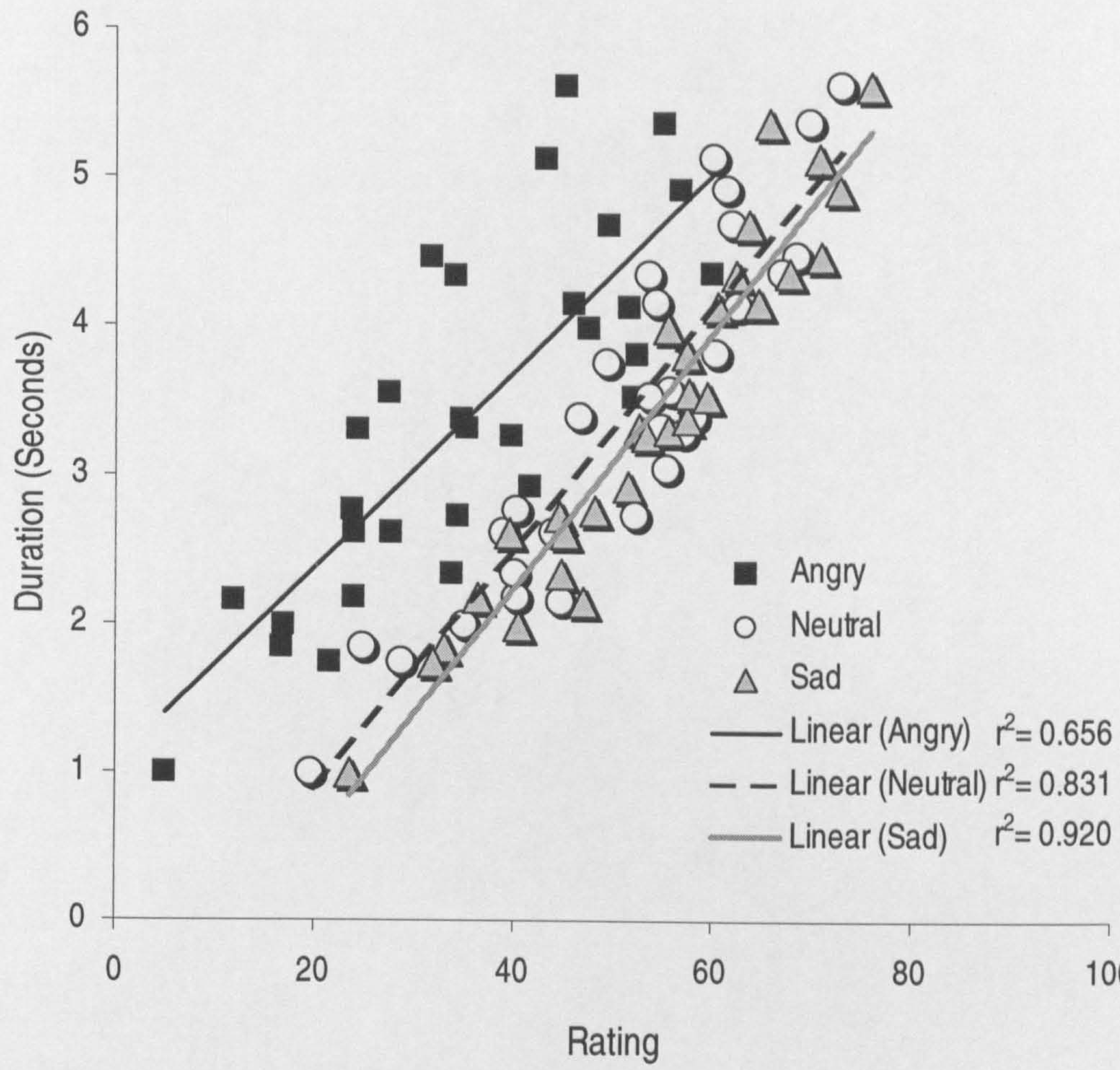

Figure 4.13 Correlation between ratings data and movement duration. For each affect the linear regression is shown with unadjusted r^2.

Regression analysis was done for each of the emotions that had been represented in the movements. To do this speed scores were again calculated from Principal Components Analysis (PCA) for use in the regression. Results for *angry* emotion showed that for the PCA one factor accounted for most of the variance in the kinematics values (84.41%).

The resultant factor scores accounted for 51.4% of the variance in the intensity ratings by observers of *angry* movements and time morphs [r^2=.514 (adjusted), F(28,1)=31.995, p<.001]. For *neutral* movements a single factor accounted for 84.64% of the variance in kinematics markers. The resultant speed scores accounted for 75.2% of the variance in intensity ratings [r^2=.752 (adjusted), F(28,1)=89.147, p<.001]. For *sad* movements the results were that the PCA of the kinematics markers resulted in a single factor that accounted for 84.93% of the variance. The factor scores from this analysis accounted for 82.3% of the variance in the intensity ratings

[r^2=.823 (adjusted), $F(28,1)=136.041$, $p<.001$]. In each case kinematics markers made a significant contribution to the regression model [*sad*: β (standardized coefficient)=-.911, t=-11.664, p<.001; *neutral*: β (standardized coefficient)=-.872, t=-9.442, p<.001; *angry*: β (standardized coefficient)=-.728, t=-5.627, p<.001].

These results from the regression analysis showed a relationship between kinematics and judgements of emotion. This relationship was very strong for *sad* and *neutral* movements, but weak for *angry* movements. A reason for the above result may be that the velocity of *angry* movements was not perfectly modulated by changing duration. From this it can be concluded that temporal characteristics seemed to largely describe *sad* and *neutral* movements. *Angry* movements, however, can only be partly described from a temporal perspective. It is however still the case that perceived intensity of emotion was modulated through a modulation of temporal characteristics.

General discussion

Overall the picture gleaned from Experiment 3 indicated that the perception of emotions *could* be modulated through changing the movement kinematics in the data. Interestingly, the intensity of the *angry* emotion could be manipulated, but it was only occasionally the case that changes in duration was enough to change the categorisation of these movements from "angry" to "neutral" or "sad". Conversely, *sad* and *neutral* movements were relatively compliant to modulation though changes of movement duration. It seemed that categorisation of these movements as *angry*, *sad* or *neutral* emotion depended closely on their temporal characteristics.

Why were angry *movements more resistant to modulation?*

Firstly it may be possible that humans have a bias towards detecting *angry* movements. This bias may result from a low threshold for detection of threatening movements. In this case even weak signals with the right speed or spatial configuration would lead to interpreting a movement as *angry*. Secondly, and leading on from the last point, humans may be more attuned to *angry* movements. This would certainly fit into an evolutionary perspective where a "safe rather than sorry" perspective could be of great value as a survival strategy.

Expounding this argument, humans would have either an innate ability or early in life learn a simple heuristic that speed means angry/threatening. Additionally, this sensitivity to angry/threatening movements may extend to sensitivity for other cues, such spatial or phasic relations among the points.

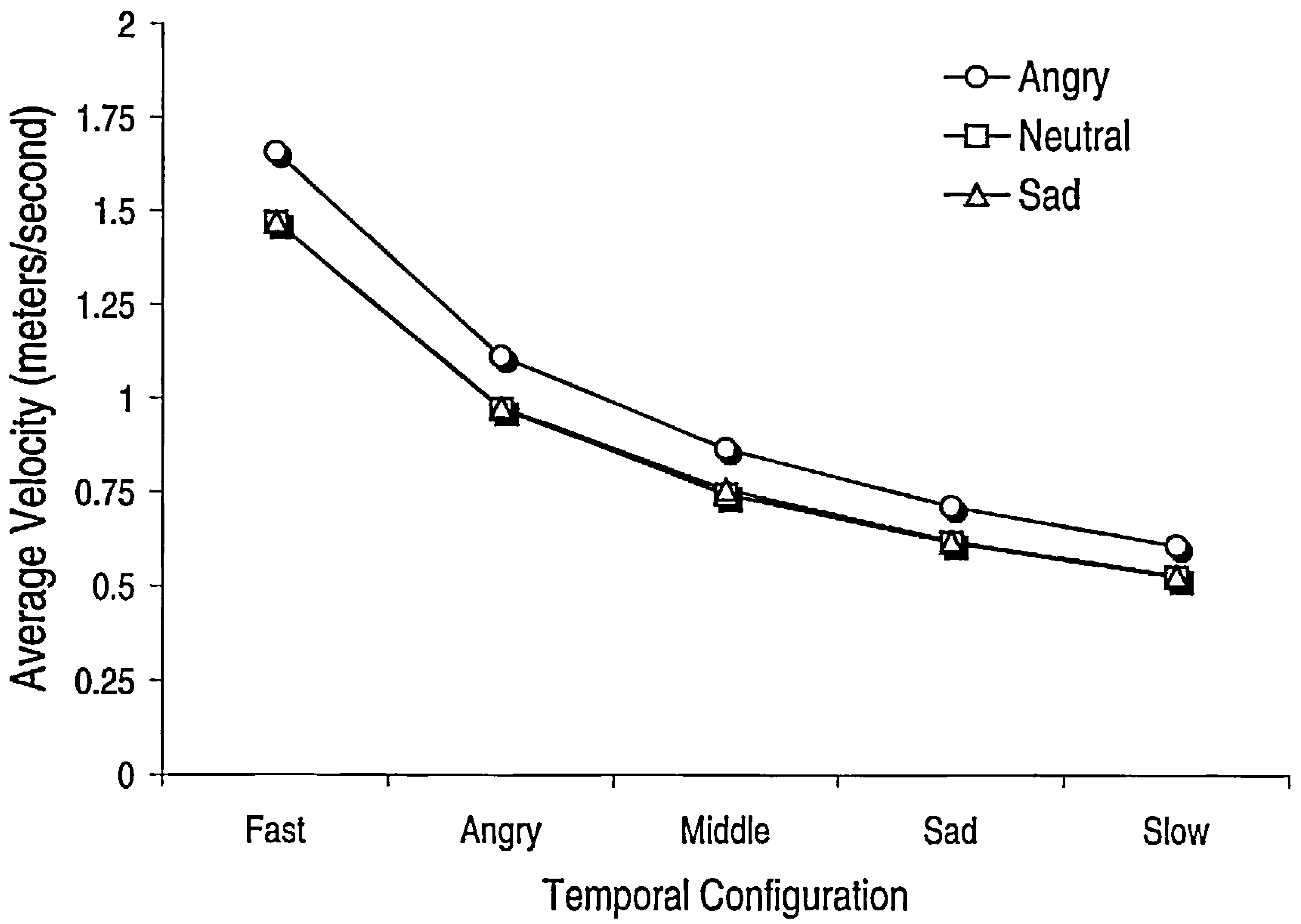

Figure 4.14 Average Velocity for Angry Neutral and Sad Movements

These latter may be weaker than the speed cue, hence the misinterpretation of *neutral* and *sad* movements as *angry* from only speed manipulations. However, in the absence of the stronger speed cue (such as our slowed down *angry* movements), the spatial cues would still be enough to encourage recognition as *angry* movements (hence the consistent correct identification of *angry* movements).

A different, related explanation may be that although the duration of movements was changed, velocity was not as dramatically affected. In comparing velocities of *angry* movements with those from *sad* and *neutral*, it is clear that *angry* movements were always faster than the other two (see Figure 4.14 above).

Faster velocity for *angry* movements was most probably due to the spatial component of velocity. This aspect of the movements was purposefully not manipulated during the morphing procedure, however the natural differences in the movements were not uniform. A visual examination of the displays is enough to indicate that an actor's hand moves further during the knocking phase of an *angry* movement than it does during the same phase of *sad* and *neutral* movements (see the demo on http://staff.psy.gla.ac.uk/ ~helena/demos/Actor1.mov).

Our findings regarding the modulation of perception through temporal features further suggest that there is a fairly direct mapping between perception and the basic physical properties of movements.

The Role of Other Properties in Perception of Emotion

In Experiment 3 the speeds of affective human arm movements were manipulated by changing movement duration. When new and original *angry* movements were viewed, they generally retained their identity as *angry* emotion, but the intensity of perceived emotion was modulated by velocity changes. *Sad* and *neutral* movements were similarly affected by changes in velocity, but faster movements were categorised as *angry*. These results further emphasised the role of velocity in perception of emotion. However, it was clear that there are other properties of the movements that were not controlled by velocity, but that play a role in the discrimination of emotion.

Our findings about speed are partly supported by the research of Kamachi et al. (2001) who showed that for facial expressions, different durations for morphed image sequences affected the perception of emotion differently. For angry movements, the best results were obtained when they were played at a medium duration (.87 seconds) and for sad facial expressions, slower duration (3.37) enhanced recognition. Additionally, for happy and surprised expressions, the most appropriate duration was fast (.2 seconds).

Since a completed sequence had to take place for each stimulus, the stimuli with fast duration were made up of fewer frames and thus can be though to have taken place at greater speed. Similarly for longer duration more frames were included, resulting in a slower change of expression over time. The duration of expression change used by Kamachi et al. (2001) may, however, not have reflected naturalistic duration of expression change.

Pollick et al. (submitted), however, found that for point-light displays of facial expressions, speed alone did not enhance recognition of emotion. For their stimuli they had captured natural human expressions, changing over time. They found that the speed for changes of expression was not diagnostic for categorisation of expression, instead spatial exaggeration of facial expressions enhanced recognition, suggesting that configural processing of facial expressions take place. Such configural (or holistic) processing of faces is usual in identity recognition from faces (Maurer, Le Grand, & Mondloch, 2002), and there is also evidence from static faces for configural processing of facial expression (White, 1999).

In biological motion perception, there is evidence for configural processing of stimuli, since upside-down point-light displays are recognised with lower accuracy than the upright version (Sumi, 1984). It is, however, unclear whether the disruptive effect of inverted point-light displays is due to the new and unusual dynamic configuration, or to more spatial factors. This is because the motion and structure from motion information are largely confounded (O'Toole, Roark, & Abdi, 2002).

The above findings regarding configural processing of biological motion and role of spatial features in facial expression recognition along with our findings from Experiment 3, suggests that it would be reasonable to more closely examine the role played by spatial characteristics in the perception of emotion from arm movements.

Our findings so far in Experiments 1-3 also suggest that there was direct perception of emotion from motion, particularly regarding the relationship between movement speed and $D1_{Canonical}$. It therefore stands to reason that $D2_{Canonical}$ would also represent a feature in the movements. This feature may well turn out to be some spatial characteristic of the movements.

Summary and Conclusions

In Experiment 2 it became clear that speed plays an important function in the perception of emotion from movement because of it's role as a perceptual indicator of physiological arousal. This finding has been exploited in Experiment 3 to modulate judgements about movements with emotion through a modulation of speed. Movements were morphed in the temporal domain with the purpose of modulating observers' judgement of emotion to the point where they would confuse movements that were usually reliably perceived as different.

Since *angry* and *sad* movements were very different with respect to kinematics markers of speed in Experiment 2, and suffered little from confusion in the categorisation task for Experiment 1, they were ideal candidates for this research. As a baseline, *neutral* movements, for which kinematics lay between *angry* and *sad* movements, were also used in the judgement task.

Categorisation and ratings of emotion intensity were collected from observers. The analysis of these showed that modulating the temporal configuration of movements could modulate perception of emotion. However, *angry* movements were recognised as *angry* in most cases despite changing the temporal configuration. This suggested that spatial characteristics of *angry* movements were a strong enough signal or recognition.

In the next chapter possible spatial and form characteristics of the movements will be explored as features that would inform the second dimension along which emotions were categorised.

Chapter 5 - More about Dimension 2

Introduction

In Experiment 1 we showed that when categorising human movements with emotions, observers used two major dimensions along which to define emotion. The most important of these was the speed of movements. This dimension ($D1_{Canonical}$) corresponded to a dimension of arousal that has been observed in a model of experienced emotion (Russell, 1980) and face perception (Schlosberg, 1954). The other major dimension along which observers categorised the movements ($D2_{Canonical}$) appeared to be similar to the dimension of valence in these models. $D2_{Canonical}$ did not correspond as well to the valence dimension as $D1_{Canonical}$ had corresponded to the dimension of arousal. The correspondence was, however, good when considering that the task in Experiment 1 was to match movements with expressions to category labels instead of a semantic definition of experienced emotions as in Russell's tasks (Russell, 1980) or perception of still photographs (Schlosberg, 1954).

Additionally we showed in Experiment 2 that a small rotation could maximise the correlation between kinematics of movement speeds and $D1_{Canonical}$. This rotation also caused negative emotions (*sad*, *weak*, *afraid*, *angry* and *excited*) to be grouped at the opposite end on $D2_{Canonical}$ from more positive and neutral emotions (*tired*, *happy*, *strong*, *relaxed* and *neutral*). The emotions for which the configuration was the worst were *weak*, *sad* and *tired*. The reason for this was related to the speed of movements. Speed was a strong indicator of emotion and may have been used exclusively in categorising the emotions *weak*, *sad* and *tired*.

Support for this hypothesis, with regard to *neutral* and *sad* movements, was obtained in Experiment 3. In this experiment the duration of movements were manipulated to modulate categorisation of *angry*, *neutral* and *sad* movements. Such a simplistic modulation of speed for *sad* and *neutral* movements, resulted in excellent modulation of recognition. This was not the case for *angry* movements. In the psychological space of movement categorisation, *angry* emotion had a distinct location on both $D1_{Canonical}$ and $D2_{Canonical}$, indicating that both dimensions played a role in the categorisation of this emotion. Indeed, modulation of temporal movement characteristics in this case was not enough to perfectly modulate the perception of emotion. A reason for this may be that only the temporal aspect of speed was manipulated, while spatial features remained unchanged. Since a modulation of speed was not sufficient to perfectly modulate perception of all the emotions, it is likely that spatial features of movements might also play a role in the perception and cognition of emotion from motion. It is this feature of the movements that will be explored further in the current chapter.

In Experiments 2 and 3 we found that the kinematics movements were easy to measure for a quantitative comparison between different movements and for correlation or regression of stimulus characteristics with perceptual judgements. This is not the case for spatial features of the movements. One spatial feature that we were able to quantify was the total distance that the wrist moved during the course of an action. However, there are other spatial features of the movements that could also inform perception. Such features include the phase or timing relationships between points, the spatial relationships between different points over time, the form in which the movements are displayed, posture of an actor and the more cognitive considerations such as inferences about who or what performed a movement. These possibilities have been explored in the five experiments that are detailed in the following sections.

In Experiment 4, the phase relations between points were scrambled. The resultant displays were shown to observers with the task of categorising the emotion. In Experiment 5, parts of the displays were selectively deleted in 4 different conditions so that the head (which informs posture) was missing. Only the hand, and finally all but the hand, were displayed in a recognition task for comparison with a control condition. For Experiment 6, new movements were collected and used to drive a solid body humanoid model. Perception of this solid body model was compared to that of point-light displays. The aims of these three experiments were to explore the kinds of spatial features encapsulated in human movements and the way in which they could be used as perceptual indicators for recognition of emotion. In Experiment 7, an attempt was made to examine how prior beliefs about what or who was trying to express an emotion would influence perception of human emotion from scrambled human movements.

Experiment 4 – Phase Scrambled Displays of Biological Motion

Introduction

In Experiment 4 knocking movements from Experiment 1 were presented with the same task – to categorise the emotions according to 10 emotion labels. The critical difference in this case was that the relationship among individual points was distorted by displaying the movements upside down and with scrambled phase (Bertenthal & Pinto, 1994). This was done in order to disrupt the displays in such a way that form would be affected while speed remained unchanged. We examined the perception of emotion within the framework of a psychological space and related aspects of this psychological space to physical properties of the movements.

Methods

Participants

Sixteen Glasgow University student volunteers took part in the experiment. Participants were naïve to the purpose of the experiment and were paid for their participation. The task was difficult and two participants abandoned the experiment.

Stimuli

Stimuli were point-light displays of the knocking arm movements that were used in Experiment 1. The movements were presented as yellow dots against a black background on a graphics computer (Octane, SGI) and were shown from a right saggital viewpoint. Spatial configurations of movements were altered by displaying them upside down and randomly phase-shifting the 6 markers that made up each movement.

To phase-shift the markers, they were re-sampled so that each marker from the sequence started at a randomly different stage during the movement cycle from the other 5. For instance, the start point of the wrist in the phase scrambled movements, might be the 50th frame of the original movement, the 20th frame might be the start for the head movement. A random start-point was calculated for each marker in a movement and the x, y and z motion tracks were displayed from this point onwards with the beginning of the marker's motion being shown at the end of the display.

In Figure 5.1, the process of phase scrambling has been illustrated by using the wrist and hand markers of a single movement from Experiment 1 and the same points after the movement had been phase scrambled for the current experiment. The xyz tangential position for wrist and metacarpal joints have been plotted against the corresponding frame numbers to show a position index for each marker in each frame of the movement.

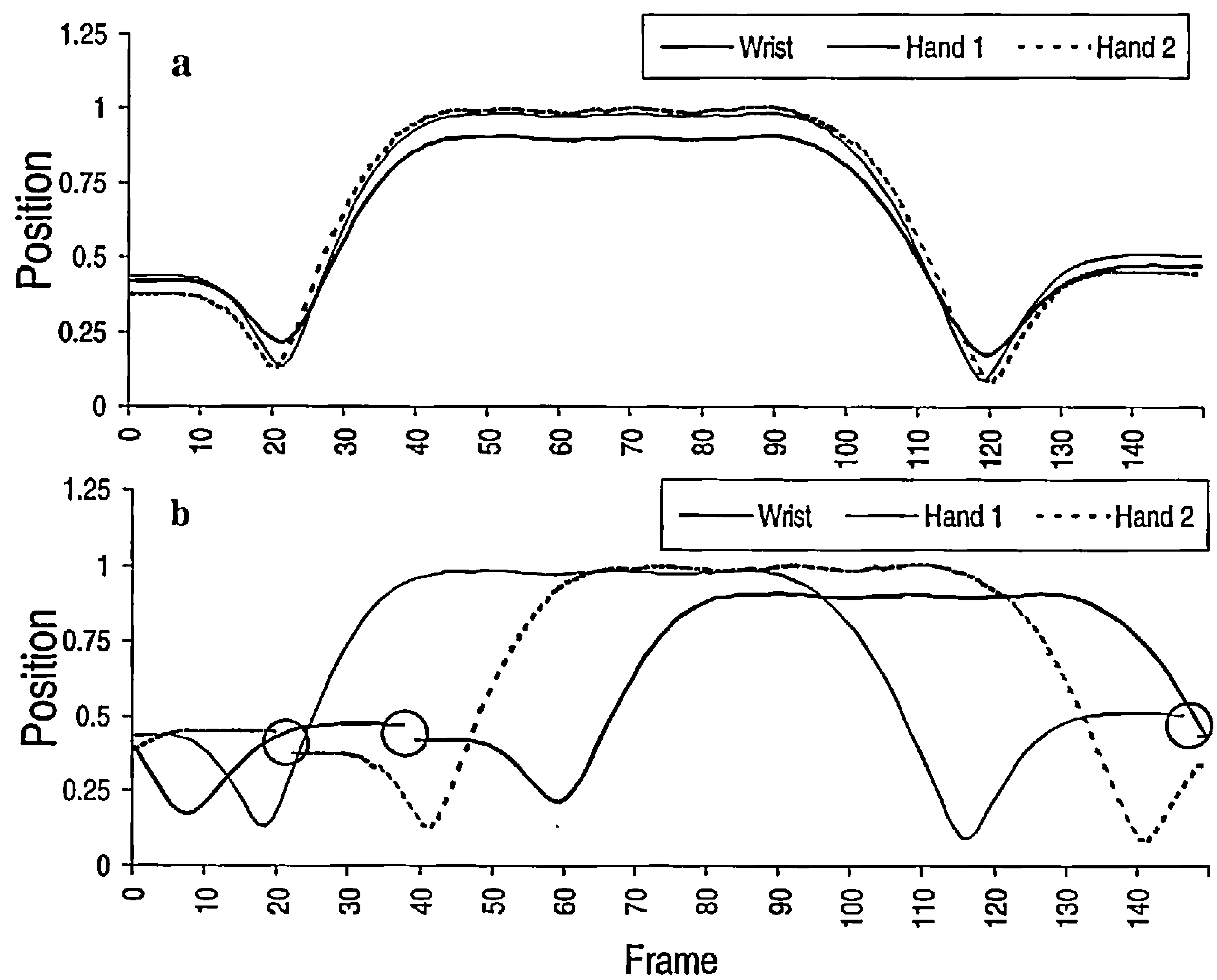

Figure 5.1 Tangential position of x, y and z co-ordinates for point on the wrist and hand for a) the canonical presentations of biological motion and b) phase scrambled displays of biological motion. Semi-transparent grey circles have been drawn on b to indicate the places at which phase shifts occurred for each of the point.

The movement illustrated here consisted of 150 different frames, and the places where a phase shift occurred have been denoted with a circle. In a movement that had not been phase shifted (Figure 5.1a) the wrist and hand markers all followed a very similar xyz position path over time. In a phase shifted movement (Figure 5.1b), however, the paths appear scrambled. The profile of each position path remained the same, but the

temporal sequence has radically changed. This manipulation of the data changed the global relationship among markers in the displays, but not the motion of individual points. This resulted in the kinematics remaining unchanged, but the form cues encapsulated in the movements being disrupted. Some examples of the phase scrambled movements can be seen on the following web site:
http://staff.psy.gla.ac.uk/~helena/demos/scrambled_actor1.mov.

The phase shift was calculated for all of the movements with the result that the 60 knocking movements each had a unique phase scrambled configuration.

A final manipulation of the displays was that the movements were shown up side down. This was done in order to prevent the recognition of the displays as originating from human arm movements, being a trivial task for observers.

In a pilot run of the experiment we discovered that observers could parse out the head and limb movement after a few presentations of the displays. This indicated that phase shifting the movements was only partially successful in causing a global disruption. The reason for this was that relatively stationary markers, such as the head and shoulder, were affected only marginally by scrambling the phase. These points offered stationary, stable elements that observers could perhaps use to recover spatial configuration. In order to maximise the spatial disruption; we presented the movements upside-down. In previous research, upside-down displays of human point-light walkers appear to be displays of strange human movements (Sumi, 1984).

Design

Stimuli were divided into 2 blocks according to the movements performed by the two different actors. For the purpose of counterbalancing, half of the 16 observers saw the movements performed by Actor1 first. In the two experimental blocks there were the 30 phase scrambled knocking movements, each repeated 3 times and presented randomly. There was also a practice block of two neutral movements. These neutral movements were randomly chosen - one from each experimental block.

Procedure

The experiment took place in a single session that lasted for about an hour. Participants were told that they would see some movement and that their task was to identify emotions from the displays. Observers were not told that the displays would be of human movement. This was done deliberately for two reasons. The first was that it would be very hard to persuade participants that these were human movements without an explanation of the experimental procedure – we wanted to avoid detailed explanation of the methods in order to prevent observers becoming wise to the purpose of the experiment. Secondly we thought that it would be interesting to discover what participants made of the scrambled stimuli.

Bertenthal and Pinto (Bertenthal & Pinto, 1994) did not report what observers thought their scrambled displays denoted, but in research by Proffitt, Bertenthal and Roberts (1984) there was a clue to what we may expect. Participants were presented with computer synthesised point-light walker displays and one of their tasks was to describe what they had seen. When displays were not recognised as a walker, they were described as balls bouncing off each other or lights attached to strings, oscillating backwards and forwards or rotating (from Proffitt, Bertenthal and Roberts, 1984 p320). In that experiment there was also a condition where the displays had been scrambled, but participants' descriptions were not reported.

At the start of the experiment a participant was presented with a set of written instructions that detailed the procedure of the experiment. In the instructions it was also emphasised to all participants that there would be *neutral* movements and that they should not use this term as a means of indicating that they could not recognise the emotion. The instructions were repeated verbally before the practice session.

In the first session of the experiment, observers were given a practice of 2 trials to familiarise them with the stimuli, equipment and procedure. Since the experimental design did not allow the observer to skip trials, it was stressed at this stage that a response should be made for each trial.

Two experimental blocks followed the practice session. For each trial the participant viewed the computer display of a movement and was then presented with a dialog box that contained the names of the ten possible emotions. Their task was to choose a single emotion by clicking the mouse on the matching label. After completing all trials in the experiment there was a short debriefing interview. An additional question was added to the standard debriefing interview to elicit from observers what they though the moving displays had represented while they were doing the task.

Results and Discussion

How did recognition of scrambled displays compare with canonical displays?

Over all trials of scrambled displays, participants correctly recognised emotions about 14% of the time and the range of performance was 4% (*afraid*) to 34% (*excited*). The average rate of recognition was significantly better than the chance value of 10% in a paired samples t-test [$t(13)=4.097$, $p<0.005$, two-tailed]. Compared with the canonical[21] knocking movements from Experiment 1 – for which the average rate of recognition was about 34% - performance with scrambled displays was significantly worse, as shown in an independent samples t-test for unequal sample sizes [$t(26)=9.885$, $p<.001$, one-tailed].

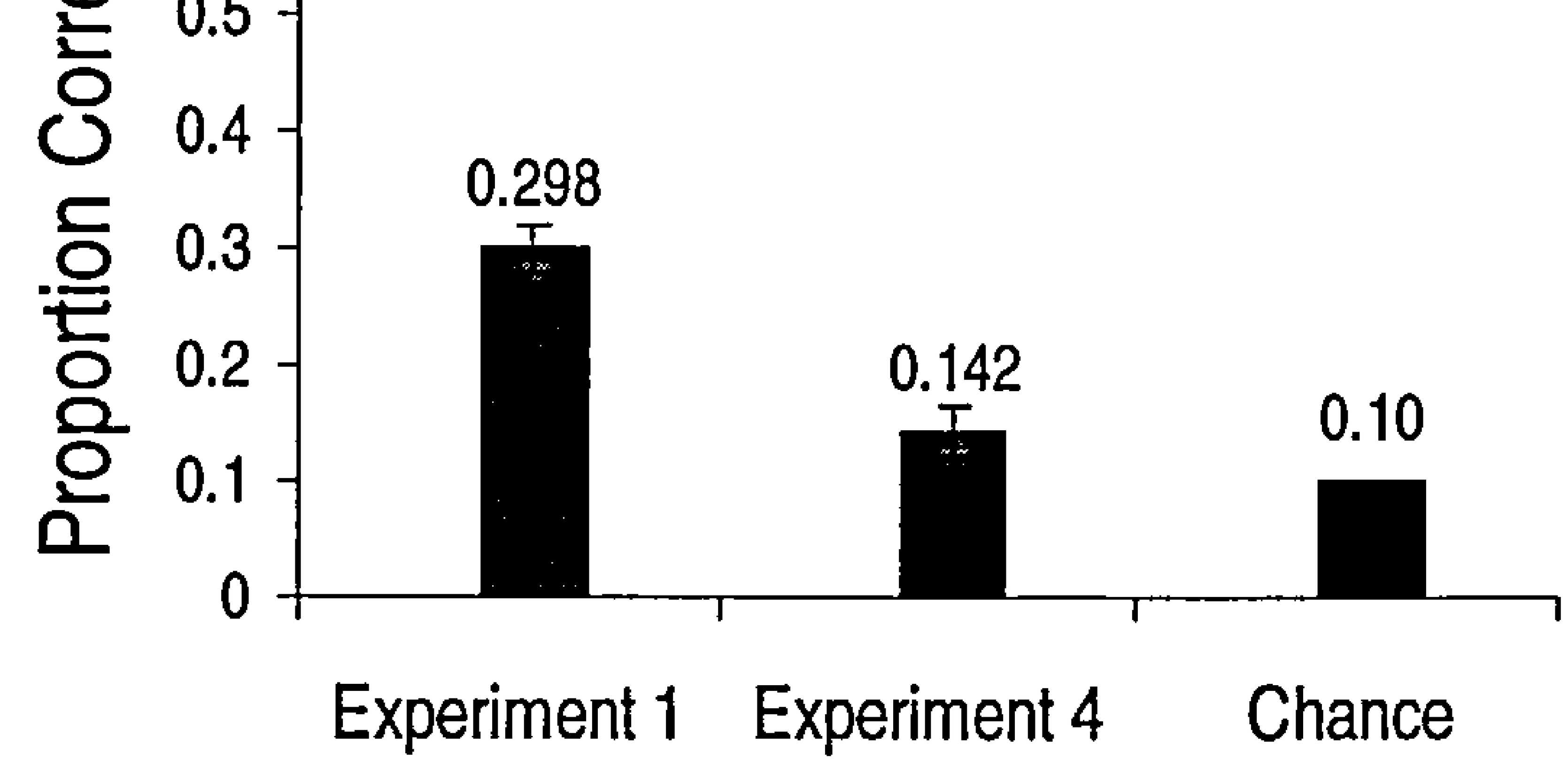

Figure 5.2 Proportion correct recognition of affects for Experiment 4 compared with the recognition task form Experiment 1 and chance levels of recognition. Standard Errors were very small so the error bars indicate the 95% confidence levels for each mean.

[21] Canonical displays were upright unscrambled arm movements as presented in Experiment 1 – in this case only knocking movements from Experiment 1 were used for comparison.

The mean of proportion correct recognition of emotions for canonical movements (Experiment 1) and scrambled movements (Experiment 4) and the level of chance has been graphed in Figure 5.2a along with standard mean errors to show the difference described above. These results echo those of Bertenthal and Pinto (1994) who also found that compared with canonical (unscrambled) displays, performance on phase scrambled movements was significantly worse while still being above chance.

How do different interpretations of the movements contribute emotion recognition?

For the upside-down phase-scrambled displays, none of the participants reported that the stimuli looked like a knocking arm movement. However, approximately half of the observers reported seeing some form of human movement in the displays. The response for each of the 14 participants who completed the experiment is shown in the list below.

Moving dots
Animal Rearing †
Person Waving* †
Person digging in a mine* †
Dots moving
Parts of a body throwing balls*
Molecules moving
Someone's abstract emotions
A man doing a single action repeatedly* †
A person cutting wood with an axe* †
Moving dots – later a person moving them*
People in small groups talking*
A man doing some action – perhaps digging* †
Abstract emotional stimuli

One issue of note in the descriptions is that participants were not consistent in their answers, however there were 2 rough themes in their descriptions. The first theme was human movement compared to non-human movement. The descriptions that were coded as human movement has been marked with stars (*) next to them. The other type of distinction that could be made in the descriptions was between coherent and incoherent movement. Six of the participants reported seeing coherent movement consistently throughout the experiment and the descriptions coded in this way have been marked with a cross (†) next to them. The mean proportion correct recognition of emotion for human and non-human groups and for coherent and incoherent groups has been graphed in Figure 5.3 along with standard errors.

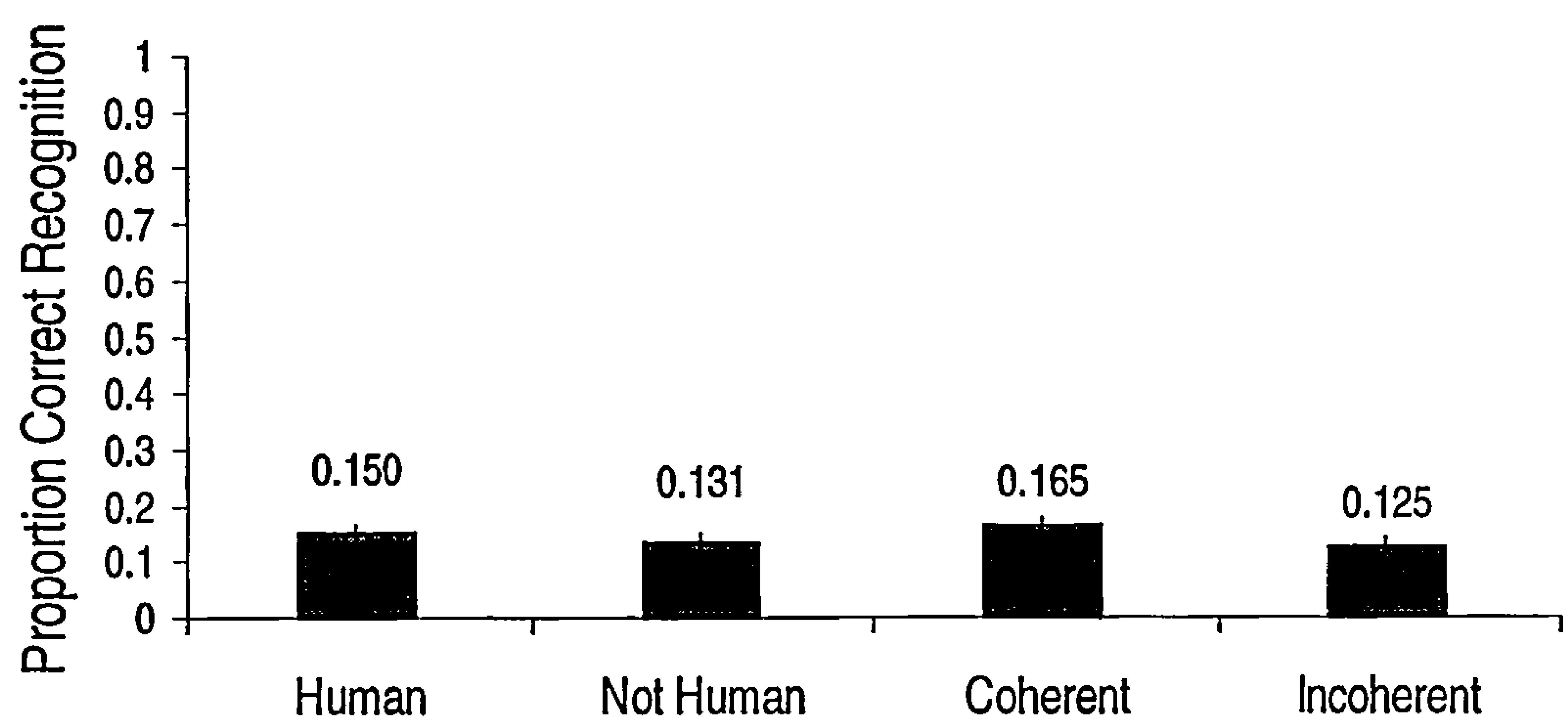

Figure 5.3 Proportion correct recognition of affects for the subgroups of observers that interpreted the scrambled displays to show human movement (n =8) compared with non-human movement (n=6) and coherent movements (n=6) compared with incoherent movement (n=8). Error bars indicate the standard errors for each mean.

The proportion of correct recognition for participants that interpreted the movements to show coherent movement was compared to that of participants who saw incoherent movements. Participants who thought the movements were coherent were able to recognise different emotions with 16.5% accuracy compared with participants who thought the dots were not related to each other and were able to recognise movements 12.5% of the time. This difference was significant as tested in a between groups t-test with unequal samples [$t(12)=2.178$, $p<.05$, one-tailed]. This result indicated that observers were better at judging the emotion that was depicted when they were able to resolve the displays into a single entity rather than unrelated elements or entities.

The difference in the proportion correct responses for observers that had indicated they interpreted the displays as showing human movements (15%) was compared with that of participants who did not interpret the displays to show human movement (13%). No significant difference between the two groups of participants was found in an independent samples t-test for groups with unequal samples [$t(12)=0.576$, ns, one-tailed]. This indicated that whatever the interpretation a person might have, regarding what they are looking at, when called upon to make judgements about the emotion that is being portrayed they found it equally easy or hard to do the task.

The finding that emotions can be recognised equally well from both humans and non-humans is rather interesting since it would indicate that people have no preference in the objects to which they would attribute feelings. Of course this is a well-known phenomenon amongst animators who regularly imbue their animal characters - such as Mickey mouse or Roger Rabbit - with feelings and personality. Heider and Simmel (1944) also found that feelings and intentions were attributed to abstract objects such as circles and triangles while recent animators have had similar success with desk lamps in the Pixar animation Luxo Jnr. (Lassiter and Rafael, 1987). The findings regarding recognition of emotion from non-human animate objects are, however, limited since they were incidental to the main purpose of the experiment. For this reason these issues have been further examined and are reported later in this chapter in Experiment 7.

Correct Recognition of Individual Emotions

In the current experiment the form of movements was disturbed in order to examine the effect on perception of emotion from movements. It has already been shown that correct categorisation of movements were impeded, however, it is not yet clear from the results whether the difficulty was equal for all emotions. This is the subject of the following discussion in which the proportion of correct identification of each emotion has been compared across experimental conditions with reference to the psychological space from Experiment 1.

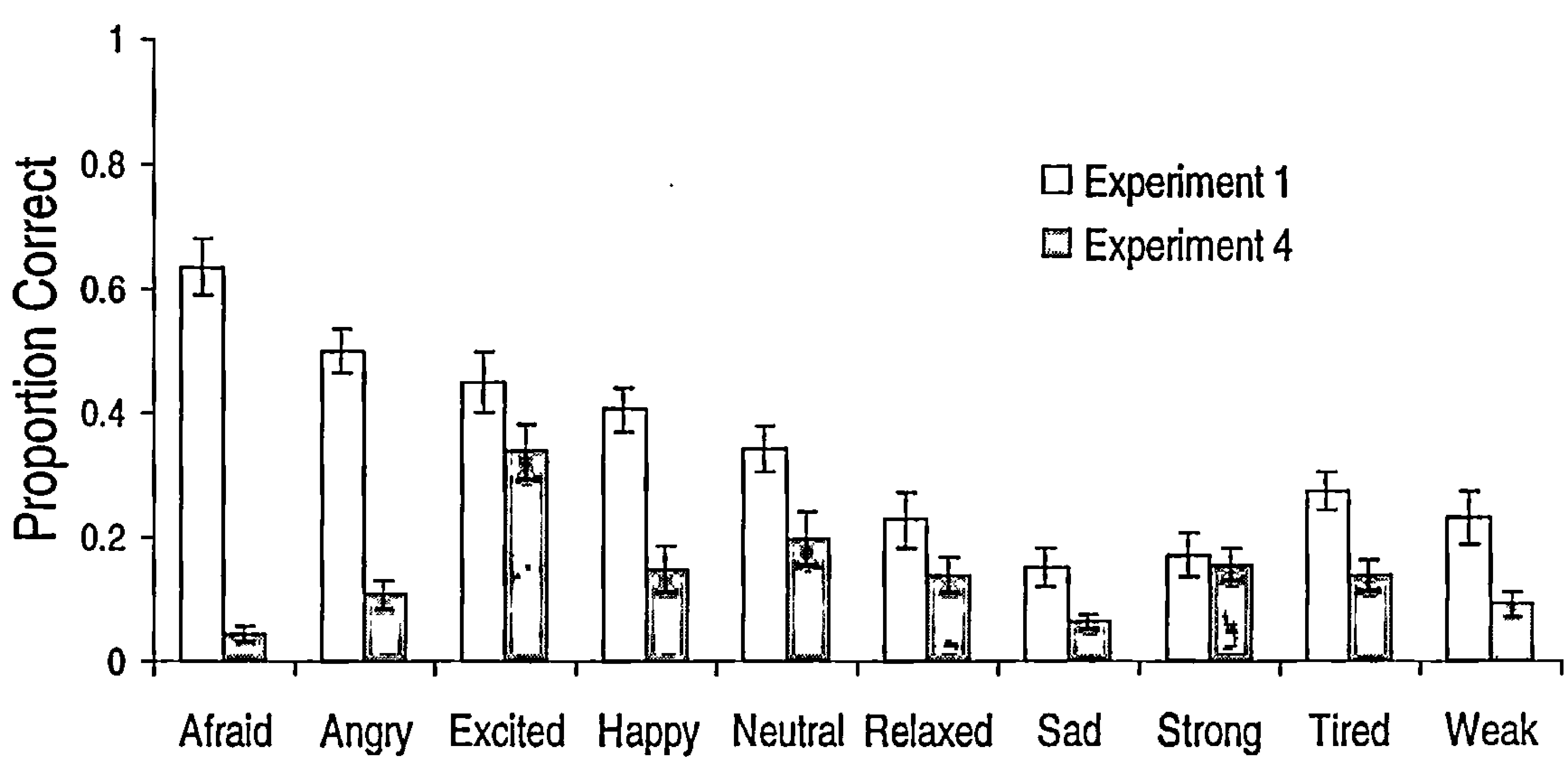

Figure 5.4 Proportion of correct recognition for different affects from knocking movements presented in normal fashion in Experiment 1 and phase scrambled in Experiment 4. Error bars depict standard errors of means according to observers (n=14).

The proportion of correct recognition of emotion has been graphed in Figure 5.4 to show the differences between Experiments 1 and 4. The most striking differences between canonical and phase scrambled movements was for the emotions *afraid*, *angry* and *happy*. Other differences were less pronounced or were non-existent, as in the case of *strong* movements. In order to test these differences, a 2 x 10 ANOVA with mixed design was calculated with the proportion correct results for canonical and phase scrambled knocking movements entered as a between groups variable with 2 levels. The 10 emotions were entered as the levels for a repeated measures variable.

The results showed a significant main effect for the between groups variable with the overall proportion of correct recognition for emotion from phase scrambled movements being much lower than for the same movements presented canonically [$F(1,26)=97.711$, $p<.001$]. This confirmed the findings in the t-test that was performed earlier. There was also a main effect of emotion [$F(9, 26)=10.228$, $p<.001$] which was again not surprising considering that it seems standard to find that there are individual differences for emotions in the recognition accuracy of emotion from movement. Of much greater interest was a significant interaction between two factors in the ANOVA [$F(9,26)=9.424$, $p<.001$].

Simple main effects of group were found for *afraid*, *angry* and *happy* at $p<.001$ and for *neutral*, *tired* and *weak* at $p<.05$. All were recognised at lower levels of accuracy for scrambled movements. There were few significant differences in the proportion of correct responses between different emotions. The only exception was for *Excited* movements that were significantly better recognised than most other emotions at $p<.05$, this excluded only *neutral* and *strong* movements.

In summary, an analysis of the proportion correct recognition has shown that scrambled displays were not as well recognised as canonical displays. We were also interested in the confusion between emotions and presentation format and this issue will be discussed further on pp. 133.

Psychological Space of Phase Scrambled Movements

As in Experiment 1, a confusion matrix of each actor was calculated and transformed into distances between the emotions. This was input to the INDSCAL multidimensional scaling algorithm to construct a Euclidean Distance Model of the differences and similarities between emotions. The resultant MDS solution was 2-dimensional and had $r^2 = .77$ and stress = .19. The first dimension accounted for 49.5% of the variance and the second dimension accounted for 27.3% of the variance. The psychological space for phase scrambled movements is the $RS_{Scrambled}$ and has been replicated in Figure 5.5a.

For comparison with the results of Experiment 4, a new psychological space was constructed from the subset of knocking movements that has been presented to participants in Experiment 1. This is the $RS_{Canonical}$ and is shown see Figure 5.5b. It had stress=.136 and r^2=.891 and apart from an anti-clockwise rotation of about 45°, the configuration was very similar to that of the combined knocking and drinking movements from Experiment 1 (see pp. 54).

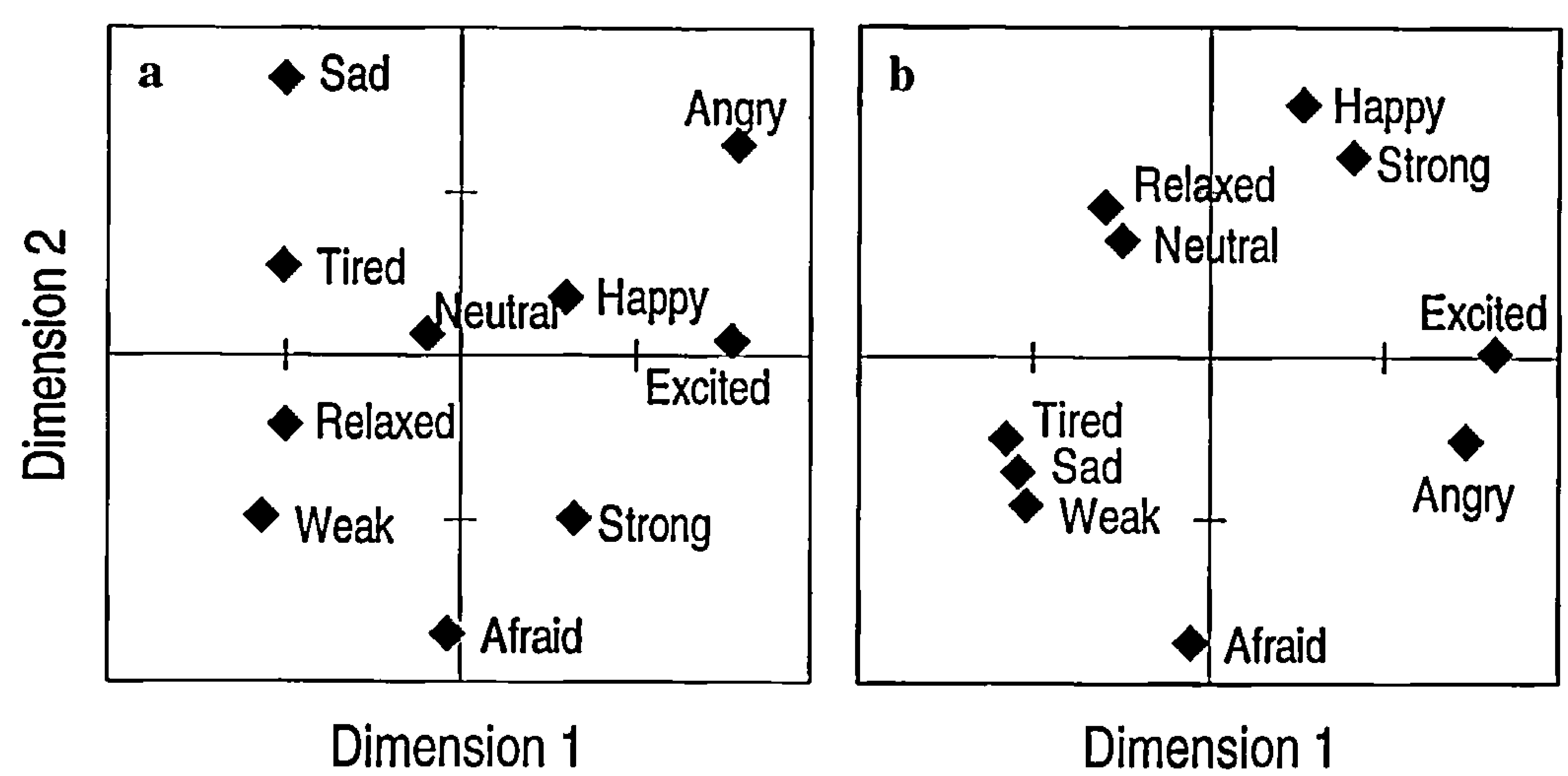

Figure 5.5 Psychological spaces for **a)** phase scrambles ($RS_{Scrambled}$) and **b)** canonically ($RS_{Canonical}$) presented knocking movements. Both psychological spaces have been plotted on the same axis for accurate comparison.

Comparison of the psychological space for phase-scrambled movements and canonically presented movements show a similar structure for Dimension 1, and a scrambled structure for Dimension 2.

This was confirmed by correlation between the dimensions of the two psychological spaces. The correlation between dimension 1 from both spaces yielded a r^2=.81 showing the close relationship between the two psychological spaces for this dimension. In Comparison the correlation between Dimension 2 for the two spaces yielded r^2=.01. From these results it is possible to conclude that when the form of movements were disrupted, there was a dramatic effect on the perception of emotion. This effect, however, was localised in the psychological space to only one of the dimensions along which observers based their judgements. In Experiment 1, $D2_{Canonical}$ conformed roughly to the pleasantness axis of the circumplex model while no such structure was apparent in Experiment 1.

The above, along with the results from Experiment 3, suggests that that the two dimensions along which human observers make their decisions about emotions, function more or less independently. However, it is still the case that when both spatial and temporal characteristics were presented without interference, the categorisations or emotions are much more accurate (according the feeling an actor was trying to convey). Additionally, the speeds of movements are most important for the accurate perception of emotion from motion.

Correct Recognition and the Psychological Space

We were interest the interaction between emotion and presentation of the movements (as canonical or scrambled) since we expected different confusions between pairs of emotions for the two conditions, according to their relative location on the psychological space for the perception of affect from emotion.

The psychological space from Experiment 1 gave both an indication of the relationship between different emotions in representation and the relative importance of the different dimensions for each of the emotions. In other words the diagnostic difference between pairs of emotions were indicated on the psychological space. Movements that were far apart on both dimensions were diagnostically different on both dimensions (such as *strong* and *weak*) and hence were confused only a little with each other. In comparison, movements that were similar along one or both dimensions were confused with each other more often. The relationships between emotions changed with the phase scrambling.

The aim of the following analysis was to confirm that the confusions between emotions could be predicted from the relationships between emotions on the $RS_{Canonical}$.

From the $RS_{Canonical}$ the relationships between different pairs of emotions can be characterised in 4 ways and we examined the confusions between emotions according to these categories. The four categories of emotion pairs were based on the way in which movements were confused with each other along $D1_{Canonical}$ and $D2_{Canonical}$ in the psychological space for knocking movements from Experiment 1 ($RS_{Canonical}$). In the first category were movements that had not been confused along either dimension. Examples of such pairs included *sad, tired* or *weak* compared with *happy* or *strong*. A characteristic of these pairs in the $RS_{Canonical}$ was that they were located far from each other on both the dimensions.

The second category of movements were those that were often confused for each other and included pairs such as *happy* and *strong*, or *relaxed* and *neutral*. They were typically close to each other on the $RS_{Canonical}$ and were very similar on both $D1_{Canonical}$ and $D2_{Canonical}$. The third and fourth categories of affect pairs were differently confused on $D1_{Canonical}$ and $D2_{Canonical}$. In the third category, pairs of affects were very different along $D1_{Canonical}$, but similar along $D2_{Canonical}$. Such pairs included *sad, tired* or *weak* compared with *angry*. In the final category of movement pairs there was similarity along $D1_{Canonical}$ and difference along $D2_{Canonical}$ and examples being *afraid* compared with *relaxed* or *neutral*.

To quantify the pairs of motions that belonged in each category, we used equations from vector geometry based on the Pythagorean Theorem. A right triangle was made for each pair of emotions on the $RS_{Canonical}$ with $D2_{Canonical}$ as the side opposite θ and $D1_{Canonical}$ adjacent to it. For each pair of affects the Hypotenuse was calculated with the following formula:

$$H_{a+b} = \sqrt{D1_{a-b}^{\;2} + D2_{a-b}^{\;2}} \qquad \text{Equation 10}$$

Where *a* and *b* were two emotions being compared and *a* represented the value of greatest magnitude. $D1_{a\text{-}b}$ and $D2_{a\text{-}b}$ were defined as:

$$D1_{a\text{-}b} = D1_a - D1_b \qquad \text{Equation 11}$$

$$D2_{a\text{-}b} = D2_a - D2_b \qquad \text{Equation 12}$$

The hypotenuse gave an estimate of how far apart in the $RS_{Canonical}$ two affects were located. An arbitrary criterion was set on the size of H_{a+b} for including movements in Category 2 pairs. This criteria was 1.075 and was a quarter of the maximum distance two points could be from each other, based on the range of values in the $RS_{Canonical}$. This was a unitless range of numbers produced by the multi-dimensional scaling algorithm to depict the xy ($D1_{Canonical}$ and $D2_{Canonical}$) co-ordinates for the 10 affects. Six pairs of affects fell into this category and they are listed in Table 5.1 along with pairs of emotions in the other categories.

For pairs of emotions that had $H_{a+b} > 1.075$, θ was calculated with the following formula:

$$\theta = \tan^{-1} \frac{D1_{a\text{-}b}}{D2_{a\text{-}b}} \qquad \text{Equation 13}$$

The range of θ was split into 3 equal parts (θsmall, θmedium and θlarge) that were used to categorise the remainder of affect pairs.

Pairs of affects that had θmedium were movements that were different on both $D1_{Canonical}$ and $D2_{Canonical}$ – these pairings of Category 1. Pairs of affects that had θlarge were Category 3 pairings and were different along $D1_{Canonical}$ but similar along $D2_{Canonical}$. Finally, pairs of affects that had θsmall were similar along $D1_{Canonical}$ and different along $D2_{Canonical}$ and were Category 4 pairings. The pairs of emotions in each category have been shown in Table 5.1.

So for Category 1 pairs there were diagnostic differences in both temporal ($D1_{Canonical}$) and spatial ($D2_{Canonical}$) configurations between movements and this resulted in few confusions between the movements in Experiment 1. After scrambling the displays, we did not expect that these pairs of movements would suffer much additional confusion. This was because the movements' speeds should be a sufficient cue with which observers could tell them apart when spatial features became ambiguous. For Category 2 affect pairs there were no diagnostic differences. Instead both the temporal and spatial configurations were similar and movements were often confused for each other. We expected that they would be affected the most in the scrambled condition. This is because although the temporal configuration of movements would remain unchanged, their spatial configuration would now be much different and there would be less confusion between movements.

Table 5.1 Pairs of affects grouped into four types according to the dimensions along which they were confused in the psychological space for knocking movements in Experiment 1.

Type 1		Type 2		Type 3		Type 4	
Both Dimensions		Neither Dimension		Dimension 1		Dimension 2	
Afraid	Angry	Angry	Excited	Angry	Sad	Afraid	Happy
	Excited	Happy	Strong		Tired		Neutral
	Sad	Neutral	Relaxed		Weak		Relaxed
	Tired	Sad	Tired	Excited	Neutral		Strong
	Weak		Weak		Relaxed	Angry	Happy
Angry	Neutral	Tired	Weak		Sad		Strong
	Relaxed				Tired	Neutral	Weak
Excited	Happy				Weak		Sad
	Strong			Happy	Relaxed	Relaxed	Sad
Happy	Neutral			Neutral	Strong		Weak
	Tired			Relaxed	Strong		Tired
	Sad						
	Weak						
Neutral	Tired						
Sad	Strong						
Strong	Tired						
	Weak						

Pairs of movements in the Category 3 category were classed together as being more different from each other on $D1_{Canonical}$ than on $D2_{Canonical}$. The diagnostic differences between these movements would be based on temporal properties of the movements. Since scrambling the movements had no effect on temporal properties of the movements, we expected little if any change in the number of confusions between movements with these affects.

Category 4 affect pairs were classed together as being more different from each other on $D2_{Canonical}$ than on $D1_{Canonical}$. The diagnostic differences between these movements would be based on physical properties of the movements that could inform $D2_{Canonical}$. If such information were encapsulated in features based on form, scrambling the movements would have some effect on the confusions between pairs of movements. We could, however, not predict what this effect would be since the phase shift procedure was not applied systematically.

After re-coding the pairs of movements into the four sub-groups above, the average proportion of times that each pair of movements was confused was calculated for each participant. The means for the four sub-groups are plotted in Figure 5.6.

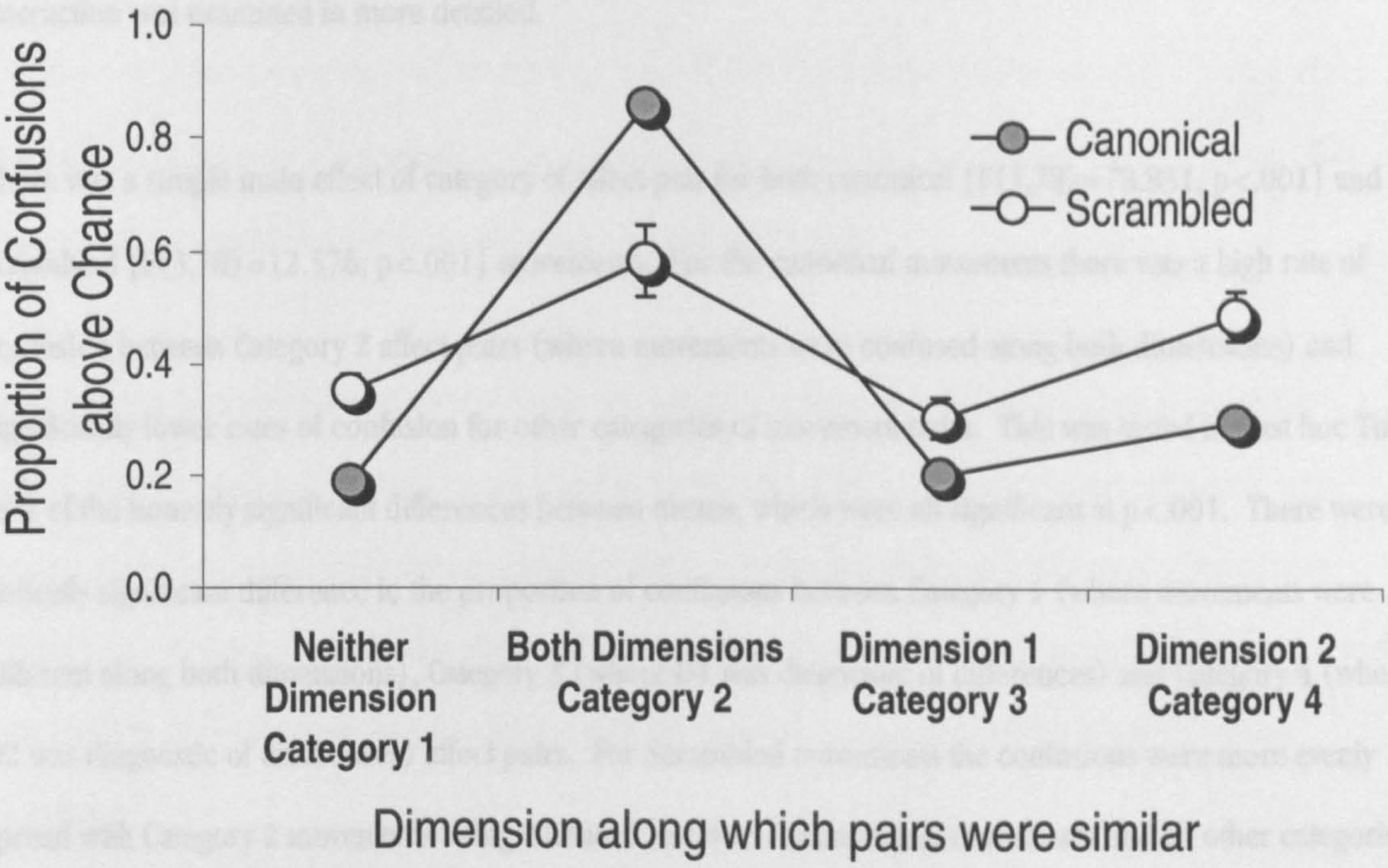

Figure 5.6 Proportion of confusions for different pairs of affective movements have been plotted for Experiment 1 (Canonical Displays) and Experiment 4 (Scrambled Displays). Means are based on confusions by 14 participants for pairs of affects that were rated as 1 similar along neither dimension (15 pairs), 2 similar along both dimensions (6 pairs), 3 similar along dimension 1 (11 pairs) and 4 similar along Dimension 2 (11 pairs).

To test whether the number of confusions changed from Experiment 1 to Experiment 4, we calculated a 2 (canonical vs. scrambled displays) x 4 (categories of emotion pairs) mixed design ANOVA with presentation type as a between groups variable and affect pair category as a repeated measures variable. The results showed a main effect for the four categories of affect pairs [$F(3,78)=73.194$, $p<.001$] and an interaction between category and type of presentation [$F(3,78)=18.213$, $p<.001$], but there was no main effect of presentation type [$F(1,26)=2.933$, $p=.099$]. This meant that there was no difference in the average proportion of confusions participants made between pairs of emotions when movements were presented in the usual way or when they were scrambled. A closer look at the means for canonical and scrambled movements did show that there were generally more confusions when presenting scrambled movements (43.2%) than when canonical movements (37.8%) were presented.

The main effect of category of affect pair was not very informative since the $RS_{Canonical}$ had been calculated from distance data that represented dissimilarity between movements in Experiment 1. Instead the confusions from Experiment 1 were used as a baseline against which to compare confusions in Experiment 4 and hence the interaction was examined in more detailed.

There was a simple main effect of category of affect pair for both canonical [$F(3,78)=78.831$, $p<.001$] and scrambled [$F(3,78)=12.576$, $p<.001$] movements. For the canonical movements there was a high rate of confusion between Category 2 affect pairs (where movements were confused along both dimensions) and significantly lower rates of confusion for other categories of movement pairs. This was tested in post hoc Tukey tests of the honestly significant differences between means, which were all significant at $p<.001$. There were no honestly significant difference in the proportion of confusions between Category 1 (where movements were different along both dimensions), Category 3 (where D1 was diagnostic of differences) and Category 4 (where D2 was diagnostic of differences) affect pairs. For Scrambled movements the confusions were more evenly spread with Category 2 movements being confused less than for canonical movements and all other categories being confused more. HSD's were found in a more complex pattern than before with category two pairs being confused significantly more than Category 1 and 3 at $p<.001$, but not for Category 4 pairs. Again there was no HSD between Category 1 and 3 pairs, but there were HSD's between both of these and Category 4 pairs at $p<.05$. These results reflected a lower rate of confusions for Category 2 pairs and a higher rate for Category 1,3 and 4, with the increase in Category 4 pairs being much more than for the affect pairs.

There were simple main effects of presentation type at all levels of affect pair categories and for Category 1, 2 and 4 affect pairs – for which spatial characteristics played a part in categorising movements – these effects were significant at $p<.001$. For Category 3 movements – in which the diagnostic difference had been in the temporal characteristics of canonical movements - the level of significance was lower at $p=.045$. A closer look at the means confirmed that although there was a significant increase in confusions for Category 3 affect pairs when scrambled movements were compared with canonical movements, this was a smaller increase than for all other categories of affect pair. The greatest effect was on affect pairs of Category 2, where movements were confused much less for each other after scrambling. Disrupting the spatial cues also caused category 1 and 4 pairs to be confused more for each other.

From this analysis we concluded that the difficulty in telling apart movements after scrambling the displays, was different for different categories of movements pairs. In the case where speed had represented a diagnostic difference in canonical movements, scrambling the spatial features of movements had only a limited effect. In cases where Dimension 2 had played a role, scrambling the movements had a more dramatic affect. This illustrates the independence of dimension 1 and 2 from each other.

By comparing the confusions observers made, it is possible to draw only limited conclusions since the scrambling process had been applied randomly to movements. This issue has not been addressed in the current thesis, but represents a possible future expansion of the research that has been conducted here. Another way of looking at confusion data is to translate all confusions into distances between affects for analysis with an MDS algorithm to make a psychological space that can be compared to that of the canonical movements.

Comparison of Psychological Space to Movement Kinematics

Up to this point, results have been reported in terms of a) the proportion of correct responses and b) the psychological spaces for perception of canonical and phase scrambled presentations of movements with emotion. In this section the results of a comparison of the movement kinematics to perception are reported.

The aim of the following analysis was to explore the relationship between physical properties in the movements and the psychological space in order to test further the observations regarding independence of Dimensions 1 and 2. This independence could be shown in two ways. Firstly in the relationship between the dimensions from the two psychological spaces. If dimension 1 from Experiment 1 corresponded with dimension1 from the current experiment and, but dimension 2 from the two experiments did not. There would be additional proof in the case where the relationship between the kinematics and dimension 1 from the psychological space for the phase scrambled movements were the same as that found in Experiment 1.

In order to establish the nature of the relationship between the psychological space of Experiment 4 and the kinematics markers, we correlated the kinematics markers measured from the movement data with the co-ordinates of the emotions from the psychological spaces.

We first took the instantaneous measures of the wrist kinematics (velocity, acceleration and jerk) and obtained the kinematics markers of duration, average velocity, peak velocity, peak acceleration, peak deceleration and jerk index. These kinematics markers were next correlated to the Dimension 1 and Dimension 2 co-ordinates of the 10 emotions in the psychological spaces.

From Figure 5.7 it can be seen that for both Experiments 1 and 4, Dimension 1 (activation) co-ordinates of emotions correlated well with the kinematics markers. This table shows that energetic motions were positively correlated with shorter duration and greater magnitudes of average velocity, peak velocity, acceleration, deceleration and jerk. For Dimension 2 in Experiment 1, it was found that to a lesser extent there was a tendency of longer duration and smaller magnitude of the other kinematics markers to be correlated with Dimension 2. For Experiment 4, in contrast, there was no correlation with the kinematics markers.

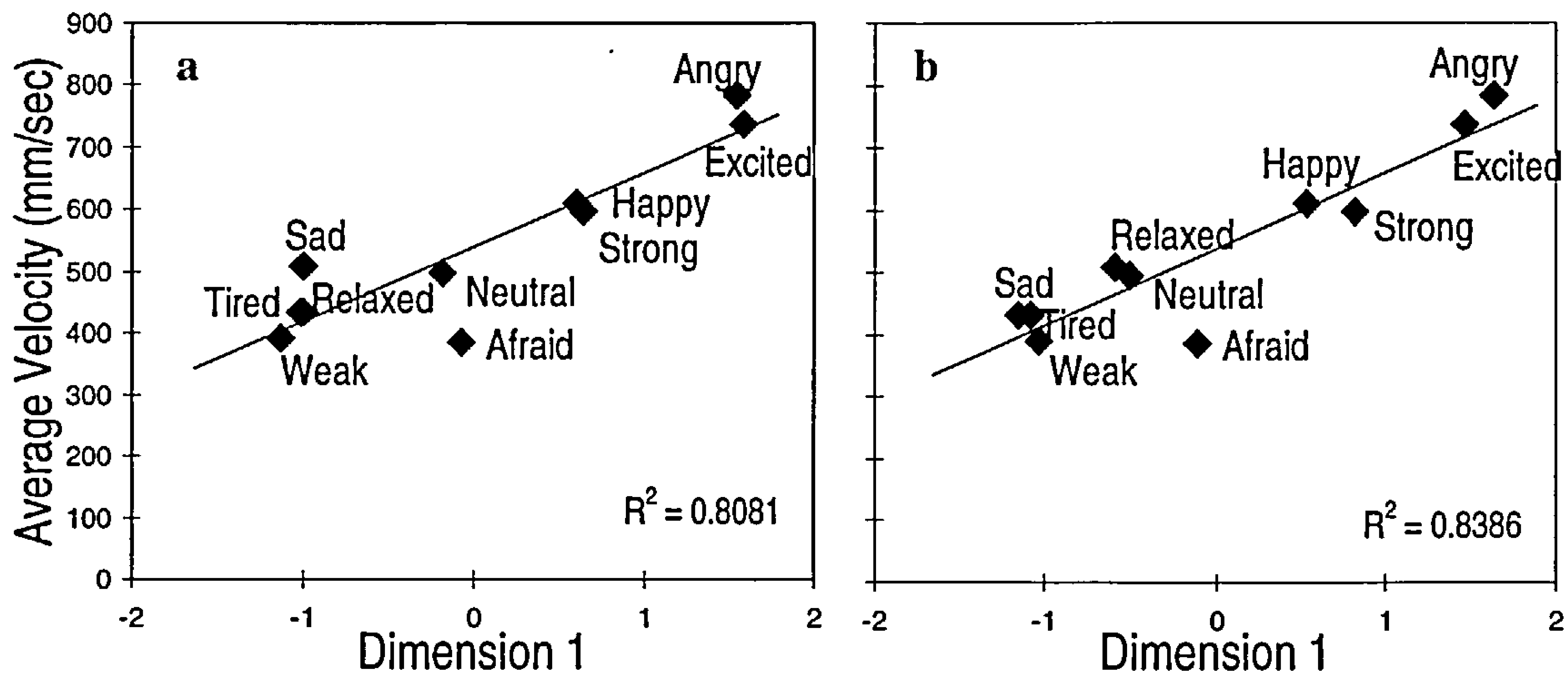

Figure 5.7 Correlation between average velocity and perception of movements with emotion presented as a) phase scrambled movements and b) in the normal way.

Table 5.2 Correlation between kinematics markers and dimension 1 from the Psychological Space. All values are Pearson's correlation coefficient.

Kinematics Property	Experiment 1		Experiment 4	
	$D1_{Canonical}$	$D2_{Canonical}$	$D1_{Scrambled}$	$D2_{Scrambled}$
Duration	-0.92	-0.37	-0.91	-0.27
Average Velocity	0.94	0.36	0.91	0.17
Peak Velocity	0.92	0.41	0.90	0.32
Peak Acceleration	0.94	0.18	0.93	0.27
Peak Deceleration	-0.88	-0.13	-0.88	-0.33
Jerk Index	0.94	0.19	0.93	0.27
Distance Moved	0.38	0.76	0.37	0.45

This was examined further by rotating the psychological space to find the orientation that maximised the r-squared values of the correlation with the kinematics markers. It was found that for Experiment 1 a 27° counter-clockwise rotation, and for Experiment 4 a 12° clockwise rotation, resulted in the highest correlation with the kinematics markers. These rotations resulted in average r^2values of 0.88 for Dimension 1 (compared with 0.73 for the un-rotated configuration) and 0.03 for Dimension 2 (compared with 0.35 for the original configuration) in Experiment 1. For the phase scrambled movements the average correlation improved for Dimension 1 to r^2=.86 (compared to 0.73 for the original configuration) and decreased to 0.01 for Dimension 2 (compared to 0.09). From these results it can be seen that while the original psychological space was roughly oriented so that energy in Dimension 1 is correlated with the speed of the movement, a relatively small rotation of the space can improve the correlation.

Summary

In general the results for this experiment, and the comparison with the results from Experiment 1 has shown that disrupting the form in a movement had an effect on the perception of emotion from movement. However, this disruption was limited, and the speed of movements, which had not been changed, still conveys enough information to allow for categorisation that is better than guessing.

Disrupting the phase destroyed form information. However, it is unclear at this stage which aspect of form informs perception of emotion. This is because both global and local spatial features were disrupted. Global features include the hierarchical relationships between joints and the form from motion information that encourages the perception of a human arm. Localised spatial features in the movements that were disrupted are the temporal sequences for the position of each point throughout the movement and the coherence of informative features such as the hand. Issues concerning local and global form features are examined further in Experiments 5 and 6.

Experiment 5 – Local Features of Movements with Emotion

Introduction

In Experiment 4 the movements were disrupted by phase scrambling the whole display. It became clear that complementary to categorisation of movements with emotion according to their speed, there is also form information that allows a finer grained categorisation. Thus the disruption of form information had a dramatic effect on the categorisation of movements. It was, however, not clear whether this was because local features in the movements were disrupted, or because the disruption was to global features of form in the movements. In Experiment 5 local features of the movement form were changed so that the effect on performance could be observed.

When phasic relationships changed there was a change on two levels: First a global disruption of the form and secondly numerous local disruptions in the relationship between joints. Natural phase shift is a feature of human motion (Amaya, Bruderlin and Calvert, 1996) and in computer animation is regarded as being an important consideration (Lasseter, 1987). Amaya et al (1996) also believed that this phase shift between points would be characteristically different for different affective movements. As phase changes the form of an object is revealed in the rigid relationships between moving joints and because of this there is a confounding factor in Experiment 4. This confound arises because when phase relationships are disrupted, cues about form are distorted as well. The question remains, is dimension two modulated by the subtle phase relations amongst points, the forms as described by these relations, the form regardless of these relationships or some other subtle interaction between form and phase relationships?

It seems almost impossible to manipulate form and phase separately, however, since phase relations suggest form. Deleting parts of the form in the fashion of individual point-lights should allow the phase relations to be maintained, but the form to be more flexible in an observer's imagination. In the following experiment, the technique of omitting local features in the movement from the display was used to make three new types of display. These displays were compared to a control condition in which movements were displayed with all features present.

Methods

Stimuli

Stimuli were the *afraid*, *angry*, *happy*, *neutral* and *sad* movements described in Experiment 1 and were both drinking and knocking actions. Along with the original movements - the control stimuli in this experiment – three new types of stimuli were constructed from the original movements. The new stimuli were made by changing the IVis inventor file to display only the required points from the 6 possible motion tracks. All stimuli were displayed as described in Chapter 2 and the new types of movements are describes below.

Movement type A – The Headless Man

On many occasions, when the displays from Experiment 1 were shown in presentations or experiments, both casual and experimental observers have commented that they looked at the posture of point-light people when making judgements of emotion. Posture can be derived from the configuration of the head and arm points. This information is not very accurate, since it is based on an inconsistent cue deriving from the recording sessions where it is impossible to place marker so that they are always in the same spatial location with reference to the cameras and each other. Instead the physiology of an actor was used to gauge marker locations. The consequence of this was that over different recording sessions the spatial relationship between the points depicting the head and shoulder changed, resulting in an artificial and inaccurate representation of body posture.

Nevertheless, the configurations of points were the same for all the movements of each actor, since markers were placed only once at the start of a recording session. This meant that if there were differences in position of the head for different emotions, these differences would constitute a valid cue for the categorisation of emotion within each recording session. A case in which posture may add to the perception of emotion would be when *sad* movements are compared to other movements. Actors may hang their heads to indicate that they are *sad* or hold their head up high to indicate confidence.

It is a relatively simple manipulation to miss out the point of light that depicts the head.

The movements of a headless man were made by instructing the display software to render the movement of only the five points depicting the joints on the arm and hand. Comparing results from this condition to those from the control condition would shed some light on the role a local cue such as the head position played in the perception of emotion from motion. In particular, if it were an important cue, there should be a cost in performance for movements in which it informed perception of emotion. In comparison if it were indeed an inaccurate cue, this bogus information would be removed from the displays and result in improved performance.

Movement type B – The Handy Man

Mather, Radford and West (1992) showed that the end-effectors of a movement (in their case the ankle and wrist for gait) were the most important element of the movement to the extent that disruption of these elements had the greatest effect on judgement tasks. In the case of affective arm movements, the wrist velocity has proven important in emotion recognition. The hand alone, however, can describe only local features in the form.

For this reason, type B movements described only the path travelled by the three points that depict the movement of the hand so that the elbow, shoulder and head were not shown. Seeing only the hand, observers would be able to base their judgements of emotion only on local form and speed cues. Again, this was achieved by displaying only the three relevant points, depicting the wrist and metacarpal joints, to observers. If it were the case that the hand was of primary importance in the perception of emotion from motion and the other points only contribute minimally, there would be little cost to observers' abilities in categorising the movements.

Movement type C – The Handless Man

In comparison to the handy man, when the local cue that primarily contributes towards assessment of speed (the hand) is not displayed, performance should be impaired. In this case form cues would also be attenuated, but postural information would be still be available. To show movements without a hand, all point were displayed apart from the one on the wrist and the two describing the metacarpal joints. It was hypothesised that the loss of the end-effector of movements would cause a deficit in the participants' ability to categorise movements.

Participants

Forty-eight volunteers were recruited from the student population at the University of Glasgow to take part in the experiment. They were paid for their participation and were naïve as to the purpose of the experiment.

Design

The experiment had a mixed design with movement conditions coded as a between groups factor with 4 levels. These 4 levels were made up of the different types of movement that were presented to participants: regular movements (control), headless movements (headless), handless movements (handless) and movements depicting just a hand (hand). Each participant saw trials for only one movement type in the experiment, which mean that there were 12 observers in each level of the between groups factor.

Within each level of the between groups factor, there was a 5 (emotions: *afraid*, *angry*, *happy*, *neutral*, *sad*) x 2 (actions: drinking, knocking) x 2 (actor: Actor1 and Actor 2) repeated measures sub-design.

The stimuli were arranged into 4 blocks according to all combinations of actor and emotion. Each block contained the movements of a single actor depicting either drinking or knocking actions and hence there were 15 stimuli. Each participant saw a total of 60 movements and each of these were repeated 3 times. There was also a practice session in which the *neutral* movements from all the 4 different experimental blocks were shown to familiarise participants with the displays and procedure.

Across participants the order of experimental blocks were randomised and within each block of experiments the order of stimuli was random.

Procedure

The experiment took place in a single session that lasted for about an hour. On arrival in the lab, participants were randomly allocated to one of the 4 experimental conditions. They were then presented with a set of written instructions in which the task and procedure were described.

Following this there was a practice session in which participants were asked to simply categorise the movements they saw as either knocking or drinking movements. This was done to familiarise participants with the displays and procedure. The length of this practice session was adjusted online according to when a participant was correct in their categorisation for three times in a row. It was decided prior to the experiment that if a participant could not distinguish actions after a single presentation of all the *neutral* stimuli, they were to be disqualified from the experiment. This never occurred, and typically after only a few (about 5) presentations, the practice session was ended.

After the practice session the experimental blocks were presented. For each trial the participant viewed the computer display of a movement and was then presented with a dialog box that contained the names of the 5 possible emotions. Their task was to choose a single emotion to which the movements corresponded by clicking the mouse on the matching label. After completing all trials in the experiment there was a short debriefing interview.

Results and Discussion

The proportion of correct recognition for each condition in the experiment is shown in Figure 5.8. Observers that viewed displays in which the point on the head was missing showed the highest proportion of correct recognition. For the other three conditions, there was little difference in the rates of correct recognition, but the lowest recognition was evident for movements in which the hand was absent.

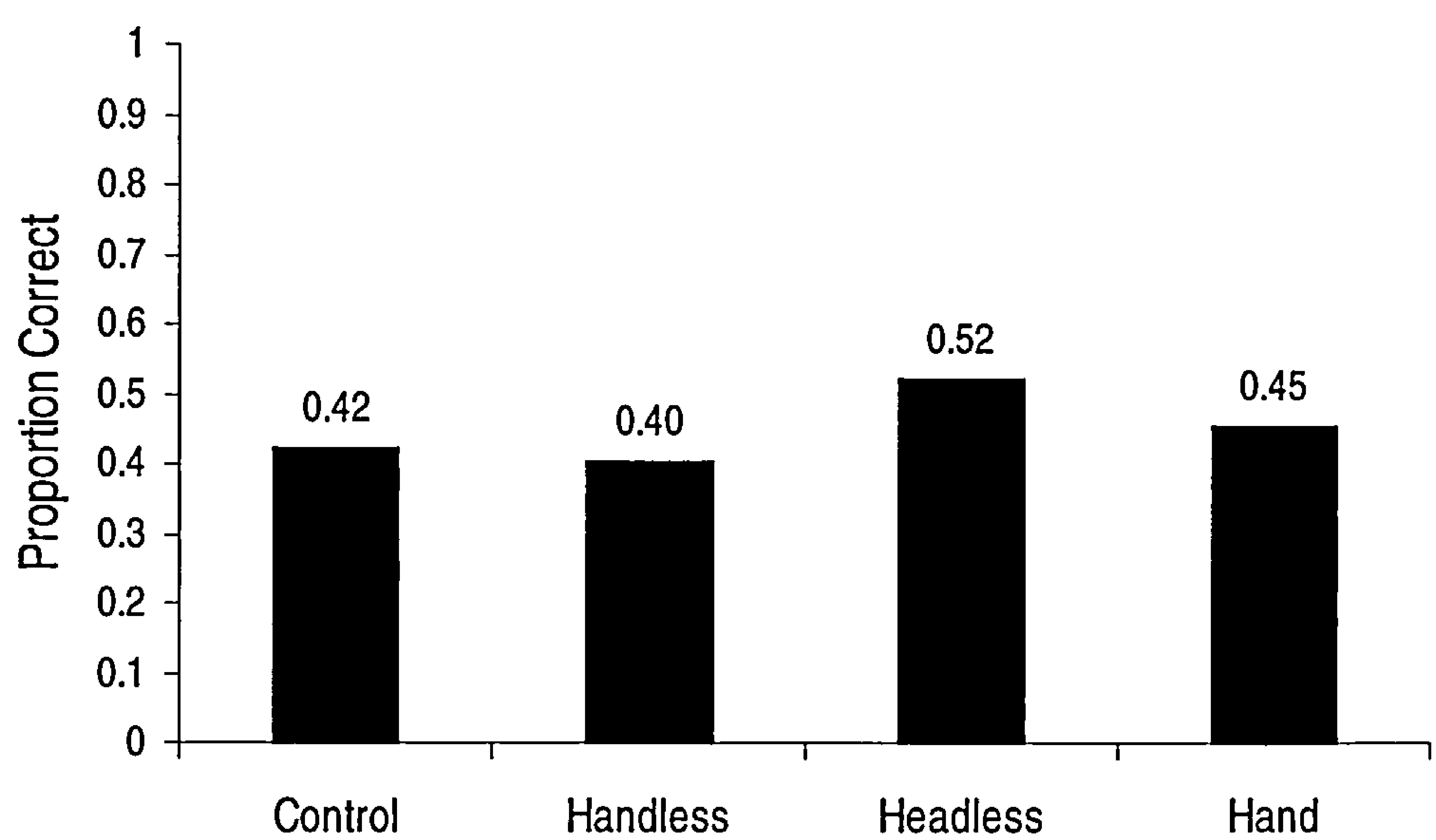

Figure 5.8 Mean proportion of correct recognition for affects plotted for each condition in Experiment 5.

These observations were formally tested in a 4 x 5 mixed design ANOVA in which the different conditions constituted the levels for a between groups variable and the emotions were presented as five levels in a repeated measures variable. The data was averaged over actor and action since these factors were not of great interest in the current experiment and have been discussed in Chapter 2.

The results of the ANOVA showed main effects of condition [F(3,440=4.773, p<.01] and emotion [F(4,44)=4.204,p<.01]. The interaction between condition and emotion, however, was non-significant [F(2,44)=1.406, p=.167]. Post Hoc Tukey tests of the differences between means of different conditions showed only two honestly significant differences. The first of these occurred for the difference between the control group and the headless group, with headless stimuli being correctly categorised more than control stimuli at p<.05. Headless movements were also better recognised than movements without a hand with a HSD at p<.01. There was no HSD for the difference between movements with only a hand and those without a head. Post hoc Tukey tests of the mean differences between pairs of emotions showed HSDs for *angry* movements compared with *happy* and *neutral* movements at p<.05. In both cases *angry* movements were better recognised.

The findings from this experiment show that local features play a minor role in the perception of emotion from movement. Our results do not agree with those of (Mather, Radford, & West, 1992) who found that for judgements about coherent movements and walking direction, the wrist and ankles were vital for accurate recognition. Similar to our design Mather, Radford, & West (1992) selectively deleted some movements of joints from their displays. When shoulder and hip or elbow and knee joints were missing, there was no effect on accuracy in either task. In comparison, when the wrist and ankles were missing, performance dropped by about 30%. They concluded that for walking, the characteristic movement of wrists and ankles were detected by the visual system for use in making judgements. The characteristic movement was based in asymmetry of the xy position of the wrist and ankles that were absent from the xy position of other markers.

In our movements, however, these asymmetries were also found in the elbow, and shoulder as can be seen in Figure 5.9 where the position of head, shoulder, elbow and wrist has been plotted. Unlike walking movements, arm movements are not cyclic and also rely on bending and rotation of the limbs that are not needed in walking. This introduced characteristic human movements to joints such as the elbow and shoulder, resulting in them carrying enough information for a visual analysis of emotion.

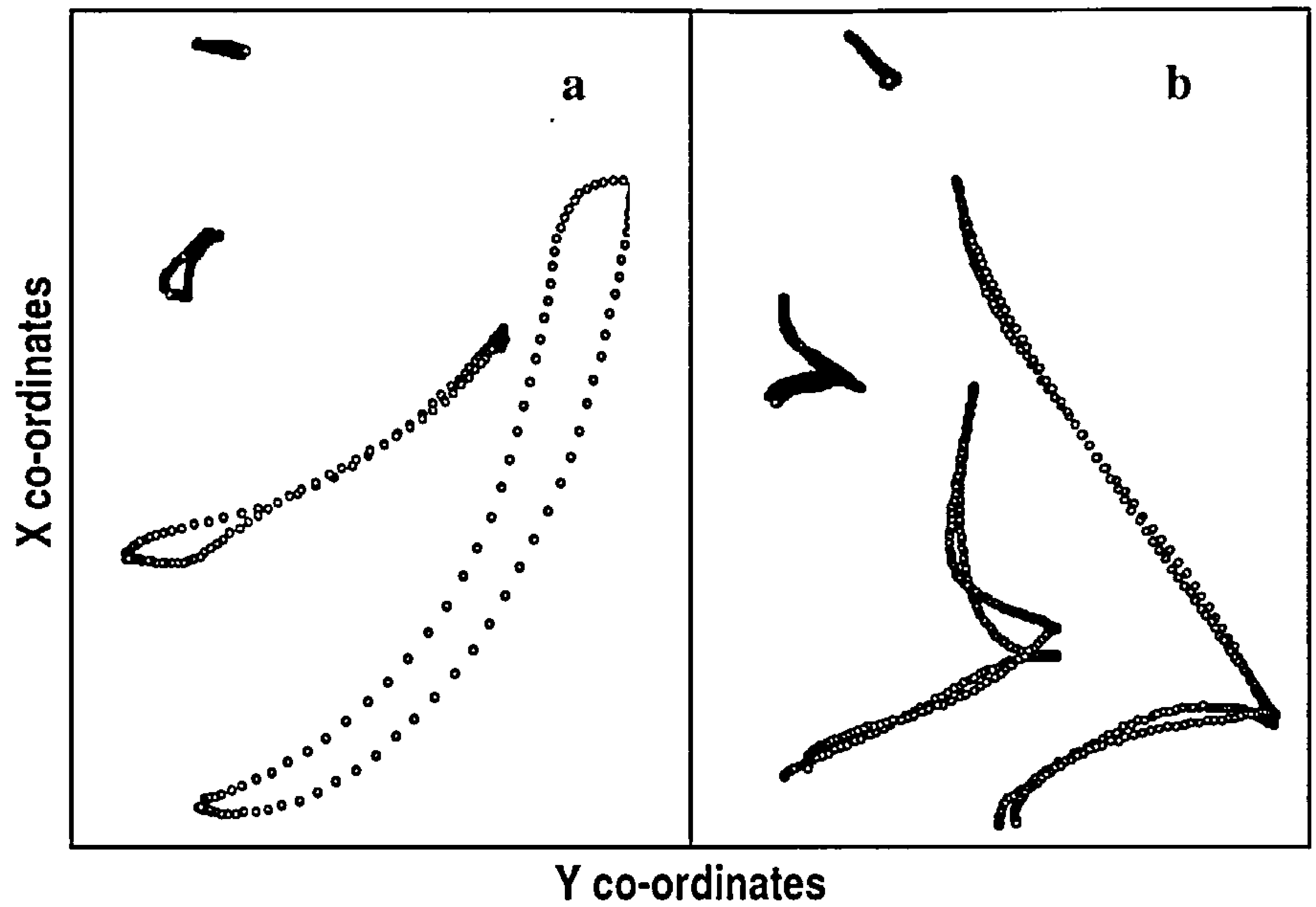

Figure 5.9 The x and y position data for a) knocking and b) drinking movements plotted against each other in the same way as Mather, Radford and West (1992). Each point represents a xy co-ordinate in time so that the full time course of a movement is plotted in one graph. The points corresponding to each limb form a unit that represents the movement of that joint. The movement of the metacarpal joints has not been plotted so in this depiction the joint that move the greatest distance is the wrist while the head and shoulder are depicted by the smallest movements sequences with the wrist in between.

Although there was attenuation of recognition when some features were not shown, this did not approach the dramatic effects observed in Experiment 4 when movements were phase scrambled. This suggests strongly that the spatial information for categorising emotion were distributed throughout displays of arm movements. The relevant global form cue seems to be contained in the time relationships between different joints. Such timing is more complex in non-cyclic movements that rely on bending and stretching and it would be interesting in the future to test the findings of Mather, Radford and West (1992).

A related finding emerged concerning the perceived posture of actors as they performed movements. When the point that depicted the head was missed out of the displays, movements were categorised with greater accuracy. This suggests that postures defined by the head and arm configuration, is an imprecise cue to emotion, which is particularly interesting since many casual and experimental observers of these displays have in the past commented on using posture to judge emotion from the movements. Common sense suggests that posture should play a role in recognition of emotion and observers seem to use this as a heuristic to judge mood. For these displays, however, this is a bad cue and instead of enhancing performance, the heuristic caused interference in the judgement task.

Experiment 6 - Point-Lights vs. Solid Body Model

Introduction

Recognition of action from point-light displays has occupied a unique position in the study of visual perception and cognition. The attraction of these stimuli is their ability to omit irrelevant aspects of the human form that might make recognition or detection trivial. In some situations, however, point-light stimuli are too abstract to reliably depict person characteristics. Recent advances in computer animation has resulted in fairly inexpensive and generic desktop computer tools such as Poser and Life Forms for the animation of human movement. These animation tools allow form information to be added systematically back into the displays. The advantage of this – over other video recording of movements - is that the same fine control is available for digitally captured human movement is extended over the form information. Increasingly human movement is being portrayed by the use of such techniques for both entertainment (Hodgins, 1998) and research (Hill & Johnston, 2001) and this trend is set to continue as animation tools become move sophisticated and user friendly.

Hodgins, O'Brien, and Tumblin (1998) describe a series of experiments that set out to test the sensitivity of human observers to detect anomalous movement from stick figure movement and movement rendered with a polygonal model. Stick figures are similar to point-light displays in that they reduce the amount of extraneous information so that only motion is available. The results showed that human observers were more sensitive to strange movements when they were presented as polygonal models rather than stick figures. The authors did, however, suggest that this might not be the case for all kinds of action recognition.

In the following experiment a solid body human model was used to present movements and the question we were addressing was; to what extent can additional form information aid the perception of emotion from human movements?

Methods

Participants

12 student volunteers participated in the experiment. Participants were naïve as to the purpose of the experiment and were paid for their time.

Movement Recording and Stimuli

The movement of a single male actor was captured using a three-dimensional movement analysis system (Optotrak) that recorded the position of infra red emitting diodes fixed to the torso and right arm of the actors. As his movement was recorded, the actor performed knocking movements with emotion. Emotions were elicited using 10 short scripts representing *angry*, *afraid*, *excited*, *happy*, *neutral*, relaxed, *sad*, *strong*, *weak* and *tired* emotion; after reading each script actor walked up to a door and knocked three times. Each movement was repeated a number of times.

Movements were processed and used to animate a humanoid solid body model (HSM's) as described by Ude, Riley, & Atkeson (2000). From the HSM's point-light displays (PLD's) were made. A start and endpoint of each movement was calculated using the tangential wrist velocity and was defined as the point where this velocity fell above or below 5% of the average wrist velocity respectively. The start and end points defined the first and last fame of the animations and were used in an animation algorithm to make a QuickTime movie of each stimulus. After processing the movements, there were 2 examples of each emotion for each condition in the experiment (PLD's and HSM's).

Design and Procedure

The experiment was programmed using PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993), a specialist software package for running behavioural research experiments in Psychology. The Program was run on an AppleMac G4 in an experimental booth. Movements were blocked according to condition (point-light displays and humanoid solid body model displays) and each block was shown alternatively, three times with half the participants seeing PLD's in their first block experimental trials. There was also a short practice block during which participants could familiarise themselves with the procedure and stimuli. There were 20 stimuli each experimental block and these were not repeated within experimental blocks.

At the start of the experiment, participants were randomly allocated to seeing point-light displays (PLD's) or humanoid solid body models (HSM's) before being given a set of written instructions to read. After a verbal repetition of the instructions, the participant was guided through the practice session. Following this the 6 experimental blocks were shown. During each trial participants were shown a QuickTime movie followed by a screen in which they could make their response by typing one of the 10 emotion words.

Results

Results from the experiment were analysed according to the proportion of correct recognition so that results from the two conditions could be compared. A psychological space was also constructed for each of the conditions.

Correct Recognition of Emotion

The proportion of correct recognition PLD's and HSM's has been graphed in Figure 5.10. The means for each condition was practically identical with point-light displays being recognised with an overall accuracy of 37.8% and displays of HSM's being recognised with an accuracy of 37.1%. As before, different emotions were recognised with different accuracy and for these movements *angry* and *happy* were best recognised.

These findings were confirmed by calculating a 2 x 10 repeated measures ANOVA with two factors. The first factor represented the two experimental conditions and the levels of the factor were the 10 emotions. Results of the ANOVA showed that there was no main effect of condition [$F(1,11)=.111$, ns] or any interaction between condition and emotion [$F(9,99)=1.344$, ns], there was, however, a main effect of emotion [$F(9,99)=5.858$, $p<.001$]. Post hoc test of the differences between emotions showed that *angry* movements were significantly better than *excited*, *afraid*, *neutral*, relaxed, *sad* and *weak* and that *happy* movements were better recognised than *afraid*, *sad* and *weak*, all at $p<.05$.

Further t-tests that compared only the first two blocks of stimuli presented to participants, similarly showed that there was no difference between recognition of PLD's and HSM's. This was the case for the 6 participants that saw PLD's first [t=-1.695, df=5, ns] and the other 6 that saw HSM's first [t=-1.074, df=5, ns].

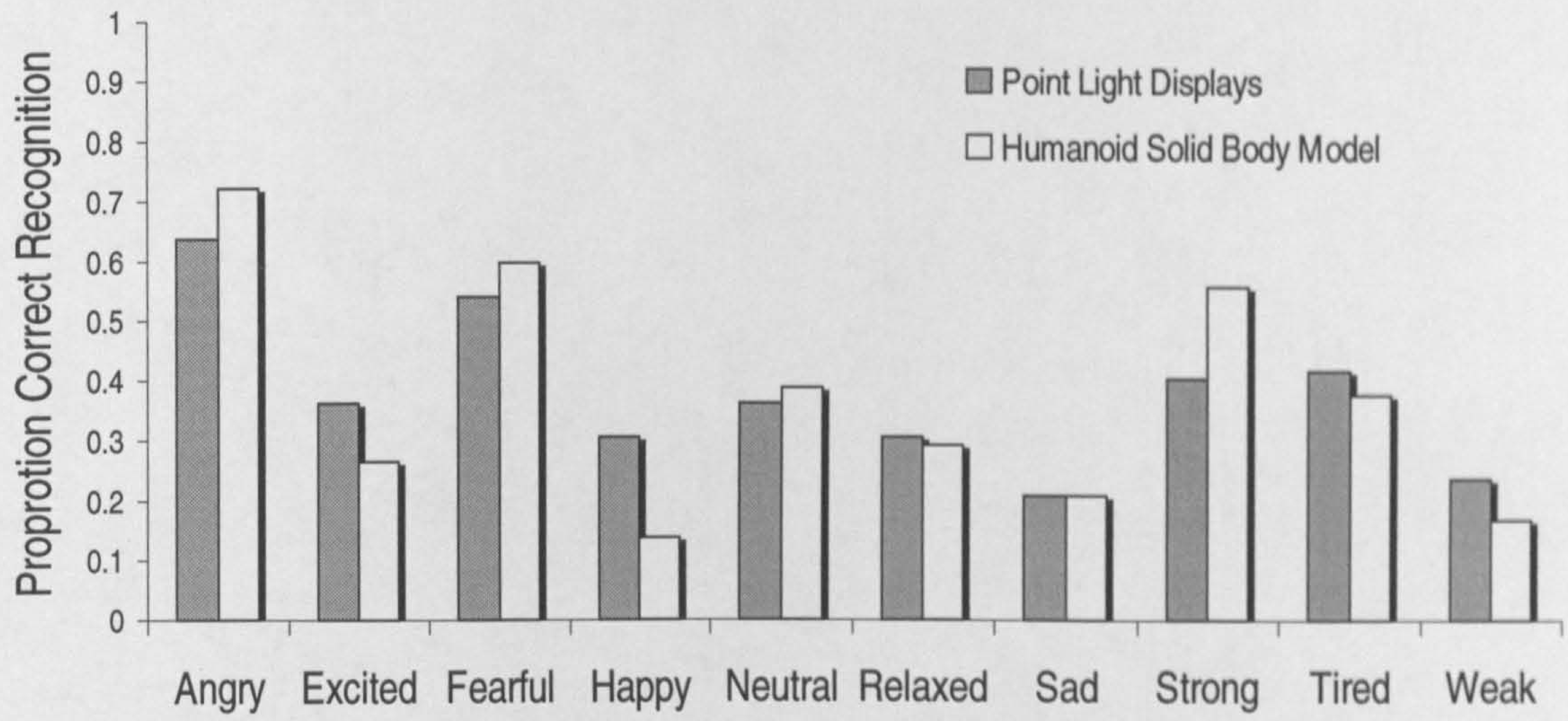

Figure 5.10 Proportion correct recognition of emotion for Point-light displays and Solid Body Model.

From these results it would seem that adding more form information poses no benefit for recognition of emotion. It may however still be the case the strategy participants used categorising emotions would be different. For this reason a psychological space was constructed from the data for each participant.

Comparison between Psychological spaces for PLD' and HSM's

The data for each participant was processed to form a confusion matrix each condition. The confusion matrices were converted to measures of dissimilarity and input to the INDSCAL multidimensional scaling algorithm. For each condition the resultant psychological space was 2-dimensional these are both shown Figure 5.11.

The fit for both models was reasonable with the model for PLD's having stress=.125 and r^2=.897. for HSM's the stress was .139 and r^2=.888. The two psychological spaces bear a very close relationship to each other, an observation that was confirmed by a *strong* correlation between the dimension of the two models. For dimension 1 the correlation between the two models gave r^2=.95 and for the second dimension the correlation gave r^2=.965. No kinematics markers have been measured for these movements; however, it seems that dimension 1 again encapsulates arousal as perceived from movements. For dimension 2 there is a different configuration from that in Experiment 1. In the present case dimension 2 does not clearly show valence. It has previously been observed that there are individual differences in the way that actors portray movements and this may be why the second dimension for this model is so different from previously.

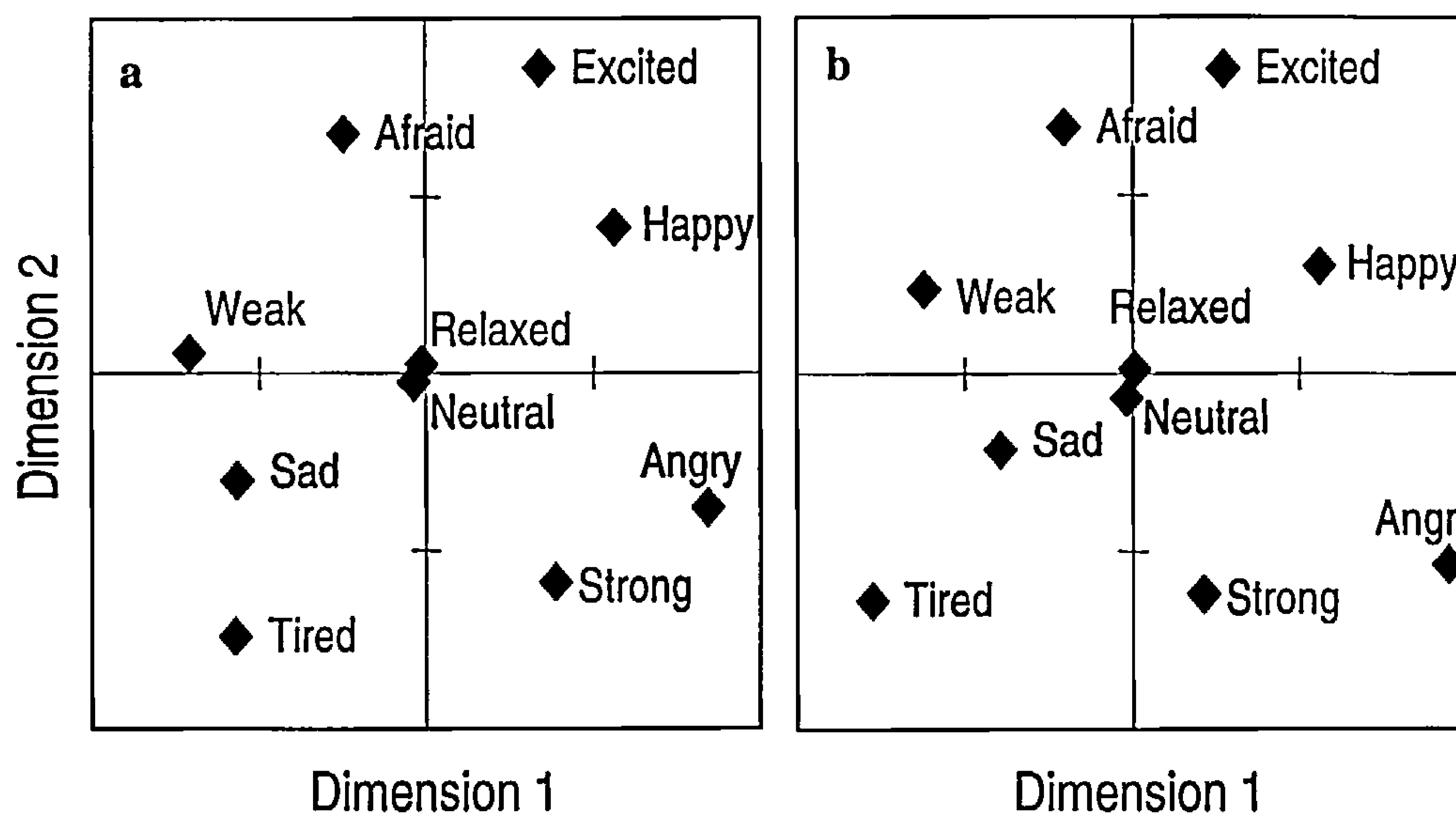

Figure 5.11 Psychological spaces for a) point-light displays and b) humanoid solid model displays.

It seems from these results that enhancing form information does not emotion the perception of emotion from the movements.

Experiment 7- Beliefs about Animacy

Introduction

During the course of Experiment 4 it became clear that in a judgement task about emotion, the display does not have to look like a person or be interpreted as being a person for observers to still make judgements about emotions. In the following experiment, we were concerned with testing this observation.

The result that humans have little trouble in attributing social characteristics, such as emotion, to objects or entities that they do not categorise as human, is not really that surprising. This is a characteristic of human cognition that has been under investigation since 1944 with the publication of Heider and Simmel's seminal paper on apparent behaviour (Heider & Simmel, 1944). In their experiments, they animated geometric forms to move as though they were interacting in a social way. Participant viewing these displays related stories about what the characters in the animations were doing and these descriptions included attributions about personality, intentions and emotions. Although research into this phenomenon has progressed slowly over the last 60 years, a vast repository of examples of this form of attribution can be found in animations that have been made for entertainment.

For the displays in Experiment 4, participants were told nothing about what was being represented, they were only instructed to categorise the movements according to emotion. For participants who thought the moments were human, there seemed to be no difference in the perception of emotion than for participants that did not think they saw biological motion. In the following experiment, we made an attempt to manipulate participant's beliefs about what was being portrayed in the displays so that the effect of a person's belief about the object making moments could be assessed. There were three conditions in the experiment between which the only difference was the instructions that were given to participants at the start. The task given to participants was the same as in Experiment 4 and the proportion of correct responses were measured in each experimental condition.

Methods

Participants

Participants were 42 individuals from the student population at the University of Glasgow. This number included 14 participants for each experimental condition. A 43rd participant was rejected because it later emerged that she had taken part in a previous biological motion experiment, another participant was recruited in her place. Of the 42 participants, about half (19) were excluded from the analysis because they did not believe the instructions presented to them by the experimenter – for further information, see the results and discussion.

Stimuli

Stimuli were a selection of the phase-scrambled movements from Experiment 4. Movements with *afraid*, *angry*, *happy*, *neutral* and *sad* emotion that were recorded from 2 actors were presented to participants. For each emotion there were three examples yielding 30 stimuli in total.

Design

The stimuli were presented to participants in three different experimental groups and were arranged into two experimental blocks. Each block of movements consisted of the movements recorded from a single actor. There was also a short practice block in which participants were shown about 4 randomly chosen displays so that they could become familiar with the displays and procedure. Each stimulus was repeated three times during an experimental block and was presented in random order. The order of experimental blocks was also randomly allocated so that half the participants saw the movements for Actor1 before they saw those of Actor 2.

Procedure

On entering the lab, each participant was randomly allocated to one of three experimental conditions. They were then presented with a set of written instructions. The instructions for the three conditions were identical apart from for a single critical passage. In the instructions the participants were told that they would see some moving stimuli and that their task was to categorise the displays according the emotion being represented. They were then presented with the 5 labels or emotions.

The critical passage in which stimuli were described, is shown below for each condition:

Condition 1 – Scrambled Arm
Each display is made up of moving dots that represent an arm movement that we have scrambled. Your task is to assign the display with one of 5 feelings. This will allow us to judge the accuracy of our displays.

Condition 2 – Molecules
Each display is made up of moving dots that represent a simulation of molecules, moving in an emotional way. Your task is to assign the display with one of 5 feelings. This will allow us to judge the accuracy of our displays.

Condition 3 – Abstract Concepts
Each display is made up of moving dots that represent a simulation of abstract emotional concepts. Your task is to assign the display with one of the 5 feelings below. This will allow us to judge the accuracy of our displays.

Descriptions were based on the observations made by participants in Experiment 4 about what they thought the displays had represented. Condition 1 represented a control condition since most participants had described the movements as some form of coherent human or animal movement. In Condition 2 an attempt was made to make it seem as though the displays were of something that was not human or animal in any way, molecules represent a good example of this. It was hoped that participants would interpret the displays as being made up of individual objects that were either unconnected or rigidly connected. In the final condition no attempt was made to define the displays, instead they were labelled as abstract.

After reading the instructions participants were given a short practice session so that they could become familiar with the displays and procedure. The experimental blocks followed this practice session. Each trial started with the presentation of a movement and participants had to make their response by selecting one of the 5 labels that described the emotions.

At the end of the experiment there was a short debriefing session which included two questions designed to elicit whether participants had interpreted the movements as they were described, or whether they interpreted the displays in a different way. In the first question they were asked if the displays had been effective displays of human movements/molecules moving/abstract concepts of emotion, if they indicated "no" they were asked to explain what they thought the displays had represented. In another question participants were asked whether they had used any particular context, situation or story when deciding what emotion was being portrayed.

Results and Discussion

During the course of the experiment it became clear that a large proportion of participants did not accept our descriptions of stimuli, but instead made their own judgements about what was being portrayed in the displays. For this reason, participants were firstly split into two groups; those that believed the displays were of scrambled arm movements, molecules/dots or abstract concepts according to the instructions that were presented to them. The other group of participants did not believe the instructions, but instead they interpreted the displays for themselves. An example of this was the case in which an observer was told that the displays showed the movements of molecules interpreted them to be people. In order to avoid any possible confounds in the analysis, it was decided to exclude such cases. This yielded 23 participants, 7 each in Conditions 1 and 2 and 9 in Condition 3.

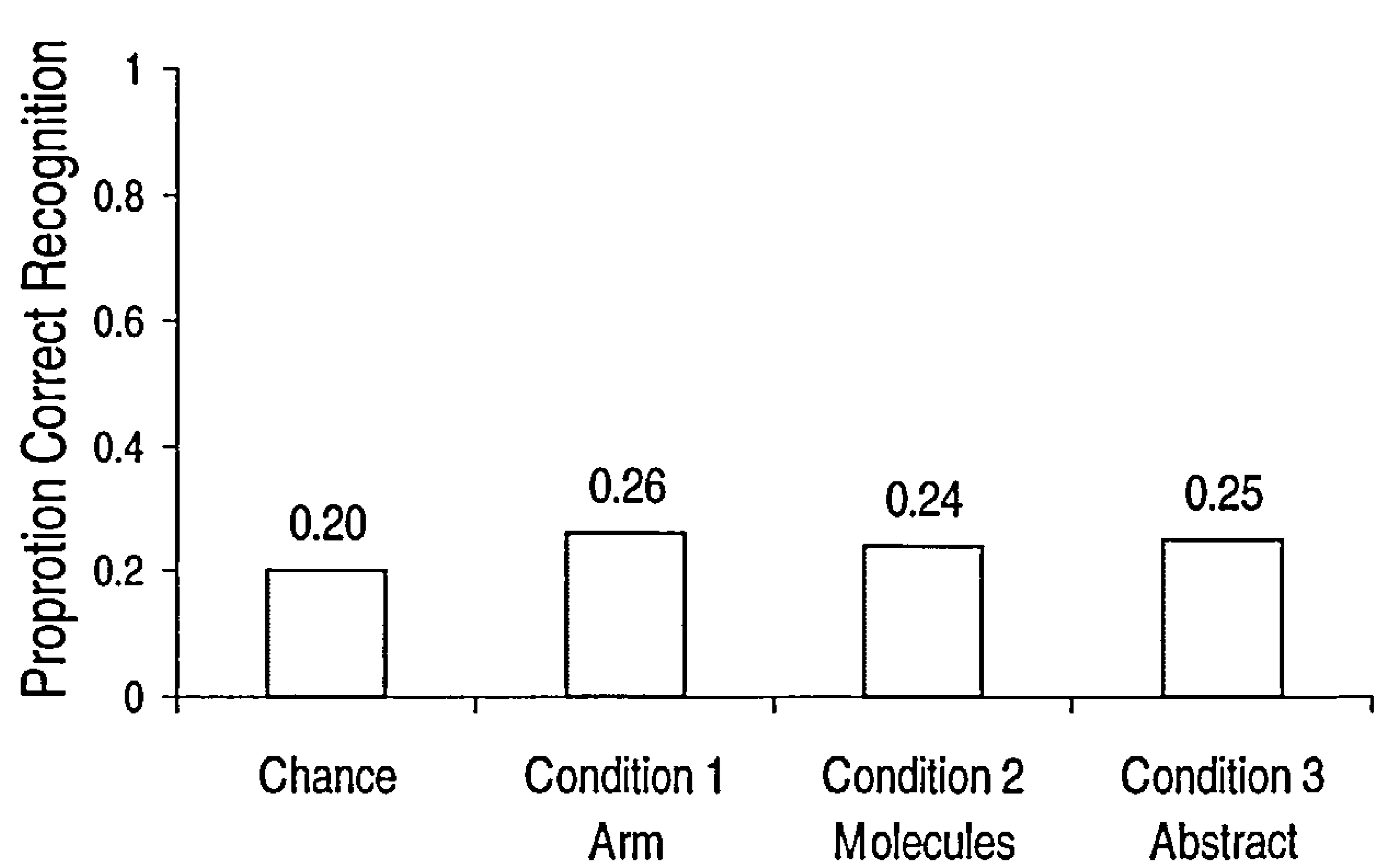

Figure 5.12 Means for the three conditions in Experiment 7 plotted against chance

The means of each condition has been plotted in Figure 5.12. From this figure it is clear that, as in Experiment 4, recognition for emotion was severely impaired by scrambled displays. Tested against chance, however, all three means were found to be significantly better than the chance level of .2 in a one-tailed paired samples t-test [Condition 1; $t=2.334$, $df=6$, $p<.05$; Condition 2: $t=2.707$, $df=6$, $p<.05$; Condition 2: $t=2.746$, $df=8$, $p<.05$].

The differences between conditions were tested in a 3 x 2 x 5 mixed design ANOVA. The three conditions formed the levels for a single independent groups factor while the two actors and the 5 emotions were the levels of the two repeated measured factors.

Results showed main effects for actors and emotions [actor: $F(1,20)=21.295$, $p<.005$; emotion: $F(4,80)=4.208$, $p<.005$] and a significant interaction between actors and emotions [$F(4,8)=6.678$, $p<.005$]. There were, however, no main effect of the instructions that were given to participants [$F(2,20)= 0.427$, $p=.66$] and no interactions between the three conditions of instructions and the other two factors [3-way interaction: $F(8,80)=1.256$, $p=.28$; 2-way interaction with actor: $F(2,20)=1.824$, $p=.19$; 2-way interaction with emotion: $F(8,80)=1.564$, $p=.15$]. These findings meant that our different descriptions of the displays did not influence participant's ability to recognise emotions. This is an encouraging result since it supports the observation that participants perform equally well – or badly – independent to their beliefs about what is making a movement.

Summary and Conclusions for Chapter 5

The four experiments above were undertaken in an attempt to understand better what type of stimulus characteristics are explained by the second dimension along which humans categorise movement with respect to emotion. In particular we focused on spatial features of the movements since these had previously been hypothesised as presenting a valid physical cue along which to judge differences in movements.

In Experiment 4 we found that a global disruption of the movements, through temporal phase shift of individual points, had a dramatic effect on the perception of emotion from motion. However, since both global and local features were affected by this manipulation, it was not clear what spatial feature had been important for categorising emotion. In Experiment 5 we found that the effect of missing out local features in the movements had only a small effect on the accurate perception of movements.

Additionally it was interesting to note from Experiment 5 that the posture of actors was a cue that distracted observers from accurate recognition of emotions. This suggested that more global features, such as the relationship among points, or the form described by the movements, were an important cue in judging emotion. To test whether form information was important for the perception of emotion, human-like features were added to movements. To do this a humanoid solid body model was constructed and human movement data was use to drive it. The result from this experiment seemed to indicate that there was no benefit to adding feature information. In the case where form was of paramount importance, there would have been a marked change in the categorisation of movements. Since this was not the case, it seems possible that the temporal relationships between points play a major role in the categorising emotion from movement. Moreover, it seems that this information informs the second dimension along which emotion is categorised from movements.

An interesting additional finding from Experiment 4 was that observers did not naturally assume emotional stimuli to be of human movements. However, whatever their beliefs about a stimulus, performance on the recognition task was the same. This observation was tested further and confirmed in Experiment 7 by changing the instructions about what the displays represented.

Chapter 6– Cognition of Emotion from Motion

Introduction

The results from Experiments in Chapter 5 highlighted the fact that although low level stimulus properties might drive perception of emotion to a large extent, there are a complex interaction with top-down cognitive processes. An example of this was shown in Experiment 5 where it seemed that observer's beliefs about the role of posture interfered with accurate recognition of emotion.

In the following experiments we will attempt to highlight some of the cognitive processes that could influence recognition of emotions. In Experiments 8 and 9 the effectiveness of out stimuli will be tested, firstly in a limited exposure free response (Experiment 8) task and secondly in comparison to fully lit video displays (Experiment 9). In Experiment 10 observers were given the opportunity to freely name emotions in order for us to assess the appropriateness of our stimulus labels. Finally in Experiment 11 the similarity ratings of emotion labels were compared to the ratings of similarity between movements to test the correspondence between the representation space of these two stimuli.

A. Stimulus Efficacy

Experiment 8 – Free Categorisation I

Introduction

A common feature in all the research so far has been that the tasks for participants involved categorising movements according to a number of pre-set labels, focusing on any emotional content the displays might have. This kind of task has allowed us to study whether observers could extract the information needed to make judgements about emotion. In contrast to this, we thought it would also be valuable to discover whether the information about emotion is actually obvious in the displays when observers have no clue about the intentions of actors to convey emotions. The following experiment was designed for this purpose.

The design of Experiment 8 was based on experiments that were conducted by Johansson (1973, 1975) and Sumi (1984). In this research displays of biological motion were shown to a large number of observers and data was collected about the actions performed. In our experiment, movements with *angry* and *sad* emotion were presented to students at the end of a lecture. The observers were asked to freely comment on the displays by writing on a short question sheet.

Methods

Participants

Participants were the attendees to a Level 1 Psychology lecture at the University of Glasgow. Since the undergraduate course is very large, lectures are delivered twice a day, at 9am and 5pm. In the morning lecture 290 students took part in the experiment. In the evening lecture, 196 students participated. Of the total number of participants, 77 either did not complete the question sheet properly – by missing out an answer to one of the questions – or indicated that they had taken part in a prior experiment on emotion from motion. These participants were excluded from the analysis.

Stimuli

Stimuli were three of the point-light displays of human movement data from Experiment 1 that were transferred to video and displayed on the video projection system in each lecture theatre. In order to minimise the number of exposures to stimuli with emotion, only two movements with emotion were selected for presentation to observers. These were *angry* and *sad* **knocking** movements.

Angry movements were selected mainly because of their place at one extreme of dimension 1 in the psychological space from Experiment 1. *Angry* was also preferred as a biologically relevant stimulus according to the alarm hypothesis (Walk and Homan, 1984). *Sad* movements were selected because they represented the opposite end of the arousal dimension from *angry* movements. We hoped that by presenting extremes of dimension 1 we would encourage the perception of emotion by observers. A third stimulus was a *neutral* **drinking** movement that was included to allow some practice during which observers could become familiar with the display format.

In order to present the stimuli to a large audience, graphics software was used to make the displays into 2 different movies that will be called stimulus clip 1 (SC1) and stimulus clip 2 (SC2) henceforth. The two stimulus clips were transferred to a VHS videocassette with video capture software. The stimulus clips were identical apart from the order of *sad* and *angry* knocking movements on the critical trials.

The first presentation in both clips was of the *neutral* drinking movement – this was repeated once. For SC1 there then followed a knocking movement with *angry* emotion and finally the *sad* movement. In SC2 the *sad* movement was presented before the *angry* movement. After each of the 4 displays, the video clip was paused so that observes could make their responses on a question sheet. There were three questions that were also labelled on the video clips to minimise confusion for participants.

Design

For symmetry in the experiment, there were two conditions to the order of critical trials. In the first condition, which was presented to observers at the morning lecture, *angry* movements were shown first, followed by *sad* movements. The second condition was presented to students in the evening lecture and in this case *sad* movements were presented first.

Question Sheet and Procedure

At the start of the experiment observers were presented with a short question sheet. On the front page there was a brief set of instructions that detailed the procedure of the experiment. This question sheet has been reproduced in Figure 6.1 and was stapled so that only the front sheet could be seen. Participants were asked not to open their sheets or look at the other questions until instructed to do so. All instructions were verbally repeated before the start of the presentation and participants were encouraged to use their own words to describe what they saw in the context of the questions in the sheet.

Practice Session: After reading the instructions, participants were shown the first display. This display was of a drinking movement with *neutral* emotion. Their task was to answer the first question after viewing the movement. Question 1 was:

Q1 What did you see in the display?

After all participants had finished writing, they were told that the movement had been of a person making a drinking action. The display was then shown again along with a verbal description of the movement. Participants were told that they could now open the question sheet and look at the other questions.

Thank you for taking part in this experiment.

You will see 3 moving displays and for each there are some questions to answer.

Please do not look at Q2 or Q3 before answering Q1. There are no right or wrong answers to these questions, but it is important that you answer them honestly and with the first impressions that you have

Q1. What did you see in the display.

Q2. Complete the three lines below, each with a different answer:

The person is...

The person is...

The person is...

Q3. Complete the three lines below, each with a different answer:

The person is...

The person is...

The person is...

Please enter the time here ....

Please enter your name and matric number ....

....

Have you ever seen these stimuli before? Yes/No

If Yes, where and when?

Thank you for taking part in this experiment. All the information here is confidential and your name and matric number will be used only to ensure that you have not taken part in this experiment more than once.

Figure 6.1 Reproduction of the question sheet presented to observers for Experiment 8.

Experimental Session: The task in the experimental session was to look at displays presented and answer questions 2 and 3. Both questions were as follows:

> ***Q2/3 Complete the three lines below, each with a different answer:***
> ***The person is...***
> ***The person is...***
> ***The person is...***

The first knocking movement (for SC1 this was an *angry* movement and for SC2 it was a *sad* movement) was presented and then the movie was paused so that observers could write down their answers. The second knocking movement was then presented and participants wrote down their answers. Finally some personal information was collected on the participants so that observers from previous experiments could be eliminated from the analysis.

Results and Discussion

The observer's descriptions were coded according to whether they described displays as depicting actions with emotions or without emotion. Most of the descriptions with emotion were in response to *angry* movements and were descriptions of aggressive actions rather than recognition of *angry* emotion. Some examples of the aggressive actions that were described are:

> **The person is shaking their fist (this was the most common description)**
> **The person is gesturing aggressively**
> **The person is hitting someone**
> **The person is swearing/making an aggressive/rude sign**
> **The person is making an *angry* movement**

We were rather pessimistic about finding that participants would be able to accurately perceive emotion from these brief presentations. For this reason it was surprising to find that about a 34% of observers in the morning lecture and 11% of observers in the evening lecture described the *angry* movements in terms of aggressive actions. In comparison, *sad* movements were described as having aggressive actions only 7% of the time in the morning and on 3% of occasions in the evening lecture. The proportion of occasions on which participants described displays as being of aggressive actions has been illustrated in Figure 6.2.

For each condition a paired samples t-test was calculated to confirm that *sad* movements were recognised as aggressive fewer times than *angry* movements and the test was significant in both cases [SC1: $t(251)=127.82$, $p<.001$, one-tailed; SC2: $t(156)=3.063$, $p<0.001$, one-tailed]. These results show that participants did not just randomly describe displays as depicting aggressive action, instead this description was reserved for displays in which an actor intended to convey an expression of *angry* emotion.

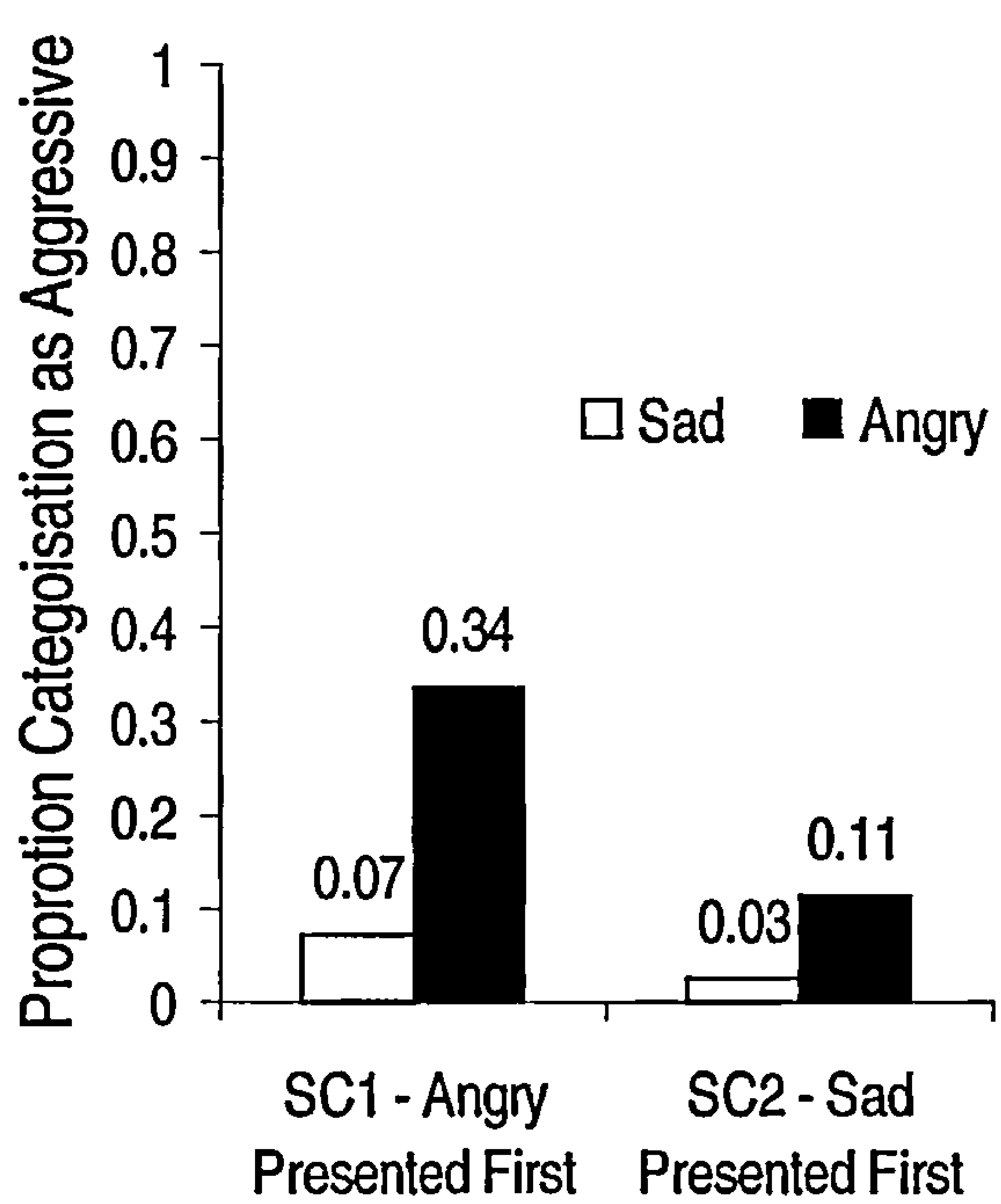

Figure 6.2 Proportion of categorisation as aggressive actions for sad and angry movements In Experiment 9.

In contrast to these results for *angry* movements, recognition of *sad* movements was very poor. There were only two occasions during which participants indicated that they thought the person seemed *sad*. Instead, there was no indication that participants perceived emotion when they looked at *sad* movements. Descriptions were of *neutral* events with only about 16% of participants in SC1 and a single participant in SC2, noting that the *sad* movements were slow.

The most probable reason for this was that the *sad* movements from this corpus presented a very *weak* signal. In Experiment 1 there had been correct recognition of knocking movements with *sad* emotion only 15% of the time, for the most part *sad* movements were confused as having *weak* (33%) and *tired* (15%) emotions and incorrect classification as movements with *neutral* action on 21% of occasions. In comparison *angry* movements were correctly recognised 48% of the time and were consistently confused only as having *strong* emotion (32%) and occasionally *excited* emotion (12%).

The results from the current experiment also showed an asymmetric transfer effect (Poulton & Freeman, 1966) so that *sad* movements seemed to interfere with accurate perception of *angry* movements. In the case where observers were presented with the *angry* movements first (in SC1) they described the *angry* movements as aggressive more times than the observers that saw the *sad* movements first. The proportion correct recognition for *angry* movements from the morning and evening sessions were compared in an independent groups t-test with unequal samples and was significant [$t(404.413)=2.233$, $p<.001$, one-tailed] [22].

It is unclear why the results were not symmetrical, but one explanation could be that humans have a decreasing sensitivity to the potential for socially significant aspects to a stimulus, as a situation becomes more familiar. In other words, they may be more attentive for emotion when the situation is unfamiliar, but as things become more familiar this sensitivity would decrease. This would certainly explain the difference in the attribution of aggression to *angry* movements by participants in the two different conditions. Although this hypothesis does not fall within the remit of the current thesis, it could certainly do with future investigation, particularly with more direct reference to the alarm hypothesis (Walk and Homan, 1984).

In summary, we showed that point-light displays of arm movements with *angry* emotion ***do*** elicit a perception that the displays are of actions with emotion. *Angry* movements presented a *strong*er signal for the perception of emotion and were more likely to elicit a description that entailed some form of aggression. Participants also seemed more sensitive to possible emotional content in the movements when *angry* actions were presented before *sad* movements. In contrast, *sad* movements did not present a *strong* a signal and this meant that they were rarely described as conveying emotion. Additionally, presenting these movements first seemed to detract from the perception of the *angry* movements that followed, as being actions with emotion.

Experiment 9 – Perception of Emotion from Video

Introduction

In Experiment 1 and subsequent experiments it became clear that the overall recognition rate for the accurate perception of emotions is above chance, but rather low. This low rate could be partially accounted for by the fact that some primary motion signals are just not effective at communicating emotions. In Experiments 1, when drinking and knocking movements were compared with each other, we found that emotion was recognised much better from knocking actions. Similarly for actors we have consistently found individual differences in the perception of different actor's movements. The variation caused by individual differences between actors and actions is a reflection of the differences that may be observable in the world and hence we have not tried to systematically limit their impact.

Another factor that may have contributed to the low accuracy were consistent misidentifications and confusions between different emotions. For example, *weak* movements were identified as *weak*, *sad* or *tired* with equal frequency and *excited* movements were almost exclusively misidentified as *angry*.

Another third possible explanation for the low rate of recognising emotion could be that arm movements in general are not an effective stimulus for this task. Alternatively, the arm movements are good actions for displaying emotions, but reducing the information so that only movements of the joints are evident could cause the very low recognition rates.

---

[22] Levene's test for equality of variance was significant indicating that equal variance for the two samples could not be assumed. For this reason the degrees of freedom were adjusted for unequal variance, however, this made no difference to the outcome.

To evaluate these last two hypotheses a control study was undertaken in which the recognition of a smaller number of emotions (*afraid*, *angry*, *happy*, *neutral* and *sad*) were selected for presentation as both full-video and point-light displays.

Methods

Participants

Participants were 8 volunteers from the student population at the University of Glasgow who were naïve to the purpose of the experiment and were paid for their time.

Stimuli

The stimuli used for the current experiment included point-light and full-light video displays of *afraid*, *angry*, *happy*, *neutral* and *sad* knocking movements. The point-light displays were obtained from the original data set used in Experiments 1 and were made into QuickTime movies. Since video clips of these original data did not exist, two new actors were filmed as they performed affective arm movements. By using video capture software, the video recording of movements was transferred to a graphics computer. They were then edited to parse out individual movements. On each of the movement clips a grey oval was pasted over the face of the actor and the clip was made into a QuickTime movie. The process resulted in 30 movie clips (5 emotions x 3 repetitions x 2 actors) for each of the stimulus conditions.

Design

The stimuli were arranged into blocks according to the actor who had performed them. This meant that there were 2 blocks of movements for each of the conditions in the experiment. The blocks for each condition was presented in succession so that a participant always saw the movements of both actors for one condition before they saw any of the movements from the other condition. In each block of trials there were 15 stimuli that were each repeated three times, yielding 45 displays. The order of displays was random within each block and the order of presentation of experimental conditions was alternated, as was the order of blocks for each condition.

Procedure

The experiment was programmed using PsyScope (Cohen et al., 1993) and was run on an Apple Mac G4.

On entering the booth participants were allocated to either seeing the displays of video sequences or the point-light displays first. They were then presented with a written set of instructions in which the task and procedure was detailed. Their task was to watch each display and categorise it according the emotion being depicted.

Participants were told to choose from 5 labels for emotions – *afraid, angry, happy, neutral* and *sad* – when categorising the emotions. Each trial started with the presentation of a display followed by an opportunity to enter an answer, participants could then carry on to the next trial. At the start of each experimental block there were also 2 practice trials consisting of *neutral* movements to familiarise observers with the procedure and stimuli.

Results

Results showed an average correct recognition rate of 59% for point-lights and 71% for full-video, both of which were revealed by a t-test to be significantly different from the chance value of 20% ($p<.005$).

To test the differences in means for the various factors in the experiment 2 x 2 x 5 mixed design ANOVA was performed on the proportion correct. For the between groups factor the participants were split into the group that saw video's first and those that saw PLD's first[23]. The repeated measure factors were the display condition (Video and PLD) and the emotions (*afraid, angry, happy, neutral* and *sad*). It was decided to miss the variance arising from differences between actors out of the analysis since there was no direct correspondence between Actor1 and Actor2 for the two display types[21].

The results from the ANOVA revealed a significant effect of display type [$F(1, 7)=19.768$, $p<.005$], showing that video sequences were better recognised than point-light sequences. There was no main effect of emotion from

[23] We were concerned about possible practice effects and hence included the order of experimental blocks as a factor in the ANOVAS.

this analysis [$F(4, 28) = 1.288$, $p=0.34$] or for the two different groups [$F(4, 28) = 1.191$, $p=.747$]. There were also no significant interactions. Figure 6.3 shows the means for video and PLD sequences for each emotion and indicates a trend for *happy*, *angry* and *afraid* movements to be better recognised from video displays.

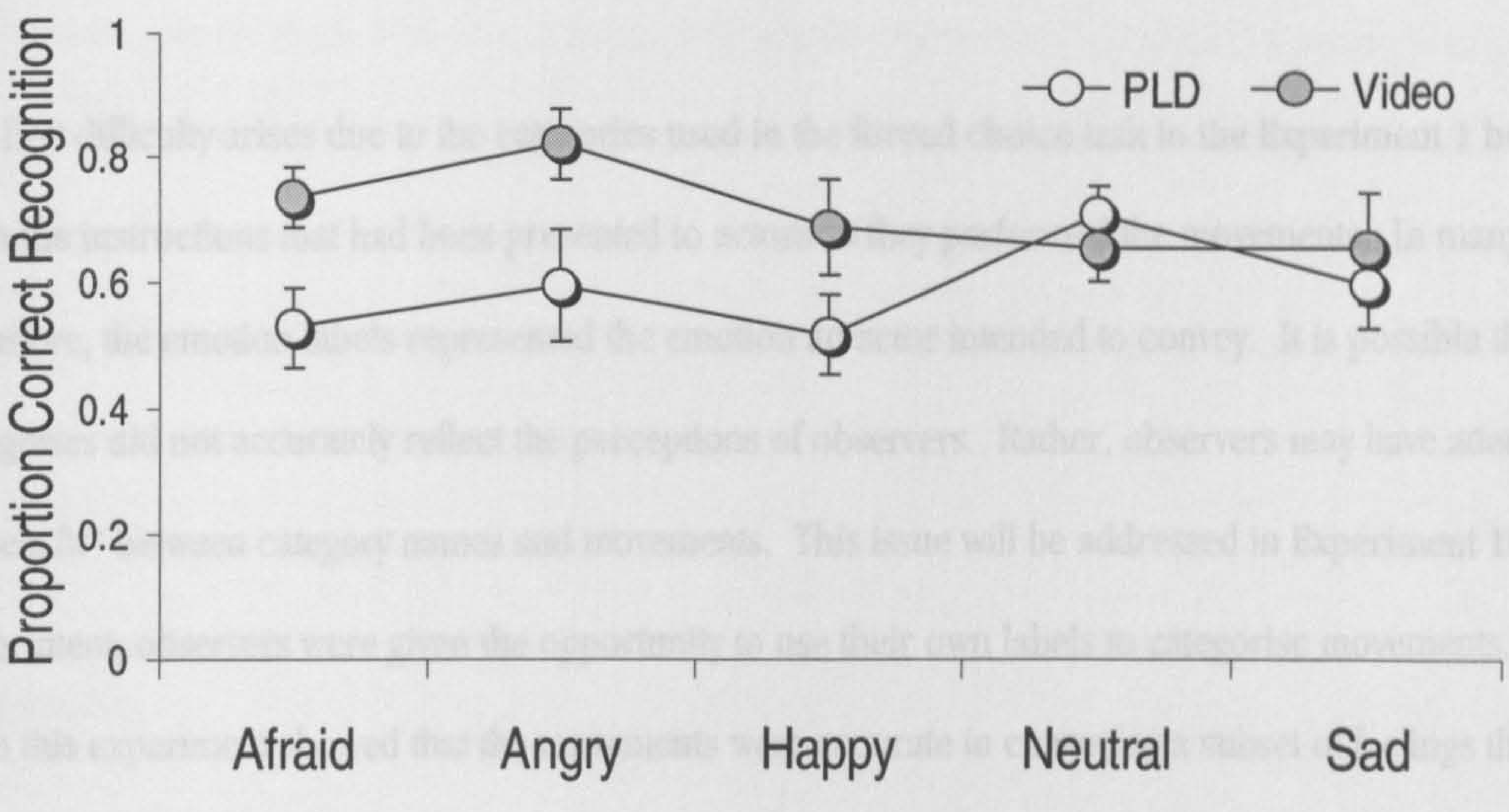

Figure 6.3 Proportion of correct categorisations for video display and point-light display conditions in Experiment 10.

These results are reminiscent of those reported by Dittrich et al. (1996) for full-body, stylised dance movements that showed recognition rates of 63% and 88% correct respectively for point-light and full-video displays. It would thus appear that simple arm movements, while not as effective as stylised dance movements, are by themselves effective in communicating emotion and that both confusions among similar emotions and point-light presentation contributed to the low recognition rates.

[24] As discussed earlier, the actors for the video sequences were different from the actors for the point-light displays. For this reason it would not be sensible to compare the different actors in the main analysis.

B. Free Categorisation and Dissimilarity

In nearly all of the experiments up to this point, observers were instructed to classify human movements according to pre-determined categories. There remain 2 major difficulties with this task and will be addressed in the following experiments.

The first difficulty arises due to the categories used in the forced choice task in the Experiment 1 being derived from the instructions that had been presented to actors as they performed the movements. In many ways, therefore, the emotion labels represented the emotion an actor intended to convey. It is possible that these categories did not accurately reflect the perceptions of observers. Rather, observers may have attempted to find a "best fit" between category names and movements. This issue will be addressed in Experiment 10. In this experiment, observers were given the opportunity to use their own labels to categorise movements. Results from this experiment showed that the movements were accurate in conveying a subset of feelings that corresponded to the emotions that actors had been instructed to depict in their movements.

In experiments 11.1 and 11.2 a second issue arising from previous research will be addressed. This relates to the psychological space discovered in Experiment 1. This psychological space was used to show that movement speed played an important role in the perception of emotion. It is important, however, to establish that the psychological space was not an artefact of either the movements or the categories that were presented to observers, but rather reflected accurately the model that a person might use in the perception of emotion. In this case, dissimilarity data was collected for the movements and categories that had been used in Experiment 1.

Experiment 10 – Free Categorisation II

Introduction

In Experiment 2 observers were given a few very specific categories into which they had to classify human movements with emotion. However, it is not clear whether these labels accurately described the perceptions observers may have had. In the following experiment we allowed observers to use their own labels in describing the emotion depicted in movements. A good correspondence between the spontaneous labels and emotions that actors intended to convey would indicate that point-light displays of arm movements were a reasonable stimulus for the current and future research.

Methods

Participants

Participants were 32 volunteers recruited from the student population at the University of Glasgow. They were compensated for their time and it was required that they had never before taken part in an experiment that used point-light displays in which judgements about emotion was made. The data for 4 participants was excluded from the analysis. Two of these did not fill out the response sheet properly. The other 2 observers were rejected on the basis that in most of their responses they either indicated that they did not know what feeling was being portrayed or answered using labels and phrases that did not describe an emotional state or feeling. Examples of such labels are "thoughtful", "thirsty" and "sober".

Stimuli and Design

Stimuli were the same as those used in Experiment 1. The experimental design was also very similar, the only exception being that each stimulus was shown only once. As before, there were 4 experimental blocks and a short practice session in which observers became familiar with the display format and procedure. The task was completed in one experimental session that was 1 hour in duration. The order of stimuli was random within experimental blocks and the order in which blocks were presented to observers was also random.

Procedure

At the start of the experiment participants were presented with a set of written instructions in which the experimental procedure was explained. They were instructed to look at each display and decide what feeling was being portrayed in the movement presented on the screen. Before going on to the next stimulus they had to write down their response on a response sheet.

Apart from 2 restrictions, there was no constraint on the terms or labels that a participant could use to describe a movement. The two restrictions were 1) the term had to describe a feeling or emotional state and 2) it had to be a single label or very short phrase. Participants were also told that there would be some *neutral* movements and that they should not use this response to indicate that they did not know what feeling was being portrayed. At the end of the experiment, there was a short debriefing session similar to the one in Appendix II. Two additional questions were:

What did you interpret the label "felt" to mean?
An Emotion (i.e. happiness)
A state (i.e. excitements)
A feeling (i.e. boredom)
All of the above
A Combination of the above
Something else

Were you ever unsure of what the person felt?
Did you guess anyway?
If you really did not know, what label/s did you use to indicate this?

During the debriefing sessions, participants were also asked to explain the meaning they had attributed to labels or phrases with ambiguous meaning. For instance "unsure" could mean that the participant was unsure about the actor's feeling, or it could mean that the actor looked felt unsure.

Results and Discussion

Frequency Data

The data for 28 participants was included in the analysis. In total 172 different labels or phrases were used to describe the 118 movements that had been presented in the experiment. On average each observer used 25 different labels to describe the movements with a range of 13-45 labels. Labels that were used with the greatest frequency are illustrated in Figure 6.4.

There were 84 labels that were used only once or twice and an additional 49 that were used with low frequency (each used less than 10 times in the experiment) and by 4 or fewer participants. These terms accounted for about 12% of all data points in the analysis. Compared with this large number of low frequency terms, 25 labels were used with high frequency (each accounting for more than 1% of all data points) or by 10 or more participants. These high frequency labels accounted for 77% of data points in the analysis. A further 14 labels were used with middling frequency and accounted for the other 11% of data.

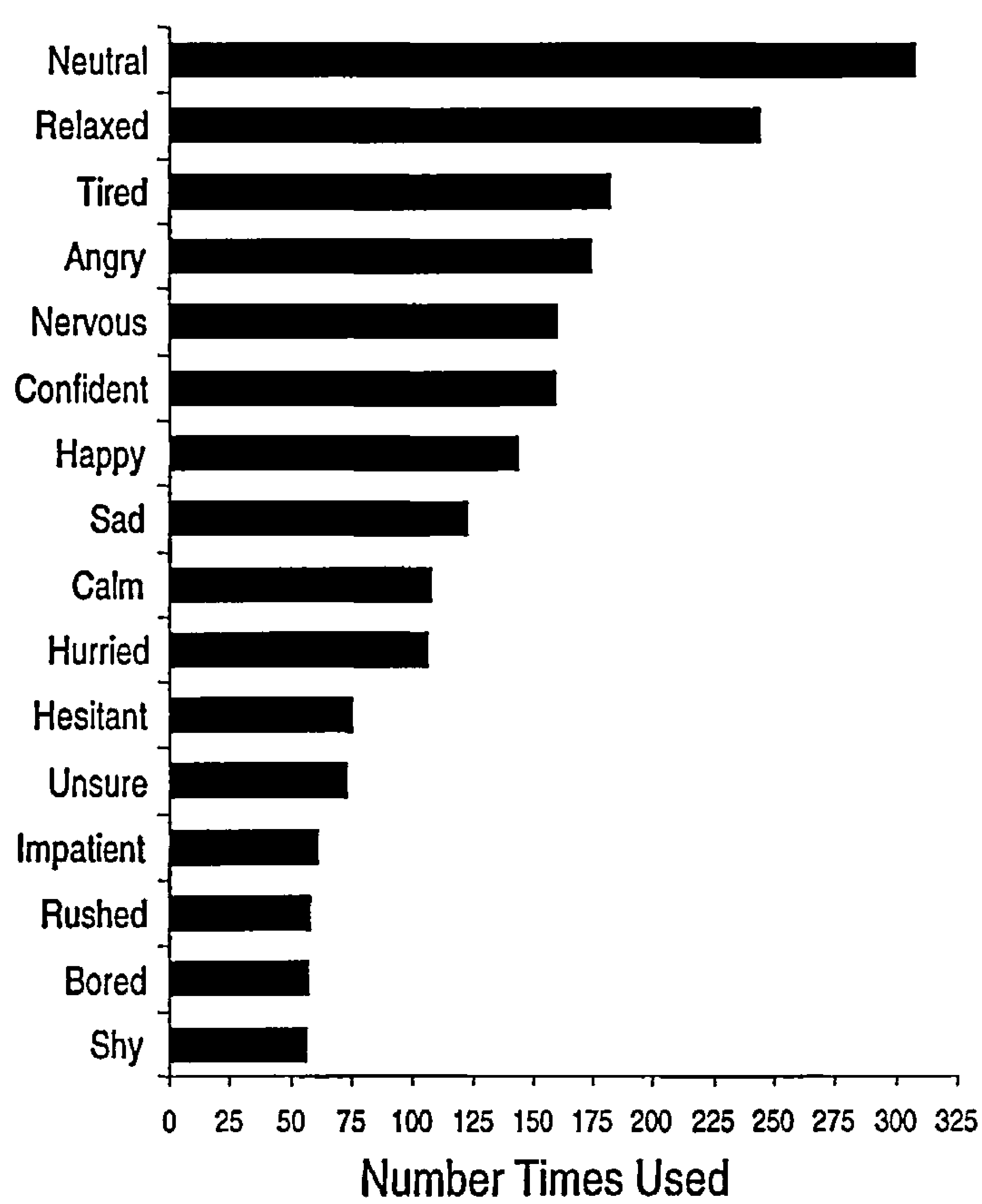

Figure 6.4 Words used most frequently to describe movements with affect.

Based on this frequency data, it seems that there was a fair amount of agreement between subjects about the subset of terms that were appropriate for describing the displays. Additionally, amongst the low frequency and middle frequency labels there were a number of terms that can be regarded as synonyms for high frequency labels. For instance "irate" and "irritable" are synonyms of "annoyed" and "joyous" can be regarded as a synonym of "*happy*". Additionally there were some labels with high frequency labels that are also synonyms of each other. An example of this would be "rushed" and "hurried".

Synonymous Labels

It was clear that there were some synonyms among the labels that were used and it seemed that similar labels might be grouped together to form categories. This was a necessary step in order for further analysis to make sense.

Four independent judges were recruited to sort the labels into groups of synonyms. The judges were all native speakers of English and were paid for their time. Each of the labels that had been used in the experiment was printed on a piece of A6 paper and the judges' task was to sort the labels into as few sensible categories as possible. Judges were instructed that a label's inclusion in a group should only be on the basis that it fitted well with the other labels in terms of its meaning. The cards had to be sorted twice, the first time in order for judges to become familiar with the terms. The second sorting was used for further analysis.

The final task for judges was that each category had to be named, preferably by using one of the labels in the grouping. Labels that fitted in no other category were classed into a category known as "other". These labels were typically terms such as "thirsty" and "elsewhere". There was a good correspondence between the 4 judges and their categories were collated in the following way.

Firstly the categories of different judges were coupled together depending on the name that was given to a group of labels during the sorting task. For instance all four judges had a category named "happy". For each actor the words they included in this category were added to a super-category named happy.

In cases where there was no correspondence in the names of categories, groups of labels that were consistently grouped together were used as the basis for a super-category. For instance "sociable", "friendly" and "conversational" were grouped together by all judges with little agreement over a category name. All groupings in which this configuration appeared were put into the same category.

In 18 cases judges came up with unique categories. In these cases the group was added as a category used by only one judge. At the end of this process there were 42 categories. For 24 of these, more than one judge contributed a group of labels.

From these 42 preliminary categories, 17 stable categories emerged that were used in further analysis. A label was included in one of these 17 categories in cases 2 or more judges had agreed that it fitted and if the other judges had grouped it into 2 different categories. The only case in which this was a problem was for "deflated". For each of the groupings "*sad*" and "*tired*" two judges had included "deflated". A toss of the coin decided that "deflated" should be in the *tired* category.

An additional constraint for categories that were to be used in further analysis was that they had to account for at least 1% of the data points, this meant that they had to be used about 33 times in the entire experiment. Again there was an exception to this in the case of the category "*excited*". Since actors had been asked to portray "*excited*" as one of the emotions with movement, this category was included in the analysis. From the original emotions, only "*strong*" was not placed consistently in a single category, however, it had been used only twice during the course of the experiment. The 17 categories used in the further analysis can be found in Table 6.1.

In a follow up task, the 4 judges were asked to indicate if the 10 terms used in Experiment 3 fitted into any of the categories. They were instructed to only include the label if it fitted well into the category. In this judgement task the categories "Other" and "I don't know" were not included and categories were not named. The reason for not presenting the category names was to encourage judges to base their decisions on how well the original emotion label fitted well in to the group of labels as a whole. The results of this task are included Table 6.1.

Table 6.1 Categories of words and the items included in each. In the grey rows at the top of each group of words; the category name has been highlighted. The thicker grey bar at the bottom shows which of the original affect words – words used to describe to actors what emotion they had to depict – fits into each category. In the category labelled "Other" words from categories that each represented less than 1% of all the data points in the experiment has been grouped together.

Angry	Relaxed	Happy	Hurried	Nervous	Tired
Angry	Calm	Cheerful	Bothered	Frightened	Deflated
Aggressive	At Ease	Happy	Busy	Nervous	Disappointed
Annoyed	Collected	Joyous	Crowded	Scared	Disheartened
Bad Mood	Comfortable	Merry	Distracted	Shocked	Exhausted
Frustrated	Composed	Pleased	Fidgety	Uncomfortable	Lethargic
Furious	Contented	Positive	Flustered	Uneasy	Sullen
Hatred	Peaceful		Hurried	Unsafe	Tired
Impatient	Relaxed		Rushed	Worried	Weak
Irate	Satisfied		Stressed	Afraid	Weary
Irritable	Secure		Urgent	Agitated	Worn out
Irritated	Serene			Anxious	
Moody	Steady			Apprehensive	
Pissed Off				Concerned	
Angry	Happy Relaxed	Happy	Excited	Afraid	Tired Weak

Sad	Hesitant	Confident	Neutral	Shy	Bored
Depressed	Cautious	Assertive	Neutral	Awkward	Bored
Disenchanted	Hesitant	Assured	Familiar	Dainty	Fed Up
Down	Reluctant	Bold	Fine	Meek	
Hurt	Tentative	Confident In	Habitual	Reserved	
Lonely	Uncertain	Charge	Neutral	Reticent	
Low	Undecided	Sure	Normal	Self-conscious	
Mournful	Unsure			Shy	
Sad	Wary			Timid	
Solemn					
Unhappy					
Upset					
Dejected					
Sad Weak	Afraid	Strong	Neutral Relaxed		

Excited	Thoughtful	Determined	Eager	Other		
Excited	Distant	Decisive	Alert	Apathetic	Indifferent	Remorseful
Energetic	Elsewhere	Deliberate	Eager	Ashamed	Inquisitive	Resigned
Enthusiastic	Reflective	Determined	Expectant	Bitter	Intrusive	Sarcastic
Lively	Thoughtful	Purposeful	Hopeful	Cocky Comical	Jaunty	Sensitive
			Keen	Committed	Lazy	Serious
			Ready	Conversational	Light Hearted	Shaky
				Demanding	Not Bothered	Sociable
				Demure	Numb	Strong
				Drunk	Official	Superior
				Edgy	Perplexed	Tense
				Flamboyant	Pretentious	Thirsty
				Forceful	Proud	Sober
				Good	Quiet	Unwilling
				Important	Refined	Uptight
				Imposing	Regretful	Waiting
Excited	Relaxed	Strong				

In the follow-up task criteria for including one of original emotion labels terms as a member to a particular category, was that three or more of the judges agreed that it should be included in that category. In most cases, the target label had already been sorted into one of the categories, so the task was not hard and the results were no surprise. The only additional issue of note was that *excited* was sorted for both it's positive and negative meaning so that it was associated with the grouping "Excited" and "Hurried".

For 13 of 17 categories one or more corresponding terms were selected from the 10 original emotion terms. Three additional categories that did not correspond to the 10 original emotions did emerge and these were Shy, Bored and Eager. From this it is clear that most of the spontaneous labels used by observers fell into categories that corresponded well with the 10 original labels. Six of the original labels (*angry*, *happy*, *neutral*, *relaxed*, *sad* and *tired*) were directly and well represented while categories emerged that encompassed 3 of the others. The worst represented category was *excited* and this may have been because of the ambiguous nature of the label.

How did the Spontaneous Labels Compare to all Possible Labels?

Although the labels used most often by observers corresponded well with the intended emotions of actors, it was not clear whether these labels were picked randomly, or whether they accurately represented only a subset of possible emotions. This issue relates to a problem that has been previously referred to in the literature on categorisation of emotion from prosody (Mozziconacci, 2001) is that there is no commonly accepted taxonomy of emotion. Some of the terms that were used by observers in our study seem unlikely targets to be labelled as emotions, emotions or moods – for instance "thoughtful" - while others – such as "thirsty" and "drunk" – are just not affective feelings. In light of there not being an accurate taxonomy of labels and terms that can be regarded as affective states, it is not clear what to make of these labels. For the purposes of this research, we have decided to regard terms such as "thirsty" and "drunk" as responses of "other". What is of interest, however, are the labels for emotional states that were never used to describe the emotions depicted in these movements.

Whissell (1989) authored a dictionary of emotion in which a large number of commonly used terms that refer to emotions and moods are recorded. From 107 of Whissell's labels that describe emotion, our participants had used about 55 to describe the movements presented to them. Some of the remaining labels were synonyms to our labels, but the majority were new terms. As with the labels generated in the current experiment, there were recurrences of emotions through synonyms and from a single reading of the list it was possible to find a number of categories or groups of labels that our participants had not used[25]. These categories were:

Envy/Jealousy
Guilty/Embarrassment/Shame
Defiance
Affectionate
Disgust/Intolerance/Contempt
Puzzlement/Perplexity
Vengefulness
Surprise/Amazement

There currently seems to be no comprehensive taxonomy of terms that can be categorised as affective, so it is not possible to accurately compare the labels used in Experiment 10 with a comprehensive list of labels that depict the whole range of feelings, moods and emotions. Nevertheless, it seems encouraging that observers did not randomly assign feelings to the movements, but rather picked a selection from the terms available to them.

It is still prudent to be cautious, however, since the original emotions were selected to represent a range of states that could be expressed in generic movements. This precluded the use of emotions such surprise and disgust – for instance it is difficult imagine trying to knock on a door in a surprised way. These two emotions are commonly regarded to be basic emotional expressions along with *afraid*, *angry*, *happy* and *sad*. It was possible to conceive of the other 4 basic emotions as being expressed in movement, so they were included in the set of movement stimuli. It may be because of this difficulty to imagine disgust and surprise being expressed in every-day movements, that none of the participants used these terms. However, from a purely subjective perspective, it is possible to imagine making movements to represent some of the other categories that were not represented in the list used by observers. One can easily imagine knocking in a vengeful way (it

[25] Perplexed and ashamed were each used once during the experiment.

would probably seem like *angry*) or in a defiant way. With some thought it also seems possible to drink in either a disgusted (it may appear similar to hesitant) or an affectionate manner (it would resemble *relaxed*).

From these descriptive results it is clear that the labels used to describe movements were similar to the labels of emotions that actors were intending to convey. Additionally, the labels that observers used were an appropriate subset of the labels available to them in the English language. The next step was to try and establish how accurate judgements about emotions has been.

Accuracy of Attributions

So far the analysis of Experiment 10 has been concerned with the semantics of terms that were used by observers. However, it was also important to establish that the attributions observers made about feelings, were related accurately to the emotions actors intended to convey. This would show that the stimuli used in the current thesis – point-light displays of human arm movements – were rich enough to accurately convey information about emotion.

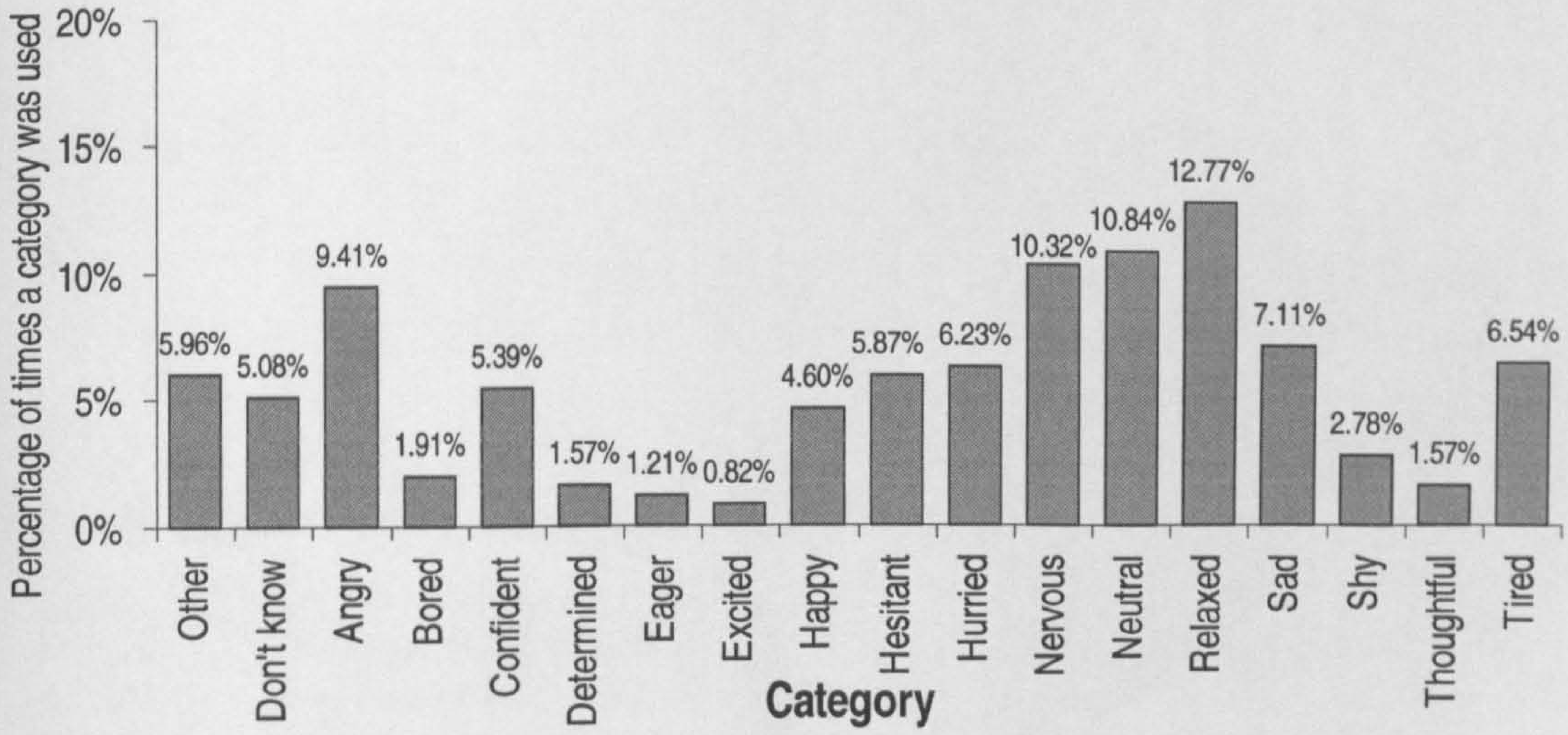

Figure 6.5 Percentage of times that each category was used in the experiment.

The free-categorisation data was re-coded according to the 17 categories described in . There were a number of instances in which observers indicated that they could not find an answer and these responses were coded under "I don't know". The percentage of times each category was used in the research can be seen in . "I don't know" and "Other" accounted for about 11% of all the data. The other 89% of the time, participants used a limited number of terms (124 distinct terms). Many of these terms were synonyms of each other and for this reason grouped easily into 16 categories as described above.

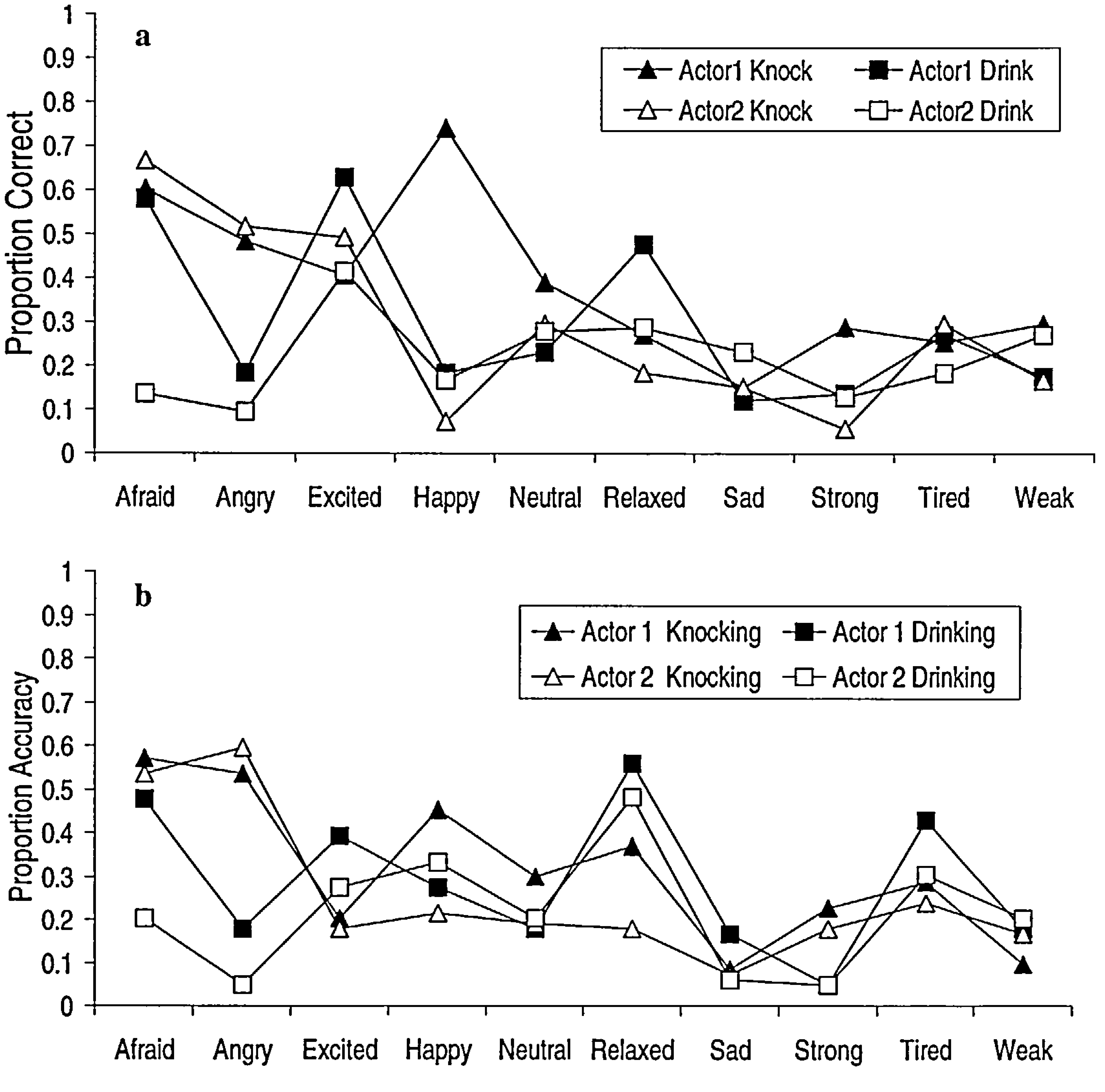

Figure 6.6 a) Replication of figure 2.4 and **b)** accuracy of judgements for the free categorisation task. Means for each movement has been plotted for different actors and actions. See the text for a summary of how accuracy was assessed.

Of the 16 categories, 13 could be paired with the original 10 labels and these 13 categories accounted for 84.6% of the responses observers made to the movements. To establish the accuracy of observer's judgements about movements, the judgements for each movement were coded according to categories that could be paired with the original emotions and those that could not. For instance, in cases where movements with fearful emotion were judges as "hesitant" or "nervous" categories, the judgements were coded as accurate. All other judgements were counted as inaccurate.

For each emotion the accurate judgements were summed and the means for movements with emotion of this is shown in Figure 6.6. From this figure it is clear that as in Experiment 1, there were some individual differences according to which actor was performing a movement and according to what type of movements was being performed. Overall, however, there was an accuracy of 26.7%. This corresponded well with the proportion of correct categorisation from Experiment 1 that was 30% and in a t-test for unequal samples there was no significant difference between the two means [$t(40)=1.86$, ns (two-tailed)].

Some of the same patterns emerged in the data as were observed in Experiment 1. For instance, in comparing Figure 6.6 with Figure 2.4 (pp. 47 and replicated above) it is clear that all *afraid* movements, apart from the drinking movements of Actor2, were well recognised in both experiments. Additionally, Actor1's *angry* movements were well recognised, while those of Actor 2 were not. There are also apparent differences between the results for the two experiments. Most notably for *excited* movements which were very well recognised in Experiment 1 (48% correct), but only recognised with average accuracy in Experiment 10 (26% correct). As noted previously, the label *excited* was used only seldomly in Experiment 10, additionally, there was some ambivalence amongst the judges about what it may mean. This most probably resulted in *excited* movements being identified in with other labels.

In order to assess these differences in the two experiments we calculates a 2 x 2 x 2 x 10 ANOVA with mixed design for unequal sample sizes. Experiments 1 and 10 were the two levels of the independent groups factor. Actors and actions emotions made up the repeated measures factors respectively with 2, 2 and 10 levels.

The results of the ANOVA showed complex high level interactions arising from the individual differences of actors and actions. There were main effects of actor, action and emotion as well as interactions between the first two and emotion and a three-way interaction between actor, action and emotion all at $p< .001$. Since these factors and interactions were discussed extensively in Chapter 2, they have not been further examined here. Instead, we have concentrated on the effects arising from the independent groups factor of Experiment.

There was no main effect of Experiment [$F(1,40)=2.753$, $p=.105$], confirming the finding in the t-test above. There were however, a 2-way interaction between emotion and Experiment [$F(9,360)=2.712$, $p<.005$] and a 4-way interaction between all the factors [$F(9,360)=1.969$, $p<.05$]. A closer examination of the two-way interaction showed that a single simple main effect of experiment or *excited* movements [$F(1,400)=1.844$, $p<.001$] as was predicted from our observations. For the 4-way interaction there was simple main effects of Experiment for the knocking movements of Actor1 with emotions excite, *happy* and *weak* at $p<.05$. For Actor2 there were simple main effects for *excited* knocks and *sad* drinking movements at $p<.005$. In all cases the movements from Experiment 1 were recognised with greater accuracy.

These results suggested very strongly that there was close correspondence between the pattern of results between the two experiments. Similarly, the confusions between different emotions in Experiment 1 seem to have been replicated in Experiment 10. The labelling of emotions were averaged over all the 4 types of movements and have been tabulated in Table 6.2.

Table 6.2 Labelling of Emotions. Only proportion of 5% and above has been included and proportions of 10% and above have been highlighted.

Emotions	Other	I don't know	Angry	Bored	Confident	Determined	Eager	Excited	Happy	Hesitant	Hurried	Nervous	Neutral	Relaxed	Sad	Shy	Thoughtful	Tired
Afraid	0.08	-	0.08	-	-	-	-	-	-	0.19	-	0.26	0.05	0.05	0.05	-	-	-
Angry	0.07	-	0.34	-	0.06	-	-	-	0.05	-	0.14	0.08	-	0.08	-	-	-	-
Excited	-	-	0.21	-	0.05	-	-	-	-	0.07	0.23	0.17	-	-	-	-	-	-
Happy	0.08	0.06	-	-	0.12	-	-	-	0.15	-	0.05	-	0.15	0.17	-	-	-	-
Neutral	-	0.08	-	-	0.05	-	-	-	-	-	-	0.08	0.22	0.23	-	-	-	0.06
Relaxed	0.06	0.07	-	-	0.07	-	-	-	-	-	-	0.05	0.16	0.22	0.06	-	-	0.11
Sad	0.05	0.06	0.05	-	-	-	-	-	-	0.07	-	0.11	0.13	0.17	0.10	-	-	0.09
Strong	0.05	0.07	0.10	-	0.09	-	-	-	0.10	-	0.08	0.05	0.19	0.13	-	-	-	-
Tired	0.08	0.07	-	-	-	-	-	-	-	0.07	-	0.08	0.09	0.13	0.15	-	-	0.16
Weak	0.05	-	-	-	-	-	-	-	-	0.09	-	0.13	-	0.10	0.20	0.10	-	0.16

From Table 6.2 it is clear that *angry* and *excited* were both labelled with words from the angry and hurried categories, *sad tired* and *weak* were most often labelled as tired and sad and *relaxed* and *neutral* were confused with each other.

Furthermore, although some of the categories above were not in the original 10 emotions, it is clear that they adequately describe the movements. For instance, hesitant and nervous both describe aspects that can be related for fearful movements. Similarly the labels under the hurried category describe a negatively stressed state that observers associated with *angry* and *excited* actions. The confident category was made up of labels such as "assertive" and "in charge" that would be associated with strong movements. In Experiment 1 *strong* and *happy* were confused with each other and this also seems to have been case in the current experiment.

The observations regarding the above confusions, along with the proportion correct show that labels were not randomly associated with movements. Rather the movements were labelled appropriately according to the emotions an actor intended to communicate.

Discussion

In Experiment 10 a free response paradigm was used to assess whether stimuli presented to observers in Experiment 1 were rich enough to convey information about emotion. The task for participants was to use a label to describe the feeling or state of an actor as they performed an action. In most cases the labels that were used described emotional states or moods. This is not surprising since the task employed a leading question that would focus observer's attributions to be affective. The data did, however, show that only a subset of the possible emotion terms available to observers were used and that this subset corresponded to the emotions that actors intended to convey. Additionally, for vast majority of the responses participants made, they used terms from categories that corresponded to the original affective terms.

By using the categories that emerged from the range of terms used by observers, it was possible to establish a rate of accuracy at which observers categorised the movements. This rate of accuracy compared well with that from the forced choice categorisation task in Experiment 1. If it had been the case that there were a great difference in the way observers responded to movements in the different tasks, we would have concluded that the movements only accurately convey information about emotion when observers are severely restricted in the way they can interpret the movements. In other words, there would be little ecological validity to the stimuli. This was not the case. Instead, in conjunction with the results from Experiment 8, it is fair to conclude that point-light displays of arm movements resulted in spontaneous and correct interpretations of emotional feelings and states.

Experiment 11 - Word and Movement Dissimilarity.

Two experiments are described that address a second issue arising from previous research. This issue relates to the psychological space discovered in Experiment 1 in which it was important to establish that this psychological space was not an artefact of either the movements or the labels that were presented to observers. Conversely we wanted to show that it reflected accurately the model that an observer would use the categorising of emotion from motion. Dissimilarity data was collected for the movements and labels that had been used in Experiment 2. This data was used to show that for both movements and categories there is a similar psychological space and that the psychological space for emotion from motion is a synthesis of the two.

In experiment 11.1 a psychological space was constructed from the dissimilarity ratings of movements with emotion. For labels of emotion, a psychological space was constructed from dissimilarity ratings in the experiment 11.2. The two spaces were then compared with the prediction that there would be a good mapping between them.

Experiment 11.1 – Dissimilarity between movements

Judgements about the dissimilarity between movements with emotion was collected with the purpose of construction a psychological space to compare with similar spaces from other research.

Methods

Participants

Participants were 15 undergraduate students recruited from the University of Glasgow and were paid for their participation. They were naive to the purpose of the experiment and potential candidates were rejected if they had previously participated in an experiment in which point-lights were employed as stimuli. During the analysis, 4 additional participants were rejected because they did not rate movements that were the same, at or above a pre-determined criterion (see the results section for further information).

Stimuli

Stimuli were the knocking movements of 2 actors that had been used in Experiment 1. It was decided to use only knocking stimuli since there was little difference in the psychological space of knocking movements compared to drinking movements. Problems with drinking movements, however, were that they did not elicit as high a rate of correct identification as knocking movements and for this experiment it was desirous to minimise noise in the data.

Design and Procedure

Stimuli were presented in 2 experimental blocks that each consisted of the 30 movements of a single actor. In each trial a pair of movements were presented one after the other and participants had to rate the similarity of the two on a 100-point scale. They were instructed that a rating of 1 would mean complete dissimilarity between movements. A rating of 100 would indicate that the movements were the same. Participants were told that occasionally the same movements would be presented twice and that these trials were "catch trials" that would be used to assess how well they had performed the task. Additionally participants were told that there was no option by which they could indicate that they were unsure or did not know how to rate the stimuli. Pairs of movements were presented in random order and the order of blocks were also randomised across participants.

Results

Similarity Ratings for Presentations of the same Movements

The scores for pairs of movements that were repetitions of each other were extracted from the data and averaged over all instances and actors. This average was obtained for each observer and compared to a pre-determined criterion. This criterion was set as an average rating of 80 since observers had been warned to be careful in their scoring. Four observers, for whom the average score of the repetitions were below this criterion, were excluded from further analysis.

Psychological Space of Similarity between Movements

Ratings for the remaining 11 observers were averaged across actors and the resultant ratings of similarity were transformed to ratings dissimilarity and input to the INSDCAL Multi-Dimensional Scaling algorithm. The resultant 2-dimensional Euclidean Distance Model had r^2=0.655 and stress=0.255. The model improved with two additional dimensions [the three dimensional solution had r^2=0.645 and stress=0.198 and a 4 dimensional solution had r^2=0.679 and stress=.151], but for the purposes of this experiment, no more than 3 dimensions have been interpreted. The 2-dimensional model have been illustrated in Figure 6.7a along with those from Experiment 11.2 and Experiment 1.

From the 2-dimesnional model, it is clear that the information about speed remains encapsulated in Dimension 1. In Dimension 2, valence had been less clearly represented. The lower left quadrant contains the most negative emotions, but *angry* movements were rated to have a similarity to *happy* and *strong* movements. *Excited* is in the negative part of this less-than-perfect dimension of valence, but so is *relaxed.* In Experiment 3 it was often the case that *excited* movements were labelled as intensely *angry*. These results would indicate that, to a very small extent, the valence dimension from the psychological space of emotion (Experiment 1) is preserved when movements with emotion are rated for similarity. The arousal dimension, however, is very well preserved as can be seen in a correlation between Dimension 1 and movement speed.

Dimension 1 of this dissimilarity space correlated well with kinematics markers for movements speed [r^2= .895 (Average Velocity); r^2 =.918 (Maximum Acceleration); r^2 =.931 (Jerk Index)]. For dimension 2 the correlation with these markers was r^2<.3 in each case. The interesting difference was that when the dimensions of the psychological space were correlated with Distance Moved – a kinematics marker that can be used to estimate the difference in spatial properties between movements in a crude manner – the correlation was better for Dimension 2 that for Dimension 1. In this case for Dimension 1 r^2= 0.107 and for Dimension 2 r^2=0.734. This would suggest that dimension 2 of the psychological space for dissimilarity between movements, encodes information about the spatial configuration of movements.

Stated another way, the observers chiefly categorised movements according to their speed and secondly according to spatial properties in the movements. This finding was supported by the reports of observers in their debriefing sessions. In all cases, both speed and spatial factors were cited as playing an important role in the strategies employed by observers. In most cases speed was also mentioned first (this was the case in 8 out of the 11 cases that were used for this experiment).

These results were almost exactly replicated for dimension 1 and 2 of the 3D psychological space and all results for the correlation can be seen in Table 6.3. Dimension 3 from this psychological space bore no relationship to the kinematics markers that were measured. It also accounted for only 12.85% of the variance that was explained by the model (the three-dimensional model accounted for 64.5% of the variance in the data) compared to 31.24% for Dimension 1 and 20.43% for Dimension 2.

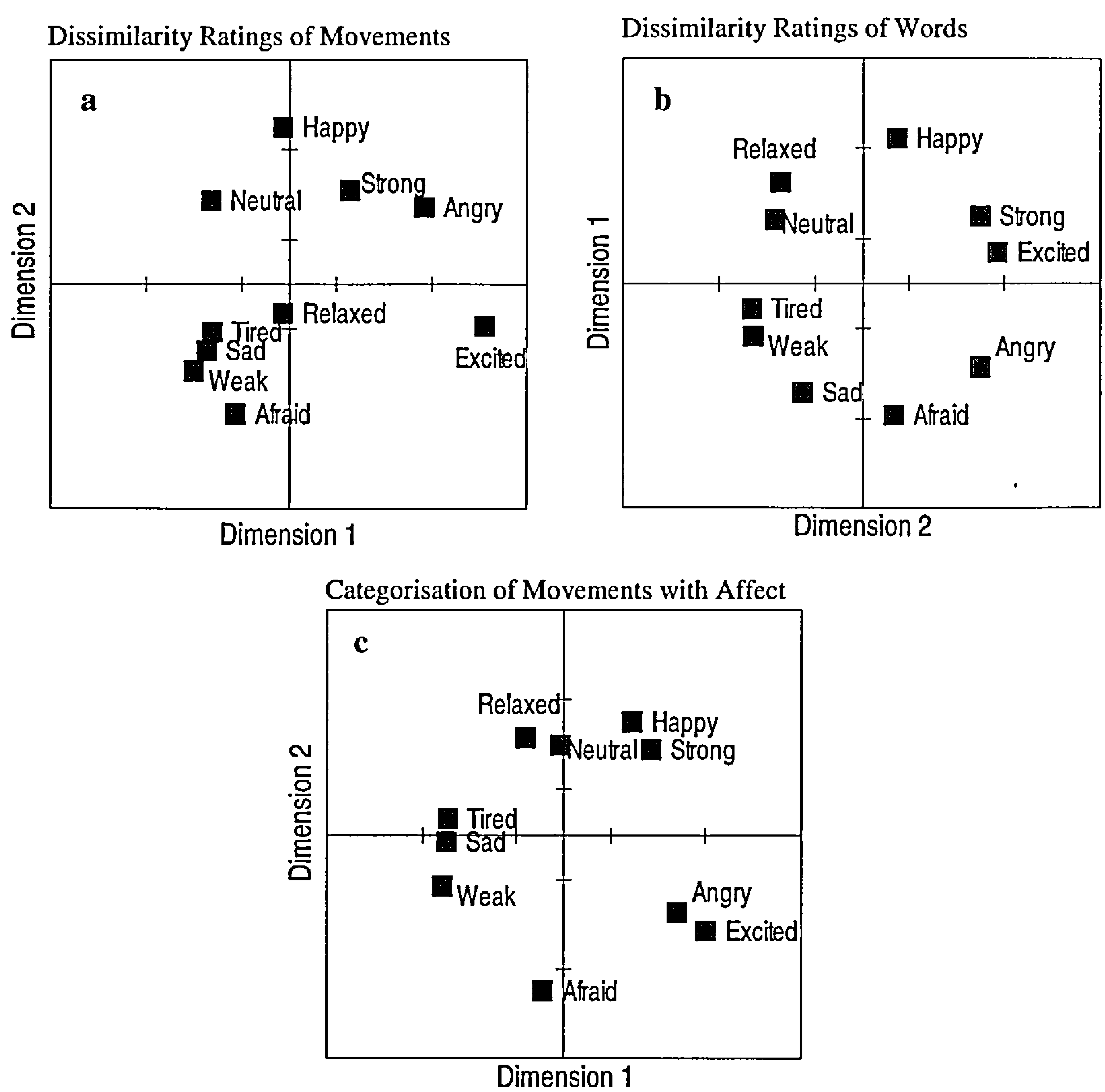

Figure 6.7 Psychological spaces from three different studies. **a)** The psychological space for dissimilarity ratings of movement, **b)** the psychological space for similarity ratings of words and **c)** the psychological space for knocking movements from Experiment 3. Note that the dimensions for **b)** have been reversed in order to have correspondence between the arousal dimensions.

Table 6.3 Correlation (r^2 values) between 3 Dimensions and Kinematics Markers

	D1	D2	D3
Duration	0.92	0.27	0.00
Max Velocity	0.92	0.25	0.00
Average Velocity	0.87	0.35	0.00
Distance Moved	0.08	0.67	0.12
Max Accel	0.87	0.28	0.02
Jerk Index	0.90	0.26	0.02

It is unclear what kind of information may be encapsulated in this third dimension. Beyond using a strategy that employed speed and spatial configuration, there was little agreement between observers about the kind of information they used. Some other features in the movements that were mentioned were the posture and the rhythm of knocks. The posture is a spatial property for the movements, but the rhythm seemed a likely a candidate for this third dimension. Unfortunately this property cannot be measured. An examination of the movement displays (on the website http://www.psy.gla.ac.uk/~frank/Talk_folder/Y2K_Talk_folder/sld015.htm) will show that there are various types of rhythm in the movements. Three major types are easily distinguished – a rather steady rhythm (*neutral* and *relaxed* movements), a rapid rhythm (*excited*, *angry* and *strong* movements) and then there are a group of movements with unusual rhythms (slow and steady: *sad*, *tired* and *weak*; with a pause: *afraid*; and a distinctive rhythm: *happy*).

The configuration of the third dimension, however, does not follow this pattern, see Figure 6.8b. It is very probable that the strategy used by different observers was wholly their own and that the similarity in psychological spaces were the result of a partial common strategy to use speed and spatial configuration.

When using multidimensional scaling it is the normal policy to accept a configuration for which it is possible to explain all the dimensions. In our research this result in a fairly poor fit to the data, however, the 2 dimensions in the configuration can be accounted for by kinematics properties in the movements. Furthermore these properties were consistently reported by observers as being properties on which they based their strategies for the rating the dissimilarity in movements. The most important thing to note from these results is that there are similarities in the psychological space from this experiment and that for the knocking movements in Experiment 1.

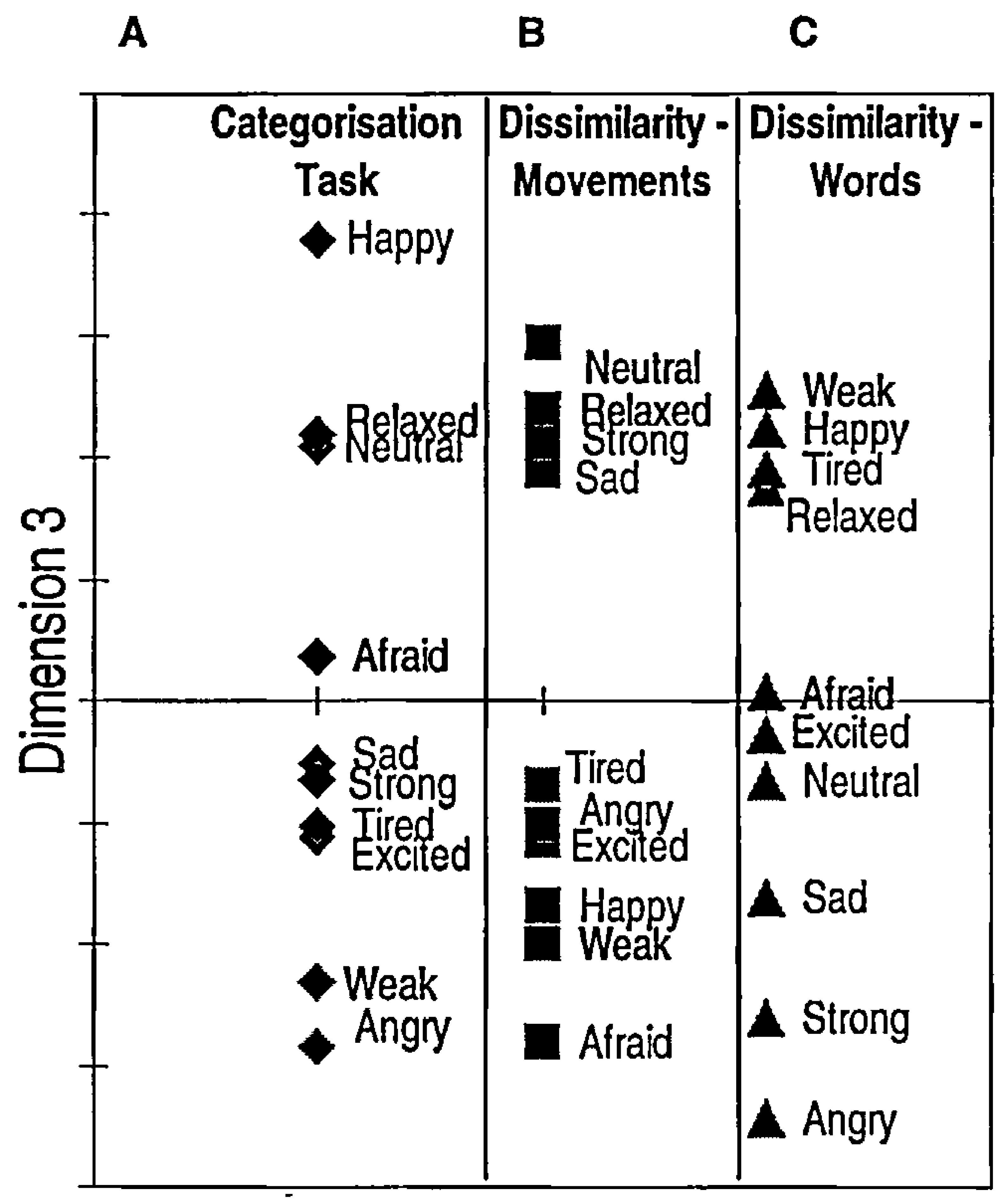

Figure 6.8 Dimension 3 for the three different psychological spaces. A) Dimension three for the categorisation task in Experiment 1 and for ratings of dissimilarity in B) experiment 11.1 and C) for experiment 11.2.

Comparison between psychological spaces of Experiments 1 and 11.1

From the categorisation task in Experiment 1, data for the knocking movements were extracted and a psychological space was constructed. This psychological space was little different than that for both drinking and knocking movements and is illustrated in Figure 6.7c. The next step was to compare this psychological space with that from the dissimilarity ratings of movements. The 2 dimensions from each psychological space were correlated and the results can be seen in the upper left third of Table 6.4. For dimension 1 the correlation was very good, but for the second dimension it was as good (both correlations have been highlighted). This supports the discussion from before in which it was suggested that the dimension of valence had not been well replicated when dissimilarity judgements were made about movements.

Table 6.4 Correlation (r^2 values) of the psychological spaces for three different tasks

		Categorisation - Knocking Movements		Dissimilarity Ratings - Movements	
		D1	D2	D1	D2
Dissimilarity Ratings - Movements	D1	0.782	0.052		
	D2	0.335	0.437		
Dissimilarity Ratings - Words	D1	0.135	0.654	0.032	0.449
	D2	0.815	0.085	0.755	0.118

Experiment 11.2 – Dissimilarity of words

It is clear from Experiment 11.1 that speed and spatial properties of movements played an important role when movements with emotion were categorised in a purely perceptual task. Russell and his colleagues (1980) used a semantic task to make the psychological space of experienced emotion and in this experiment we set out to replicate these results for the emotions that have been used in this research. Similarity ratings were collected for the words that had been used in Experiment 2. These ratings were input to a MDS algorithm and the resultant solution was compared to the psychological spaces from experiments 1 and 11.1.

Methods

Participants

Participants were 40 volunteers recruited from the student population of the university of Glasgow. Since the experiment consisted of a short questionnaire, participants were not paid for their time. Two of the participants were rejected because they rated about 95% of pairs of words to be maximally dissimilar.

Design and Procedure

Stimuli were all the possible pairings of the 10 emotion words that had been used in Experiment 2. The words were *afraid, angry, excited, happy, neutral, relaxed, sad, strong, tired* and *weak*. The pairs of words were arranges in columns on a piece of paper and participants were asked to write down how similar or different they though the meaning of the words in each pair was. They were instructed to use a 100-point scale and the experiment took about10 minutes to complete. The order of word pairs was randomised.

Results

The responses of each participant was input to a spreadsheet and processed to obtain distances that could be used in the INDSCAL Multi-Dimensional Scaling algorithm. The resultant 2-dimenaional Weighted Euclidean Model had r^2=.659 and stress=.212 and the 3-dimensional solution had r^2=.741 and stress=.131. The 2-dimensional solution is shown in Figure 6.7b and in it the arousal and valence dimensions from Russell's model (1980) can be clearly distinguished. For the 3-dimensional solution, dimensions 1 and 2 had a very similar configuration as those from the 2-dimensional solution and most of the variance accounted for by the model were captured in the first two dimensions (Dimension 1, r^2=.275; Dimension 2, r^2=.268; Dimension 3, r^2=.198). The third dimension is illustrated in Figure 6.8c.

In other research that used a facial expressions categorisation task to formalise a psychological space for emotion, this third dimension also emerged (Schlosberg, 1954, 1952). In these cases, it was shown that the dimension of arousal simply had two components – intensity and energy. It is possible that the third dimension on our research also represented intensity, however, no data was collected from participants on how they might attribute intensity to emotions. The two dimensional solution, on the other hand, is much clearer and in the literature on experienced affect, has emerged as the preferred way to describe emotion (Feldman-Barrett & Russell, 1999; Yik, Russell, & Feldman-Barrett, 1999; Watson & Tellegen, 1985; Russell, 1980).

As for the dissimilarity task with movements, the 2-dimenensional solution from Experiment 11.2 was compared to that for the categorisation task from Experiment 2. The results have been tabulated in the lower left part of Table 6.4. This time there was a strong correlation both with dimension 1 and dimension 2, however, it is important to note that the first dimension for words was valence, while the most important dimension for categorisation with movements was arousal. Hence the result is that Dimension 1 correlates with Dimension 2 in each case.

The MDS solutions of dissimilarities of words and movements were also correlated and a similar result was found as above. In this case, however, the correlation between dimension 1 for words and dimension 2 for movements was not as good. This supports the suggestion in Chapter 5 that some feature of the spatial configuration between movements is important and may inform observers of the valence of a movements with emotion. For words on the other hand, the judgements of valence were not restricted to a perceptual feature, but were rather an evaluation in the semantics of a term.

An additional point of note about the psychological spaces was that although the resemblance between them was high, it was never perfect. This may represent noise in the data due to the fact that different participants took place in each study. However, a more it may also represent the fact that the psychological space for either movement or label dissimilarity does not solely account for the representation of emotion from motion. It is more likely that the psychological space for emotion from motion is a synthesis of the representation of labels and movements, along with possible unique features to account for how the two would correspond with each other.

For instance, a feature of Russell's (1980) model of emotion is that it is arranged in a circumplex, a feature that was replicated with our labels. In comparison the movements were not as strongly arranged in a circumplex. For the categorisation task from Experiment 1, however, a roughly circular structure was again observer. A feature of the movements was that arousal was most important, in comparison, for labels valance was the primary way of categorising them. In the psychological space from Experiment 1 it is also the case that arousal was along dimension 1.

In summary, the psychological spaces for word and movements dissimilarities corresponded to each other in some respects. There were, however, differences between the representations and these were reflected in the psychological spaces. The psychological space from Experiment 1 had also incorporated aspects from the representation of both labels and movements, reflecting the characteristic aspect of each.

General Discussion

In the above research, it was decided to adopt a 2-dimensional rather than 3-dimensional solution from the MDS analysis. Primarily this was done because there was no obvious way in which to account for the third dimension in any of the experiments. This does not make the third dimension intrinsically unimportant or uninteresting, however, there was also little correspondence between dimension 3 from different psychological spaces and no obvious way in which to account for it in the stimuli.

In comparison, there was a startling similarity between the 2-dimensional configuration from different experiments. Unfortunately in MDS analysis there seems to be no definitive way of deciding how many dimensions a solution should have. The main requirement is that the stress of the model should be minimised, but also that the dimensions should be explained based on some pre-determined hypothesis.

In the case of the current research, a specific question was being addressed: would similarity ratings of words and movements result in similar configurations and would these configurations bear a similarity to the psychological space from the categorisation task in Experiment 1? We believe that the third dimension in each case of the research was an artefactual dimension connected with the unique circumstances of the different experiments. This is supported by the fact that there was little correspondence between dimension 3 of different tasks.

From an analysis of the 2-dimensional solutions, there was a basic similarity between the strategies that observes seemed to utilise when categorising stimuli with an emotional content. In this strategy arousal and a valence judgement represented the dimensions along which stimuli were described. In the case of moving stimuli, it further emerged that speed was an indicator of physiological arousal and was the most important dimension along which stimuli were categorised.

The motivation for Experiments 11.1 and 11.2 was to establish that one or other aspect of the task did not unduly bias the psychological space for categorisation of affective movements (Experiment 1). This does not seem to have been the case. Instead the multidimensional scaling solutions suggest that there is a common strategy when dealing with stimuli that have and emotional content. This strategy is to categorise movements along two primary dimensions – valence and arousal. The dimensions were orthogonal to and independent from each other and to some extent could be accounted for by temporal and spatial stimulus properties.

Chapter 7 – Discussion and Future Directions

Summary of Experimental Findings

In order to construct a model for the representation of perceived emotion, we employed a judgement task in Experiment 1 and constructed a psychological space through the use of multi-dimensional scaling. In Experiment 2 we measured movement kinematics and related these to the psychological space from the previous experiment. As with representations of experienced emotion and perception of facial expression, we found a 2-dimensional psychological space. The first dimension of this space correlated with movement speed, representing arousal. The second dimension seemed to represent valence.

From Experiment 1 it could be concluded that recognition of emotion from human movements follows a pattern that has been found in previous research in which perception of facial emotion and a person's own emotions were the stimuli. In particular, a low dimensional psychological space accounts for nearly all the variance in the data, suggesting that there are only a few dimensions that humans use to categorise emotion. The two dimensions, along which emotions were categorised, represented arousal and valence. The main difference in our representation space from those in the literature on emotion (Yik, Russell, & Feldman-Barrett, 1999) was the order of importance for these dimensions. For facial expressions and self reported affect, valence has been found to be the first dimension. In our model arousal was most important.

Additionally, in Experiment 2 we found that movement speed – which can be regarded as an indicator to physiological arousal – correlated with the first dimension. This was not the case for the second dimension. Additionally the second dimension did not seem to represent valence very well, however this could only be assessed subjectively. A rotation of the psychological space in order to maximise the fit between dimension 1 and speed also improved the subjective fit of dimension 2 to valence.

In order to test our findings regarding the relationship between speed and the psychological space, we used the computer animation techniques of time warping and morphing to modulate the speed of movements in Experiment 3. The results from Experiment 3 showed that movement duration modulated perception of *angry*, *sad* and *neutral* movements. A linear manipulation of movement durations did, however, result in a non-linear modulation of speed. In the case of *angry* movements this was attributed to spatial feature of the movements and resulted in a modulation of perception of intensity of anger, but not of categorisation. On the other hand modulation of both intensity and categorisation of *neutral* and *sad* movements were achieved through the same manipulation.

In Chapter 5 (Experiments 4-7) the role of the second dimension were further explored along with possible stimulus correlates in spatial features of the movements. The findings from Experiment 4 showed that characteristics of from could be manipulated to disrupt dimension 2 of the psychological space. Moreover the phase or timing relations amongst different markers in the displays proved to be the key to these disruptions. In control experiments (5 and 6) - which examined local features of the movements and a global enhancement of from - we found that the dramatic effects of phase shifted movements could not be accounted for by local or global form features in the movements. For instance, in the "handy man" condition for experiment 5 the local feature of the hand was sufficient to communicate emotion, in the case where this movements was scrambled, we would predict that emotion would still be correctly recognised. However, if the three points representing the hand were scrambled independently there would again be a disruption in the global for features. Since the global cues were such an important feature, it may well be advisable to find a more global measure of kinematics in place of the local wrist and hand kinematics that we used to correlate perception and stimulus features in Experiments 2 and 3. For instance some measure that relates the velocity from all markers in the display or alternatively a measure of the phase relationship amongst points. Certainly the issues regarding the local and global cues need further exploration.

In a finding related to that of Experiments 4-6 we observed in Experiment 7 that it was not necessary for an attribution of "humanness" to be made in order for participants to view the displays as showing emotion.

Finally in Chapter 6 (Experiments 8-11) we addressed some cognitive features in the research. In particular we replicated the findings from Experiment 1 with a free response task (Experiment 10), examined the role of labels and movements independently (Experiment 11) and assessed the immediacy and appropriateness of recognition of emotion from point-light displays of arm movements (Experiments 8 and 9). In Experiment 8 we showed that limited exposure to our point-light arm movements could still result in perception of emotion when *angry* movements were viewed. The results were maximised when *angry* movements were viewed first, suggesting that a diminishing bias by observers to detect biologically relevant stimuli in unfamiliar situations. In Experiment 9 we showed that arm movements compared well with dance movements (Dittrich et al., 1996) as a stimulus for presenting emotion in human movement.

In Experiment 10 we found that when observers were free to label the emotions expressed in point light displays, they picked only a subset of the emotions available in the English language. Furthermore, these labels were appropriate to the emotions actors had intended to convey and participants were accurate to the same extent as in Experiment 1. In Experiment 11 we showed that the representation of emotion from motion was not due either to only labels or movements, but was rather a synthesis of the two.

Interaction between High and Low Level Information

Evidence from the study of biological motion (Thornton, Pinto, & Shiffrar, 1998; Shiffrar & Freyd, 1993, 1990) indicates that the perception of such motion relies not only on specialised bottom-up processes of motion detection (Neri, Morrone, & Burr, 1998; Mather, Radford, & West, 1992) but also on interactions between bottom-up and top-down processes. The recognition of emotion can be regarded as a high-level process since it engages cognitive processes involved in decision making, memory and categorisation.

However, in Experiment 3 we have shown that this process may be modulated through a low-level manipulation of stimulus properties. Additionally, in Experiment 11 it became clear that the most cognitive tasks – rating word meaning for similarity – shared features with the perceptual task of rating similarity between movement stimuli. In the literature on perception of emotion it is also clear that the representation of facial expression (in cases where it is shown to be dimensional) as well as experienced affect, show similar structures to those in our research. This common representation may well serve to bind the signals from different sources in to a cohesive precept.

Support for this comes from cognitive neuroscience where a number of recent findings indicate that visual and motor processing interact whilst an observer watches an action (Gallese et al., 1996; Rizzolatti et al., 1996). In this research a typical finding has been that for both humans and monkeys there is a system of mirror neurones that are activated both when the subjects sees an action and when they make the action. There is also good evidence that this finding extends to perceived and experienced emotions (Adolphs, 2002; Adolphs, Damasio, Tranel, Cooper, & Damasio, 2000; Calder, Keane, Manes, Antoun, & Young, 2000a). Calder et al (2000a) showed additional evidence that such a neural substrate for disgust extended not just to the perception of emotion from faces, but also to perception of emotion from vocal prosody. Common neural mechanisms for perception of such diverse stimuli make it likely that perception of emotion from motion would be similarly represented.

One aspect of high level perception that has not been addressed in the current research is how a pre-manipulated cognitive set of the observer would affect perception. Such a cognitive set might include expectations about the intentions or animacy of an actor. For instance, if the movements of a human were to be acted out by an obviously robot or other non-human model, an observer's judgements may well be modulated – particularly in a free response type task. In the asymmetric transfer effect recorded in results of Experiment 8 there was a clue as to one type of cognitive set that could be manipulated. That is emotion priming, so that in an unusual situation it might be possible to enhance an observer's sensitivity to perceiving emotion in a weak stimulus. If no priming took place, it may well be the case that emotion would not be perceived.

Biologically Relevant Stimuli

For facial expressions there are a small number of emotions that are regarded as basic since they result in accurate recognition across various cultures (Ekman, Friesen, & Ellsworth, 1972; Izard, 1971). However, movements that supply signals as to emotion are unlike facial expressions in that they are not only made by our conspecifics. Instead it is possible that some biologically relevant emotions would be expressed in the movement of various species. Indeed Walk and Homan (1984) cite evidence for an alarm hypothesis by which biologically relevant stimuli, such as fear and anger would be easily communicated between species. On the other hand, socially relevant stimuli such as happiness could have species-specific signals.

From observation of pets such a distinction between biologically and socially relevant signals seem clear. For instance an angry cat adopts a rigid body stance with it's tail extended and bares its fangs and make a hissing noise. An aggressive dog also adopts a rigid stance and bears its fangs while growling. When showing affection or interest a dog would wag it's tail and sniffs or licks a person, on the other hand a cat narrows its eyes and slaps at your throat with a paw - a signal that many people interpret as aggressive. Similarly, it is easy to imagine an angry swarm of bees, but not a swarm of happy bees.

The issue of being able to rapidly recognise aggressive of fearful information becomes more important in an unfamiliar situation. From the results of Experiment 7 it seemed that initially in an unfamiliar situation observers were more sensitive to aggressive displays, but that this sensitivity rapidly dissipated when an initial judgement of the stimuli showed them to not be aggressive.

A related issue regards the role of dangerous stimuli to engage attention. It has been hypothesised that particularly threatening stimuli would capture attention in order for it to be rapidly detect and analysed (Fox, Russo, Bowles, & Dutton, 2001). However, recently it was shown in a visual search task that for still images there was no faster location of threatening stimuli than pleasant stimuli (Tipples, Young, Quinlan, Broks, & Ellis, 2002). In stead, in line with Fox et al (2001), Tipples et al (2002) suggested that dangerous stimuli engage attention in the same way as pleasant stimuli, however, attention is engaged for a longer time by threatening stimuli. It would be interesting to replicate these findings with dynamic stimuli.

Additionally in view of the suggestion that that it would be of adaptive value to accurately and quickly recognise biologically relevant stimuli, it would be interesting to find whether there are invariants in the movements and behaviour of all species when they exhibit such emotions. The findings from Experiment 4 suggest that such an invariant would fall in the form cues from movements. In the case of Experiment 4 spatial features of the movements were disrupted and recognition of angry and afraid were greatly impeded. It would be worthwhile in the future to more carefully modulate the spatial features, particularly phase relations between markers, in order to test these conclusions more rigorously.

Form and Motion

From the research in Chapters 2-6 it seems that human emotional movements are categorised and represented along 2 dimensions and that these dimensions correspond to temporal/speed and spatial/form cues in the stimuli.

To some degree these two types of information act independently (Experiment 4), however, they interact to enhance recognition as shown in the results from Experiment 1 and other experiments where form and motion were presented canonically. As discussed in Chapter 1 (pp. 18), for recognition of facial expression, form (as represented in static displays if facial expression) and motion (as represented in dynamic displays) also appear to act independently as long as the temporal and spatial cues correspond to each other. In related research Pollick, Hill, Calder and Paterson (submitted) have also looked at the separate role played by spatial and temporal facial cues. They found that exaggeration of spatial information enhanced recognition more than exaggeration of motion information, suggesting an advantage for spatial exaggerations over temporal ones.

Results by Kamachi et al (2001) suggest that there are complex interactions between form and motion information in the recognition of facial expressions. From their results it seemed that facial emotions have characteristic temporal signatures and when these are violated, recognition is disadvantaged. Research from the current thesis seems to agree with this. The temporal signature for sadness seemed to be slowness while for angry expressions the signal was best recognised from fast displays. Speed also seemed to be the primary signal used in strategies by our participants to distinguish sad and angry movements.

However, in cases where the judgements extended to movements beyond *angry*, *neutral* and *sad*, it became clear that characteristic motion signals also incorporated a spatial element which seemed to inform decisions about valence.

Categorical vs. Dimensional Account of emotion representation

It seems that the representation of emotions is closely linked to the dimensions of valence and arousal. However, this does not necessarily mean that representation of emotion is dimensional. This might seem like a paradox, however, as has already been discussed in the introduction, there is evidence for both dimensional and categorical representation of emotion. Unfortunately the above research has not added much to this debate since nearly all the analysis undertaken was from a dimensional stance. Some clear evidence for the categorical account of emotion from faces was reported by Young et al. (1997). Their main evidence was that for morphed faces there was not a smooth change in judgement over a category boundary. Rather once some threshold had been reached, there was a sharp change in categorisation to a different label. This is not a finding that was replicated in Experiment 3 as can be seen in **Error! Not a valid link.** below. If perception had been categorical, there would have been a sharp change in the ratings of *angry* and *sad* movements when some threshold for detection of *angry* compared with *sad* were reached. Instead there was a gradual change in the proportion of times a movement was rated as *angry* or *sad*.

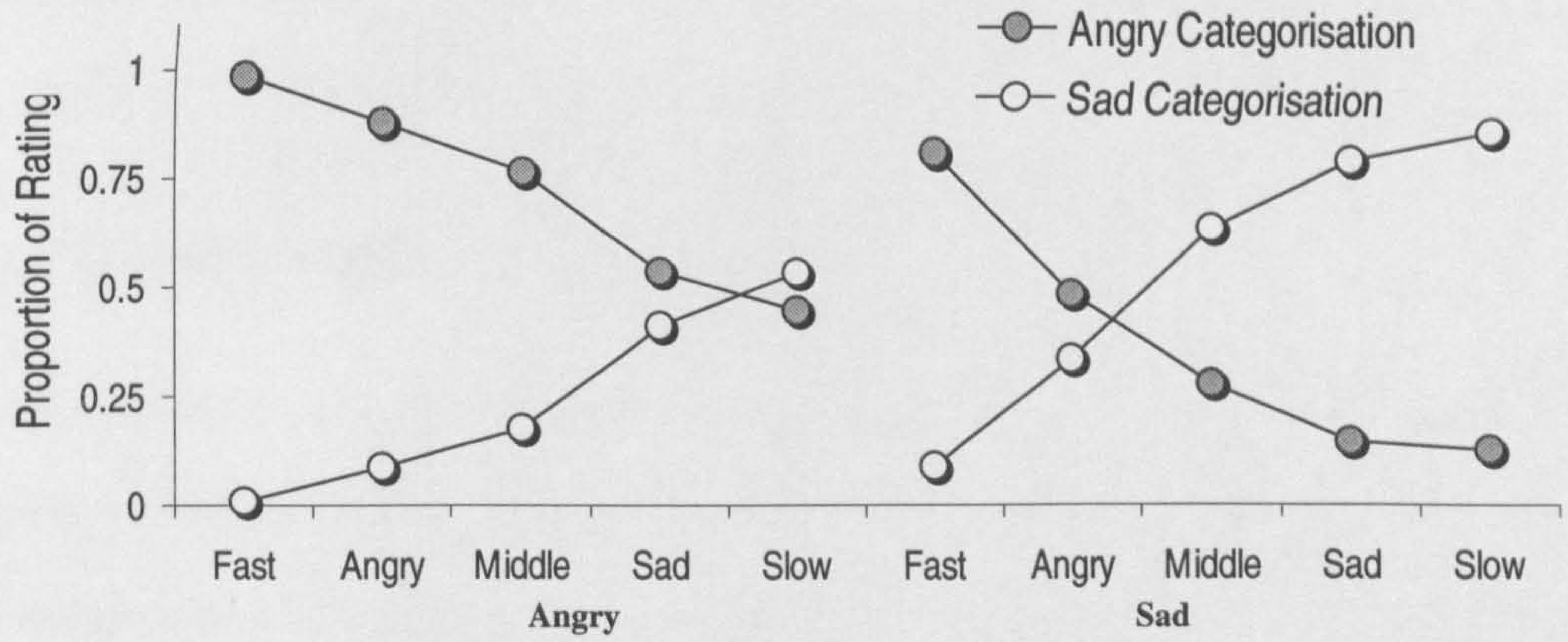

Figure 7.1 Proof for the dimensional representation of movements. The proportion of times *sad* and *angry* movements were rated as *angry* or *sad* is illustrated for data from Experiment 3.

A critical difference between our research and that by Young et al (1997), is that they morphed faces in relation to spatial configuration, while we morphed movements in relation to duration. Since spatial and motion information act more or less independently in the perception of emotion (Frank E. Pollick et al., submitted), it is possible to conclude from these results that these types of information are also represented differently. In other words, it may be the case that valence is represented in categories while arousal is represented along a dimension. A direction for possible future research would be to further explore the dimensional and categorical accounts of emotion perception from the perspective of how the perception of emotion from faces and human movements would be combined.

From another perspective the independence of the two dimensions in the current theses offers a possible future solution to the disparate findings from proponents of the dimensional and categorical accounts. It may be the case that the representation of emotion relies on two parallel components. One of these would be arousal or intensity, processed along a dimension. The other component would be valence and would rely on a categorical judgement into pleasant and unpleasant. For movements the arousal component is dominant, so it may be easy to show dimensional representation of emotion, however, for faces the valence component has proven to be dominant. If it is the case that valence is encoded in form cues while arousal is encoded in stimulus features such as speed or muscular motion, the predominant perceptual feature in static faces would be form. A morphing or other manipulation of faces would in that case take account mostly of this feature and therefore the valence component of emotion representation. It would be an interesting future project to attempt to tease apart these issues, particularly since two such different representational components would also result in different types of processing, possibly relying on independent neural architecture.

Three Directions of Future Research

We have shown that it is possible to recognise a person's emotion from their movements. Additionally, this information is represented along 2 dimensions that reflect stimulus characteristics of the movements. However, both the main experimental results and related observations has shown a number of interesting issues arising that deserve further consideration. Below are three directions of future research that addresses some of the issues that have been raised in the research.

Emotion and Animacy

Emotions are not only attributed to humans, but also to animals and even inanimate objects. In considering this, the question arises as to how perception of emotion and animacy are interrelated. For instance a desk lamp (Lasseter, 1986) can be animated to exhibit emotions, but would an observer have made the same attributions to the lamp on their own desk? Similarly, some things that move quickly and are dangerous in their natural state – such as rocks falling from a cliff – are not categorised normally as having emotions. Other examples along a similar vein regard the innateness to fear whole classes of stimuli – such as spiders or snakes (Cook & Mineka, 1989; Seligman, 1971). Does such a pre-preparedness to fear a certain type of animal also extend to their movements? Indeed, are there some kinds of movement that we find inherently abhorrent and what would distinguish such movements from un-frightening movements? Spider movements seem to be characterised by a smoothly gliding body combined with complex motion from rapidly moving limbs. Spiders also seem to move in bursts, making their movement seem fast and erratic. Could this be a signal to fear? Mori (1982) reports a phenomenon named "The Uncanny Valley" whereby the appearance of a robot that looks and moves in an almost human way can be very unsettling. Something that is exactly *like* a person or very *unlike* a person does not evoke the same feelings of disgust or fear. Thus it would appear that movements which are nearly naturalistic would be candidates for such fear-invoking signals. It seems an interesting problem to further explore the kinds of movement that would elicit fear responses.

In a related vein, it would also be interesting to further explore the alarm hypothesis (Walk & Homan, 1984). Within this hypothesis, it should be a relatively trivial problem to find movements that signal danger from a variety of sources. More socially relevant stimuli would, however, only be found in humans or in objects and animals that can be easily anthropomorphised. Such an observational study could be extended to animating the different animals and humans according to the signals that we discover and this in turn may lead to the discovery of stimulus invariants that inform emotion.

Phase Relationships amongst Joints

The ability to recognise human action from impoverished point-light displays (Johansson, 1973) has occupied a unique position in the study of visual perception and cognition. These stimuli have been used both by psychophysicists studying low-level motion perception as well as social psychologists studying non-verbal communication. What makes the stimuli attractive for all users is their ability to subtract out irrelevant aspects of the human form that might make recognition or detection trivial. For a psychophysicist studying motion perception, there is no down side to this choice of point-light stimulus since what is of intrinsic interest is how motion is used to reconstruct three-dimensional form -- the structure-from-motion problem (Ullman, 1979).

However, for researchers in the domains of visual and social cognition these point-light displays are sometimes problematic since they can appear too abstract to reliably depict person characteristics. Moreover, since in everyday life we view human motion with a convergence of form and motion information the point-light displays are deficient in that they convey form information only through mechanisms of structure-from-motion.

There are now a number of computer animation techniques that allow human motion capture data to be incorporated to animation so that the combination of form and motion may be investigated. We used one of these methods to make humanoid solid body displays in Experiment 6. Solid body models are a well-developed tool from computer animation that provides precise control over the presentation of human form as a volumetric solid. For a given motion, this control over form would allow the presentation of human movement in a variety of shapes from point-light displays to realistic human forms. A researcher will thus be able to independently vary both form and motion.

Recognition of emotion from human movement is ideally suited to study the combination of form and motion since above we have shown that form and motion contribute independently to the recognition of emotion from movement. The first research in a program that combines form and motion would be to further investigate the phase relations that are a part of human movements. Introducing noise into joint angles that make up the hierarchical configuration of a body could upset the timing of movements. If this noise was directed in such a way as to gradually change the spatial configuration of movements in a similar way as we changed the temporal signal, we would have a stimulus that may help to shed light on some of the unanswered questions above.

One possible hypothesis that could be tested with these stimuli is that temporal configurations are represented along a dimension while spatial configurations are represented in categories. It would also be possible to more carefully study the timing relationship between joints of a person and how this relates to judgements about valence.

Combination of faces and Movements

A final direction of future research would be to investigate the combination of facial and movement expression. A basic question we could ask is how does information from movement combine with information from facial expression to produce the perception of emotion. For the perception of emotion from movement it seems that two independent dimensions are involved in the psychological representation of emotion. This is in contrast with perception of emotion from facial expression, which has been characterised as being much more categorical in nature (Etcoff, Ekman, Magee, & Frank, 2000). We anticipate that these differences will lead to complicated interactions in how information from form and motion contribute to the perception of emotion.

To expand the range of stimuli one could develop morphs of both the motion and facial expressions and make hybrid displays that include different faces and motions. We predict that complex interactions will be found depending upon whether we are changing form or motion information and whether we are morphing in a direction defined by valence or activation. In the case where movements fall far apart on the activation axis in the psychological space and close to each other on the valence axis, we expect form information to be surplus to the more diagnostic motion information. In the converse case - where movements are close together on the valence axis, but far apart on the activation axis - the expectation is that form information will dominate. The most interesting cases will be those in which movements are separated more or less equally on both the axes. In these ambiguous cases, we expect the most complex interactions between form and motion and the most valuable results.

In recognition of facial identity, Burton et al (1999) suggest that the form cue of facial features – independent of motion - is the dominant cue. If recognition of emotional expression is similar to identity recognition, form information should also be dominant in the ambiguous case described above. On the other hand, in the perception of intention, movement - independent of form – has been shown to play a major role (Blakemore & Decety, 2001). If recognition of emotional expression is more akin to judgement about intention, motion information may well dominate the process. In either case, the results will shed light on the role played by perception of other human's emotions in human social cognition.

Possible Applications

Besides the theoretical advances in the area of social cognition outlined above, there is a range of applications for which our research is potentially useful. These applications include computer animation, automated recognition of action and affective computing. In the case of computer animation, full rendering of an animation sequence is expensive so there is an interest in the minimal amount of information needed by an observer to reliably preview the success of an animation sequence. Computer animators have looked at this problem within the narrow context of assessing whether a judgement about point-light and solid body models yields an identical result (Hodgins, O'Brien, & Tumblin, 1998). Although these results showed some differences, there was no systematic examination of the relative contributions of form and motion cues as proposed here.

In the case of automated recognition of action there is a growing interest in the ability to extract ever more detailed information from video data (Aggarwal & Cai, 1999) such as the emotion and intention of the actors being viewed. The results of our current and continuing research will be helpful in suggesting possible features for use in automatic recognition systems. A particularly potent application for this research is in the police surveillance of potential crimes (Troscianko, 2002). In this case a reliable invariant that describes emotion would provide the means by such automatic classification of dynamic scenes into potentially criminal or benign events could be automated. Another application is in the area of affective computing. In particular, the ability to remotely sense the emotion of a user is central to the domain of affective computing systems (Picard, 1997) where sensing emotion in a human is used to intelligently direct an interaction between human and computer. Similarly in the growing area of virtual chartrooms, an efficient online rendering of expressions would be of great benefit to visitors.

References

Adolphs, R. (2002). Neural systems for recognizing emotion. *Current Opinion in Neurobiology, 12*(2), 169-177.

Adolphs, R., Damasio, H., Tranel, D., Cooper, G., & Damasio, A. R. (2000). A role for somatosensory cortices in the visual recognition of emotion as revealed by three-dimensional lesion mapping. *Journal of Neuroscience, 20*(7), 2683-2690.

Aggarwal, J. K., & Cai, Q. (1999). Human motion analysis: A review. *Computer Vision and Image Understanding, 73*(3), 428-440.

Aglioti, S., Desouza, J. F. X., & Goodale, M. A. (1995). Size-contrast illusions deceive the eye but not the hand. *Current Biology, 5*(6), 679-685.

Amaya, K., Bruderlin, A., & Calvert, T. (1996). *Emotion from motion.* Proceedings Graphics Interface '95, Toronto, Ont.

Barclay, C. D., Cutting, J. E., & Kozlowski, L. T. (1978). Temporal and spatial factors in gait perception that influence gender recognition. *Perception and Psychophysics, 23*(2), 145-152.

Barrett, L. F., & Russell, A. A. (1999). The structure of current affect: Controversies and emerging consensus. *Current Directions in Psychological Science, 8*(1), 10-14.

Bassili, J. N. (1978). Facial motion in the perception of faces and of emotional expression. *Journal of Experimental Psychology: Human Perception and Performance, 4*(3), 373-379.

Bassili, J. N. (1979). Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face. *Journal of Personality and Social Psychology, 37*(11), 2049-2058.

Bertenthal, B. I., & Pinto, J. (1994). Global processing of biological motions. *Psychological Science, 5*(4), 221-225.

Bingham, G. P. (1987). Kinematic form and scaling - further investigations on the visual-perception of lifted weight. *Journal of Experimental Psychology: Human Perception and Performance, 13*(2), 155-177.

Blakemore, S. J., & Decety, J. (2001). From the perception of action to the understanding of intention. *Nature Reviews Neuroscience, 2*(8), 561-567.

Brownlow, S., Dixon, A. R., Egbert, C. A., & Radcliffe, R. D. (1997). Perception of movement and dancer characteristics from point-light displays of dance. *Psychological Record, 47*(3), 411-421.

Bruce, V., & Valentine, T. (1988). When a nod's as good as a wink: The role of dynamic information in facial recognition. In M. M. Gruneberg & P. E. Morris (Eds.), *Practical aspects of memory: Current research and issues* (Vol. 1, pp. 169-174). New York, NY: John Wiley & Sons.

Bruderlin, A., & Williams, L. (1995). *Motion signal processing.* Computer Graphics Proceedings SIGGRAPH 95, Los Angeles, CA, USA.

Bruner, J. S., & Tagiuri, R. (1954). The perception of people. In G. Lindzey (Ed.), *Handbook of social psychology* (pp. 634-654). Cambridge, Mass: Addison-Wesley.

Burton, A. M., Wilson, S., Cowan, M., & Bruce, V. (1999). Face recognition in poor-quality video: Evidence from security surveillance. *Psychological Science, 10*(3), 243-248.

Calder, A. J., Keane, J., Manes, F., Antoun, N., & Young, A. W. (2000a). Impaired recognition and experience of disgust following brain injury. *Nature Neuroscience, 3*(11), 1077-1078.

Calder, A. J., Rowland, D., Young, A. W., Nimmo-Smith, I., Keane, J., & Perrett, D. I. (2000b). Caricaturing facial expressions. *Cognition, 76*(2), 105-146.

Calder, A. J., Young, A. W., Rowland, D., & Perrett, D. I. (1997). Computer-enhanced emotion in facial expressions. *Proceedings of the Royal Society of London Series B-Biological Sciences, 264*(1383), 919-925.

Christie, F., & Bruce, V. (1998). The role of dynamic information in the recognition of unfamiliar faces. *Memory and Cognition, 26*(4), 780-790.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). Psyscope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25*(2), 257-271.

Cook, m., & Mineka, s. (1989). Observational conditioning of fear to fear-relevant versus fear-irrelevant stimuli in rhesus monkeys. *Journal of abnormal psychology, 98*, 448-459.

Cutting, J. E. (1978). Generation of synthetic male and female walkers through manipulation of a biomechanical invariant. *Perception, 7*(4), 393-405.

Cutting, J. E., & Kozlowski, L. T. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society, 9*(5), 353-356.

Cutting, J. E., Proffitt, D. R., & Kozlowski, L. T. (1978). A biomechanical invariant for gait perception. *Journal of Experimental Psychology: Human Perception and Performance, 4*(3), 357-372.

Dailey, M. N., Cottrell, G. W., Padgett, C., & Adolphs, R. (in press). Empath: A neural network that categorizes facial expressions. *Journal of Cognitive Neuroscience.*

Darwin, C. (1872). *The expression of emotions in man and animals.* London: John Murray.

Decety, J., & Grezes, J. (1999). Neural mechanisms subserving the perception of human actions. *Trends in Cognitive Sciences, 3*(5), 172-178.

Decety, J., Grezes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., et al. (1997). Brain activity during observation of actions - influence of action content and subject's strategy. *Brain, 120*(Pt10), 1763-1777.

Deutscher, J., Blake, A., & Reid, I. (2000). *Articulated body motion capture by annealed particle filtering.* Ieee Computer Soc, Los Alamitos.

Dittrich, W. H. (1993). Action categories and the perception of biological motion. *Perception, 22*(1), 15-22.

Dittrich, W. H., Troscianko, T., Lea, S., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception, 25*(6), 727-738.

EdVEC. (2002). Retrieved July, 2002, from http://www.edvec.ed.ac.uk/index.html

Edwards, K. (1998). The face of time: Temporal cues in facial expressions of emotion. *Psychological Science, 9*(4), 270-276.

Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to russell's mistaken critique. *Psychological Bulletin, 115*(2), 268-287.

Ekman, P., Friesen, W. V., & Ellsworth, P. (1972). *Emotion in the human face: Guidelines for research and an integration of findings*: London: Pergamon Press.

Ekman, P., O'Sullivan, M., & Frank, M. G. (1999). A few can catch a liar. *Psychological Science, 10*(3), 263-266.

Etcoff, N. L., Ekman, P., Magee, J. J., & Frank, M. G. (2000). Lie detection and language comprehension. *Nature, 405*, 139-139.

Feldman-Barrett, L., & Russell, A. A. (1999). The structure of current affect: Controversies and emerging consensus. *Current Directions in Psychological Science, 8*(1), 10-14.

Flash, T., & Hogan, N. (1985). The coordination of arm movements - an experimentally confirmed mathematical-model. *Journal of Neuroscience, 5*(7), 1688-1703.

Fogassi, L., Gallese, V., Buccino, G., Craighero, L., Fadiga, L., & Rizzolatti, G. (2001). Cortical mechanism for the visual guidance of hand grasping movements in the monkey: A reversible inactivation study. *Brain, 124*(3), 571-586.

Fox, E., Russo, R., Bowles, R., & Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *Journal of Experimental Psychology: General, 130*(4), 681-700.

Franz, V. H. (2001). Action does not resist visual illusions. *Trends in Cognitive Sciences, 5*(11), 457-459.

Freeman, J., Avons, S. E., Pearson, D. E., Meddis, R., & IJsselsteijn, W. A. (2000). Using behavioural realism to estimate presence: A study of the utility of postural responses to motion-stimuli. *Presence: Teleoperators and Virtual Environments,, 9*(2), 149-164.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119*, 593-609.

Gavrila, D. M. (1999). The visual analysis of human movement: A survey. *Computer Vision and Image Understanding, 73*(1), 82-98.

Ghafouri, M., & Feldman, A. G. (2001). The timing of control signals underlying fast point-to-point arm movements. *Experimental Brain Research, 137*(3-4), 411-423.

Giese, M. (2002). *Modelling and experiments on the recognition of biological motion and actions*. Retrieved 12 September, 2002, from http://www.uni-tuebingen.de/uni/knv/arl/arl-morph-demo.html

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology*.

Hill, H., & Johnston, A. (2001). Categorizing sex and identity from the biological motion of faces. *Current Biology, 11*(11), 880-885.

Hill, H., & Pollick, F. E. (2000). Exaggerating temporal differences enhances recognition of individuals from point light displays. *Psychological Science, 11*(3), 223-228.

Hodgins, J. K. (1998). Animating human motion. *Scientific American, 278*(3), 64-69.

Hodgins, J. K., O'Brien, J. F., & Tumblin, J. (1998). Perception of human motion with different geometric models. *Ieee Transactions On Visualization and Computer Graphics, 4*(4), 307-316.

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

Izard, C. E. (1994). Innate and universal facial expressions: Evidence from developmental and cross-cultural research. *Psychological Bulletin, 115*(2), 288-299.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics, 14(2)* (Oct), 201-211.

Johansson, G. (1975). Visual motion perception. *Scientific American, 232*(6), 76-88.

Kamachi, M., Bruce, V., Mukaida, S., Gyoba, J., Yoshikawa, S., & Akamatsu, S. (2001). Dynamic properties influence the perception of facial expressions. *Perception, 30*(7), 875-887.

Knight, B., & Johnston, A. (1997). The role of movement in face recognition. *Visual Cognition, 4*(3), 265-273.

Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception and Psychophysics, 21*(6), 575-580.

Kozlowski, L. T., & Cutting, J. E. (1978). Recognizing the gender of walkers from point-lights mounted on ankles: Some second thoughts. *Perception and Psychophysics, 23*(2), 145-152.

Lander, K., & Bruce, V. (2000). Recognizing famous faces: Exploring the benefits of facial motion. *Ecological Psychology, 12*(4), 259-272.

Lander, K., Christie, F., & Bruce, V. (1999). The role of movement in the recognition of famous faces. *Memory and Cognition, 27*(6), 974-985.

Larsen, R. J., & Diener, E. (1992). Promises and problems with the circumplex model of emotion. In Clark, Margaret S (Ed) *Emotion review of personality and social psychology*, no 13 (pp 25 59) Thousand Oaks, CA, US: Sage Publications.

Lasseter, J. (1986). Luxo jr: Pixar.

Lasseter, J., & Rafael, S. (1987). Principles of traditional animation applied to 3d computer animation. *Computer Graphics, 4*(July), 35-44.

Marey, E. J. (1895). *Movement: The results and possibilities of photography* (E. Prichard, Trans.). London: Heinemann.

Mather, G., & Murdoch, L. (1994). Gender discrimination in biological motion displays based on dynamic cues. *Proceedings of the Royal Society of London Series B-Biological Sciences, 258*(1353), 273-279.

Mather, G., Radford, K., & West, S. (1992). Low-level visual processing of biological motion. *Proceedings of the Royal Society of London Series B-Biological Sciences, 249*(1325), 149-155.

Mather, G., & West, S. (1993). Recognition of animal locomotion from dynamic point-light displays. *Perception, 22*(7), 759-766.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences, 6*(6), 255-260.

Michotte, A. (1963). *The perception of causality*. London : Methuen.

Mori, M. (1982). *The buddah in the robot: A robot engineer's thoughts on science and religion* (C. S. Terry, Trans.) Charles E. Tuttle Co., Inc.

Morrison, D. J., & Schyns, P. G. (2001). Usage of spatial scales for the categorization of faces, objects, and scenes. *Psychonomic Bulletin & Review, 8*(3), 454-469.

Mozziconacci, S. J. L. (2001). *Emotion and attitude conveyed in speech by means of prosody.* 2nd Workshop on Attitude, Personality and Emotions in User-Adapted Interaction, Sonthofen, Germany.

Muybridge, E. (1893). *Descriptive zoopraxography; or, the science of animal locomotion made popular* (1 ed.). Philadelphia: University of Pennsylvania.

Muybridge, E. (1899). *Animals in motion: An electrophotographic investigation of consecutive phases of animal progressive movements, commenced 1872, completed 1885* (1 ed.). London: Chapman and Hall.

Neri, P., Morrone, M. C., & Burr, D. C. (1998). Seeing biological motion. *Nature, 395*(6705), 894-896.

O'Toole, A. J., Roark, D. A., & Abdi, H. (2002). Recognizing moving faces: A psychological and neural synthesis. *Trends in Cognitive Sciences, 6*(6), 261-266.

Picard, R. (1997). *Affective computing.* Cambridge, MA:: The MIT Press.

Pike, G. E., Kemp, R. I., Towell, N. A., & Phillips, K. C. (1997). Recognizing moving faces: The relative contribution of motion and perspective view information. *Visual Cognition, 4*(4), 409-437.

Pollick, F. E., Chizk, C., Hager-Ross, C., & Hayhoe, M. (2001). Implicit accuracy constraints in two-fingered grasps of virtual objects with haptic feedback. In *Haptic human-computer interaction, proceedings* (Vol. 2058, pp. 98-107). Berlin: SPRINGER-VERLAG BERLIN.

Pollick, F. E., Hill, H., Calder, A., & Paterson, H. (submitted). Recognizing facial expressions from spatially and temoprally modified movements. *Perception.*

Pollick, F. E., Hill, H., Folwaczny, A., Gray, G., Holmes, R., & Love, S. (1999). Contributions of facial and head motion in the interpretation of expression. *Perception, 28*(Supplement).

Pollick, F. E., & Sapiro, G. (1997). Constant affine velocity predicts the 1/3 power law of planar motion perception and generation. *Vision Research, 37*(3), 347-353.

Poulton, E. C., & Freeman, P. R. (1966). Unwanted asymmetrical transfer effects with balanced experimental designs. *Psychological Bulletin, 66*, 1-8.

Proffitt, D.-R., Bertenthal, B.-I., & Roberts, R.-J. (1984). The role of occlusion in reducing multistability in moving point-light displays. *Perception and Psychophysics, 36*(4), 315-323.

Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (in press). From mirror neurons to imitation: Facts and speculations. In W. Prinz (Ed.), *The imitative mind: Development, evolution, and brain bases.* Cambridge: Cambridge University Press.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research, 3*(2), 131-141.

Runeson, S. (1994). Perception of biological motion: The ksd-principle and the implications of a distal versus proximal approach. In G. Jansson, S. S. Bergstrom & W. Epstein (Eds.), *Perceiving events and objects* (1 ed., pp. 383-404). Hillsdale, NJ, USA: Lawrence Erlbaum Associates, Inc.

Runeson, S., & Frykholm, G. (1981). Visual perception of lifted weight. *Journal of Experimental Psychology: Human Perception and Performance, 7*(4), 733-740.

Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General, 112*(4), 585-615.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(1-sup-6), 1161-1178.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expressions? A review of the cross-cultural studies. *Psychological Bulletin, 115*(1), 102-141.

Schlosberg, H. (1952). The description of facial expressions in terms of two dimensions. *Journal of Experimental Psychology, 44*, 229-237.

Schlosberg, H. (1954). Three dimensions of emotion. *Psychological Review, 61*, 81-88.

Seligman, M. (1971). Phobias and preparedness. *Behavior Therapy, 2*, 307-320.

Shaw, R. E., McIntyre, M., & Mace, W. M. (1974). The role of symmetry in event perception. In R. B. MacLeod & H. L. Pick (Eds.), *Perception: Essays in honour of JJ Gibson*. Ithaka, NY: Cornell University Press.

Shiffrar, M., & Freyd, J. J. (1990). Apparent motion of the human body. *Psychological Science, 1*(4), 257-264.

Shiffrar, M., & Freyd, J. J. (1993). Timing and apparent motion path choice with human body photographs. *Psychological Science, 4*(6), 379-384.

Sumi, S. (1984). Upside-down presentation of the Johansson moving light-spot pattern. *Perception, 13*(3), 283-286.

Sumi, S. (2000). Perception of point-light walker produced by eight lights attached to the back of the walker. *Swiss Journal of Psychology, 59*(2), 126-132.

Tagiuri, R. (1969). Person perception. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (2 ed., Vol. 3, pp. 395-449). London: Addison-Wesley.

Takehara, T., & Suzuki, N. (2001). Robustness of the two-dimensional structure of recognition of facial expression: Evidence under different intensities of emotionality. *Perceptual and Motor Skills, 93*(3), 739-753.

Thayer, R. E. (1996). *The origin of everyday moods: Managing energy, tension, and stress*. New York, NY, US: Oxford University Press.

Thornton, I. M., & Kourtzi, Z. (2002). A matching advantage for dynamic human faces. *Perception, 31*(1), 113-132.

Thornton, I. M., Pinto, J., & Shiffrar, M. (1998). The visual perception of human locomotion. *Cognitive Neuropsychology, 15*(6-8), 535-552.

Tipples, J., Young, A. W., Quinlan, P., Broks, P., & Ellis, A. W. (2002). Searching for threat. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 55*(3), 1007-1026.

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

Tomkins, S. S., & Izard, C. E. (Eds.). (1965). *Affect, cognition, and personality: Empirical studies.*

Troje, N. F. (2002). *Analysis and synthesis of biological motion patterns*. Retrieved 13 September, 2002, from http://www.bml.psy.ruhr-uni-bochum.de/currents.htm#Project1.

Troscianko, T. (2002). Personal communication.

Ude, A., Riley, M., & Atkeson, C. G. (2000). *Automatic generation of kinematic models for the conversion of human motion capture data into humanoid robot motion.* Proceedings IEEE-RAS International Conference on Humanoid Robotics (Humanoids-2000), Cambridge, Massachusetts.

Walk, R. D., & Homan, C. P. (1984). Emotion and dance in dynamic light displays. *Bulletin of the Psychonomic Society, 22*(5), 437-440.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98*(2), 219-235.

Whissell, C. M. (1989). The dictionary of affect in language. In R. Plutchik & H. Kellerman (Eds.), *Emotion: Theory, research and experience* (Vol. 4, pp. 113-131). New York: Academic Press.

White, M. (1999). Representation of facial expressions of emotion. *American Journal of Psychology, 112*(3), 371-381.

Wolff, W. (1943). *The expression of personality; experimental depth psychology.* New York: Harper.

Yang, G. Z., Dempere-Marco, L., Hu, X. P., & Rowe, A. (2002). Visual search: Psychophysical models and practical applications. *Image and Vision Computing, 20*(4), 273-287.

Yik, M. S. M., Russell, J. A., & Feldman-Barrett, L. (1999). Structure of self-reported current affect: Integration and beyond. *Journal of Personality and Social Psychology, 77*(3), 600-619.

Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., & Perrett, D. I. (1997). Facial expression megamix: Tests of dimensional and category accounts of emotion recognition. *Cognition, 63*(3), 271-313.

Appendix I - Scripts to set the scene for movements with emotion

Below are the scripts from Amaya, Bruderlin and Calvert (1996) that were used to set the scene for movements with emotion.

Each of the other scripts has been presented as three or four parts. Firstly is presented the part of the scrip common to all actions (for some scripts there were no commonalties between actions). Secondly the parts unique to knocking movements are presented followed by parts unique to drinking movements. Finally, script changes for lifting movements are presented for neutral, angry, happy and sad movements are shown.

Neutral

It is a sunny Saturday morning. You are well rested, you just had a relaxed breakfast and now you are on your way back from the corner store where you picked up a newspaper. You walk up to the front door of your house and knock three times for your spouse to come down and open it.

Imagine yourself sitting at your dining room table on a Saturday morning. You are well rested, you just had breakfast, you poured yourself a glass of water. It is a sunny day and you are thinking about what you are going to do today. You reach for the glass of water and drink from it.

Imagine yourself sitting at your dining room table on a Saturday morning. You are well rested, you just had breakfast. It is a sunny day and you are thinking about what you are going to do today. You reach for a bottle on the table and lift it out of the way.

Angry

Today you got up a bit late, so you had to rush to get ready. Then on the way to work, a cop flags you down and gives you a speeding ticket, although you were just keeping up with traffic. You finally get to work where a note is waiting for you denying your request for having Friday off ...

... Now you are steaming are ready to tell it to your boss. You walk to his office and knock on his door three times.

... Sitting down at your desk and flaring up, you reach for the glass in front of you and drink from it.

... Sitting down at your desk and flaring up, you reach for a bottle on your desk and lift it out of the way.

Happy

It's Friday evening, you are sitting at the dining room table leafing through you favourite magazine. You feel great, because you finished your project at work today on time. Your boss was very pleased, he complimented you on it and said you could take Monday off...

... You just talked to your friend and neighbour on the phone who invited you over to watch a video. You joyously walk next door and knock on the door three times.

... You just talked you your best friend on the phone who is going to come over shortly to watch a video. You joyously reach for the glass of water in front of you and drink from it.

... You just talked you your best friend on the phone who is going to come over shortly to watch a video. You joyously reach for a bottle on the table in front of you and lift it out of the way.

Sad

You are sitting at the dining room table trying to finish up dinner, but you are not hungry at all. You don't feel like eating, your stomach is heavy and seems to be bearing all of your body's inner activity and weight. You just move the food from one spot on your plate to another. Half an hour ago, you received a telephone call that your best friend just died in a car accident...

... You know that you should tell your neighbour, who also knew your friend. Deeply grieving, you walk over and knock three times.

... Deeply grieving, you reach or the glass of water and drink from it.

... Deeply Grieving, you reach for a bottle on the table and lift it out of the way.

Fearful

As you got to work a minute ago, the secretary told you that you boss wanted to talk to you immediately. You know what it is about; it's the third time now you messed it up, and a customer was probably complaining about you again. Your boss already reprimanded you last time and warned you that there could be severe consequences if it happened again. Standing in front of his office now and feeling the sweat on your back, you knock three times.

You are about to do your driving test. This is the second time, after you failed the first attempt by going through a red light. Your mind is restless and running at full speed: what kind of examiner are you going to get this time? Is he or she going to take you though any roads you haven't been before? What are my friends going to say if I fail again? Your throat feels dry as you reach for the glass of water in front of you and drink from it.

Tired

You just got home from a long day of work. Since your car is in the shop, you had to take the bus, and it is raining. After warming up some leftover dinner from yesterday and eating it, you are sitting at the table, exhausted, and all you want to do is going to bed. But you know well you first have to do the dishes and finish up some paperwork...

... You also know you promised to return that book you neighbour lent you tonight. You walk over to his house with the book in your hand, and knock on the door three times.

... You reach for the glass of water in front of you and drink from it.

Strong

Everything is going well these days: Work is great after your promotion a few weeks ago; your relationship is balanced, rewarding and thriving; your body feels great since you joined the gym and started to work out regularly...

... It is Saturday morning now and you just arrived at your friend's place to go jogging. You feel good as you walk up to his house from your car. You knock three times.

...You just had a light, healthy lunch: a mixed salad and half a chicken sandwich. Sitting at the table and feeling your power, you reach for the glass of water in front of you and drink fron it.

Weak

You are just getting over a terrible cold you have had for 3 long weeks. You lost weight, and you appetite hasn't come back yet. You feel fragile and feeble since you haven't worked out for ages. Sometimes, if you get up too fast, your head is still spinning...

... Your neighbour just called to have you over for chicken soup. As you walk over there, you notice that your body is aching quite a bit. Standing in front of the door, you knock three times.

... You are sitting at the table trying to finish off a bowl of chicken soup, while you reach for the glass in front of you and drink from it.

Excited

You are thrilled because today is the day when your parents are flying in for a visit. It's been at least a year since you saw them last. It will be their first time to Vancouver. You prepared everything for their arrival; cleaned up the house, rented a car for them, arranged for a boat cruise, took a week's vacation from work...

... Now you are in your car on the way to the airport, and as usual you are a bit late. You promised your Aunt Theresa to pick her up. As you get to her house, you get more and ore excited. You quickly get out of your car and hastily walk up to her front door. You knock three times.

... Now you are waiting at the airport for their arrival. You got there well ahead of time. The monitors show that their flight is delayed by half an hour. With every minute you are feeling more excited. You go to the cafeteria, where you get a coffee and a glass of water. You are sitting down and reach for the glass in front of you and drink from it.

Relaxed

An hour ago, you attended a fitness class with a friend of yours. Afterwards you showered went to the sauna and whirlpool. As you are sitting now on you terrace at your friend's home, you are feeling fresh, calm and content. You inhale deeply and exhale through your mouth a few times, enjoying the warm soothing temperature of a lovely afternoon sun...

... You decide to get another glass of fruit juice; the back door is closed, you walk around the front door and knock three times.

... You reach for the glass of water in front of you and drink from it.

Appendix II - Questions for a Typical Debriefing Session

Below are the questions from the debriefing session for Experiment 1. This represents the questions in a typical debriefing session and questions 1, 4 and 5 always appeared in some form in all experiments.

1. How did you decide which category the motions fitted into?

   Did you use any particular strategy?

2. How many actors did you think there were?

3. In your opinion, were they were male of female?

4. Did you use any particular context when looking at the displays -Such as a situation or a story?

5. Did the categories seem correct to you

Appendix III – Differences in Kinematics for Emotions

Movement Duration

In summary, findings for movement duration were very similar to those from velocity, with significant main effects and interactions at all levels. The basic finding was that different actors and different actions resulted in differences between movement kinematics, leading to complex interactions with emotion. However, when the duration for actors and actions were normalised, there emerged a consistent pattern in the data across all 4 the movement types. The output form Expertstats for the ANOVA and simple main effects for the three-way interaction has been reproduced below along with a graph that shows the three way interaction and another for the z-scores of duration. Finally there is a table that shows where Tukey tests of the HSD's between emotions were significant.

Analysis of Variance Summary Table

Data from Movement Duration
Mixed Design (alias Split Plot)

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (actor)	48.300	1	48.300	693.715	0.0000
B (action)	248.098	1	248.098	2961.659	0.0000
C (affect)	71.437	9	7.937	156.994	0.0000
AB	1.210	1	1.210	14.441	0.0191
AC	3.965	9	0.441	8.714	0.0000
BC	12.000	9	1.333	20.318	0.0000
ABC	6.262	9	0.696	10.603	0.0000
Between Error	0.279	4	0.070		
(Error BxS)	0.335	4	0.084		
(Error CxS)	1.820	36	0.051		
(Error BCxS)	2.362	36	0.066		

Simple Simple Main Effects ABC

actor at	Source of Variation	Sum of Squares	df	Mean Squares	F	p
knock	afraid	5.549	1	5.549	72.347	0.0000
knock	angry	1.242	1	1.242	16.195	0.0038
knock	excited	0.976	1	0.976	12.726	0.0073
knock	happy	1.739	1	1.739	22.671	0.0014
knock	neutral	3.904	1	3.904	50.905	0.0001
knock	relaxed	5.684	1	5.684	74.113	0.0000
knock	sad	4.664	1	4.664	60.810	0.0001
knock	strong	2.747	1	2.747	35.819	0.0003
knock	tired	2.884	1	2.884	37.606	0.0003
knock	weak	5.510	1	5.510	71.846	0.0000
drink	afraid	9.127	1	9.127	118.995	0.0000
drink	angry	3.656	1	3.656	47.662	0.0001
drink	excited	1.056	1	1.056	13.763	0.0060

drink	happy	4.029	1	4.029	52.530	0.0001
drink	neutral	0.120	1	0.120	1.570	0.2456
drink	relaxed	0.161	1	0.161	2.101	0.1852
drink	sad	1.056	1	1.056	13.763	0.0060
drink	strong	4.002	1	4.002	52.174	0.0001
drink	tired	1.628	1	1.628	21.221	0.0017
drink	weak	0.004	1	0.004	0.049	0.8305
Error Term		0.614	8	0.077		
action at						
actor1	afraid	3.178	1	3.178	37.937	0.0035
actor1	angry	6.080	1	6.080	72.583	0.0010
actor1	excited	7.775	1	7.775	92.811	0.0006
actor1	happy	7.866	1	7.866	93.902	0.0006
actor1	neutral	28.398	1	28.398	339.003	0.0001
actor1	relaxed	29.733	1	29.733	354.941	0.0000
actor1	sad	16.094	1	16.094	192.120	0.0002
actor1	strong	10.305	1	10.305	123.019	0.0004
actor1	tired	20.609	1	20.609	246.020	0.0001
actor1	weak	28.037	1	28.037	334.688	0.0001
actor2	afraid	5.993	1	5.993	71.545	0.0011
actor2	angry	10.649	1	10.649	127.121	0.0004
actor2	excited	7.996	1	7.996	95.457	0.0006
actor2	happy	12.203	1	12.203	145.670	0.0003
actor2	neutral	13.691	1	13.691	163.432	0.0002
actor2	relaxed	12.042	1	12.042	143.747	0.0003
actor2	sad	8.292	1	8.292	98.980	0.0006
actor2	strong	12.625	1	12.625	150.706	0.0003
actor2	tired	16.951	1	16.951	202.354	0.0001
actor2	weak	9.053	1	9.053	108.068	0.0005
Error Term		0.335	4	0.084		
affect at						
actor1	knock	5.177	9	0.575	11.378	0.0000
actor1	drink	48.392	9	5.377	106.350	0.0000
actor2	knock	18.552	9	2.061	40.770	0.0000
actor2	drink	21.543	9	2.394	47.345	0.0000
Error Term	1.820	36	0.051			

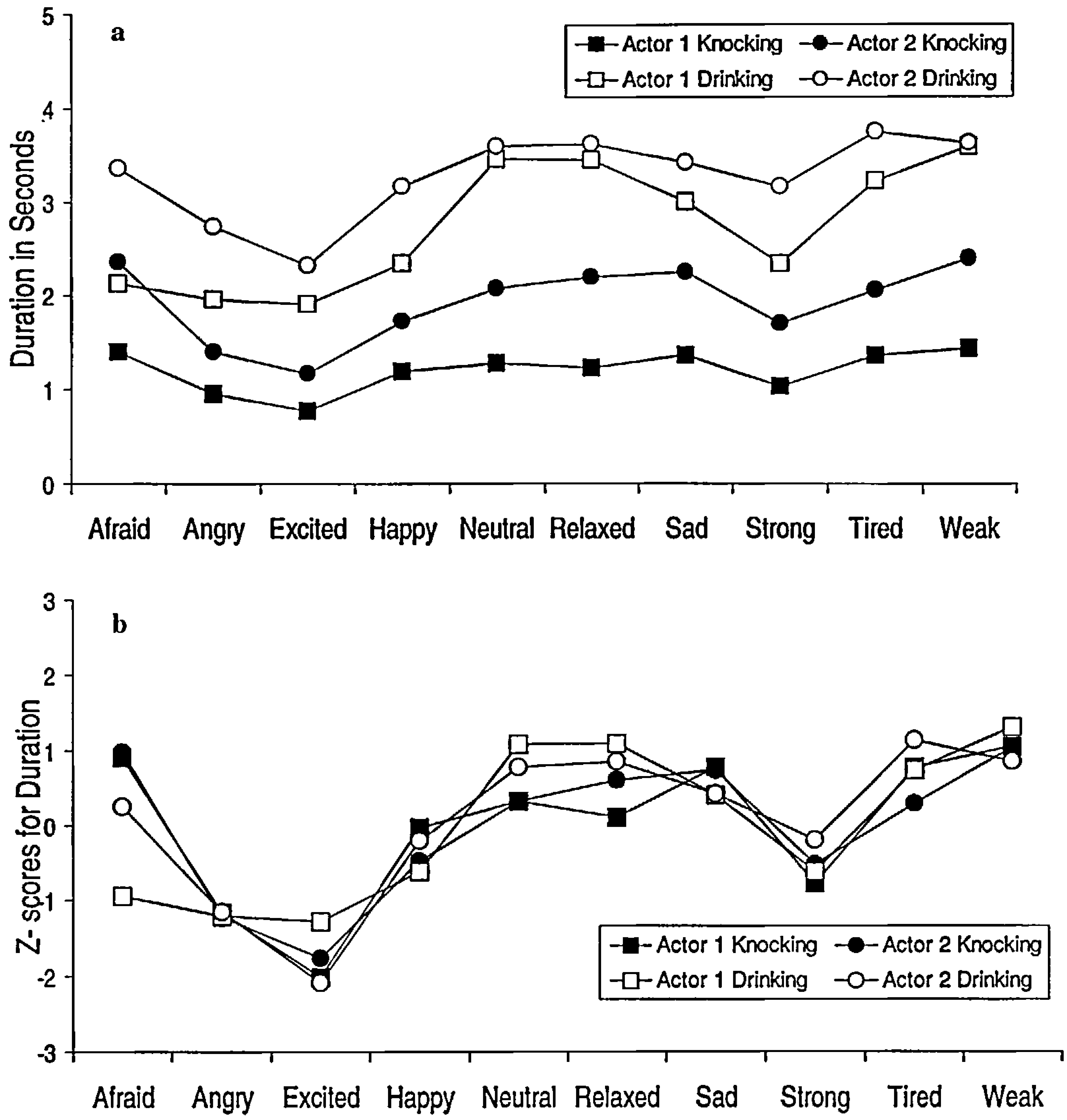

Figure III.1 Movement duration for different actors and actions a) according to the time a movements took to perform and b) as z-scores.

Table III.1 Results of post hoc Tukey tests of differences between affects at each level of actor and action for movement duration. Results are tabulated as levels of significance according to action and actor.

	Actor 1									
	Afraid	Angry	Excited	Happy	Neutral	Relaxed	Sad	Strong	Tired	Weak
Afraid										
Angry	**									
Excited	***	ns								
Happy	ns	ns	*							
Neutral	ns	ns	***	ns						
Relaxed	ns	ns	**	ns	ns					
Sad	ns	*	***	ns	ns	ns				
Strong	*	ns	ns	ns	ns	ns	ns			
Tired	ns	*	***	ns	ns	ns	ns	ns		
Weak	ns	**	***	ns	ns	ns	ns	*	ns	
Afraid										
Angry	ns									
Excited	ns	ns								
Happy	ns	*	**							
Neutral	***	***	***	***						
Relaxed	***	***	***	***	ns					
Sad	***	***	***	***	**	**				
Strong	ns	*	**	ns	***	***	***			
Tired	***	***	***	***	ns	ns	ns	***		
Weak	***	***	***	***	ns	ns	***	***	*	

	Actor 2									
	Afraid	Angry	Excited	Happy	Neutral	Relaxed	Sad	Strong	Tired	Weak
Afraid										
Angry	***									
Excited	***	ns								
Happy	***	ns	***							
Neutral	ns	***	***	*						
Relaxed	ns	***	***	**	ns					
Sad	ns	***	***	***	ns	ns				
Strong	***	ns	***	ns	*	***	***			
Tired	ns	***	***	ns	ns	ns	ns	*		
Weak	ns	***	***	***	ns	ns	ns	***	ns	
Afraid										
Angry	***									
Excited	***	*								
Happy	ns	*	***							
Neutral	ns	***	***	*						
Relaxed	ns	***	***	**	ns					
Sad	ns	***	***	ns	ns	ns				
Strong	ns	*	***	ns	*	**	ns			
Tired	*	***	***	***	ns	ns	ns	***		
Weak	ns	***	***	**	ns	ns	ns	**	ns	

Levels of α: *** α=.001, ** α=.005, * α=.05
ns = non-significant

Peak Velocity

In summary, findings for Peak velocity were very similar to those from velocity, with significant main effects and interactions at most levels. The major difference was that there was no three-way interaction. For all significant two way interactions there were also significant main effects at p<.001. The basic finding was that different actors and different actions resulted in differences between movement kinematics, leading to interactions with emotion. However, when the duration for actors and actions were normalised, there emerged a consistent pattern in the data across all 4 of the movement types. The output form Expertstats for the ANOVA (of the same design described in Chapter 4, page) has been reproduced below along with a graph that shows the two-way interactions for affect x actor and affect x action. As second graph shows the z-scores of peak velocity.

Analysis of Variance Summary Table

Data from Peak Velocity
Mixed Design (alias Split Plot)

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (actor)	11413370.184	1	11413370.184	998.090	0.0000
B (action)	33653375.528	1	33653375.528	1923.844	0.0000
C (affect)	12589479.546	9	1398831.061	82.089	0.0000
AB	4648129.880	1	4648129.880	265.717	0.0001
AC	577587.012	9	64176.335	3.766	0.0020
BC	3931458.386	9	436828.710	18.535	0.0000
ABC	268062.306	9	29784.701	1.264	0.2896
Between Error	45740.836	4	11435.209		
(Error BxS)	69971.103	4	17492.776		
(Error CxS)	613457.169	36	17040.477		
(Error BCxS)	848418.957	36	23567.193		

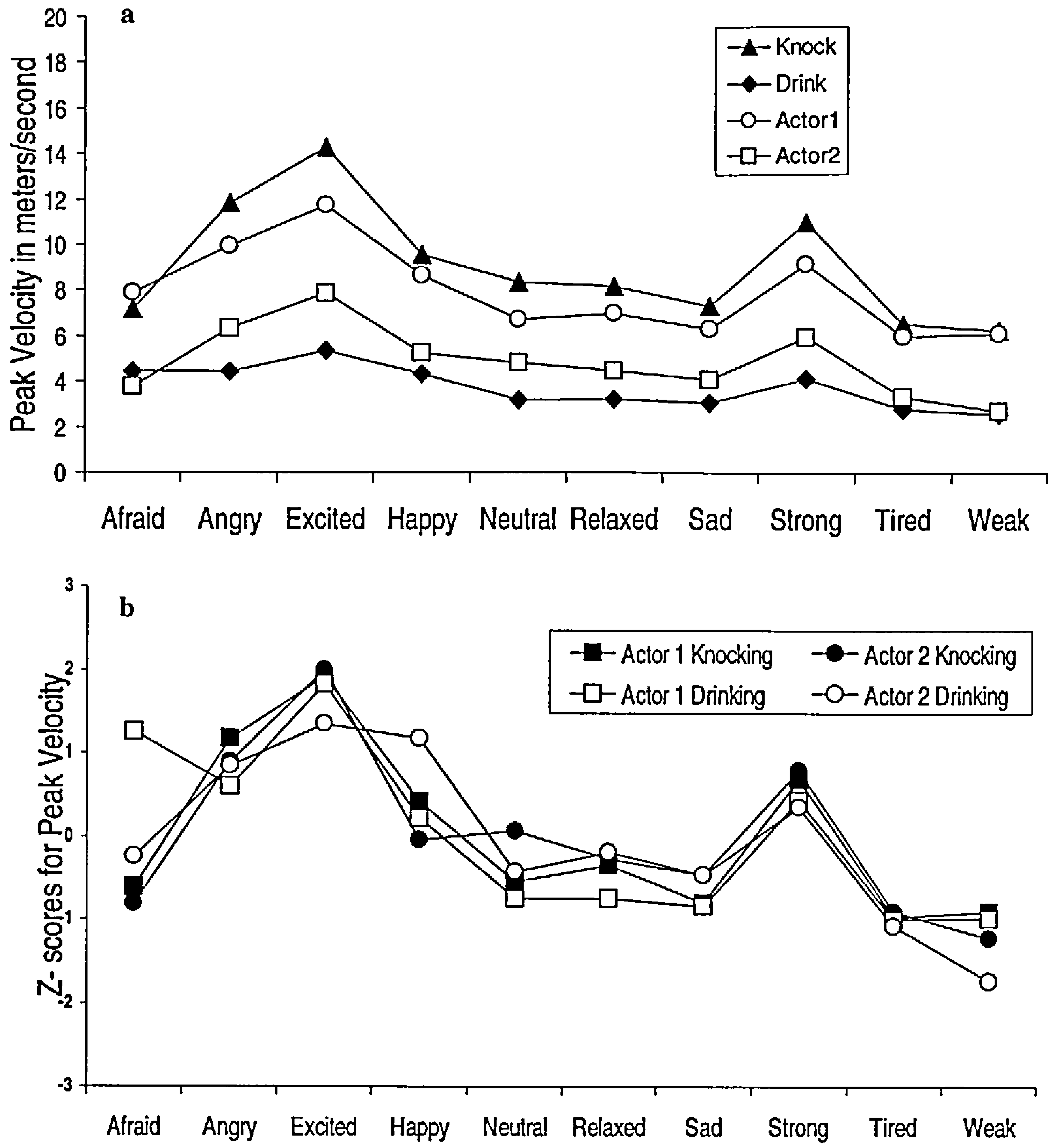

Figure III.2 Peak movement velocities for affects according to actors and actions a) with reference to the peak velocity of averaged for actions and actors and b) as z-scores of the four types of movements.

Distance Moved

In summary, findings for Distance Moved showed significant effects at all levels and interactions as shown below in the reproduced results for the ANOVA. In the graph of the distance Moved, it is clear that there is a consistent pattern in the distance moved by the arm for emotions across different movement types and that the differences in affects is preserved.

Analysis of Variance Summary Table

Data from Distance Moved
Mixed Design (alias Split Plot)

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (actor)	4879111.479	1	4879111.479	301.486	0.0001
B (action)	286806.793	1	286806.793	31.208	0.0050
C (affect)	1051156.635	9	116795.182	31.702	0.0000
AB	97230.202	1	97230.202	10.580	0.0313
AC	285956.442	9	31772.938	8.624	0.0000
BC	305011.934	9	33890.215	10.833	0.0000
ABC	199148.905	9	22127.656	7.073	0.0000
Between Error	64734.277	4	16183.569		
(Error BxS)	36760.570	4	9190.142		
(Error CxS)	132628.094	36	3684.114		
(Error BCxS)	112621.671	36	3128.380		

Simple Simple Main Effects for Selected Factors

	Source of Variation	Sum of Squares	df	Mean Squares	F	p
actor at						
knock	afraid	329158.735	1	329158.735	25.945	0.0009
knock	angry	279603.459	1	279603.459	22.039	0.0016
knock	excited	189186.629	1	189186.629	14.912	0.0048
knock	happy	744621.282	1	744621.282	58.692	0.0001
knock	neutral	123117.645	1	123117.645	9.704	0.0143
knock	relaxed	33690.027	1	33690.027	2.656	0.1418
knock	sad	192292.543	1	192292.543	15.157	0.0046
knock	strong	62945.284	1	62945.284	4.961	0.0565
knock	tired	55278.721	1	55278.721	4.357	0.0703
knock	weak	146403.136	1	146403.136	11.540	0.0094
drink	afraid	471715.462	1	471715.462	37.181	0.0003
drink	angry	271364.708	1	271364.708	21.389	0.0017
drink	excited	381684.223	1	381684.223	30.085	0.0006
drink	happy	287574.319	1	287574.319	22.667	0.0014
drink	neutral	338851.732	1	338851.732	26.709	0.0009
drink	relaxed	155920.410	1	155920.410	12.290	0.0080
drink	sad	284657.370	1	284657.370	22.437	0.0015
drink	strong	551777.931	1	551777.931	43.492	0.0002
drink	tired	136222.510	1	136222.510	10.737	0.0112
drink	weak	425380.901	1	425380.901	33.529	0.0004
Error Term	101494.846	8	12686.856			
action at						
actor1	afraid	14308.342	1	14308.342	1.557	0.2802
actor1	angry	70434.629	1	70434.629	7.664	0.0504
actor1	excited	41047.914	1	41047.914	4.467	0.1021
actor1	happy	150830.137	1	150830.137	16.412	0.0155
actor1	neutral	113.678	1	113.678	0.012	0.9168

actor1	relaxed	12960.335	1	12960.335	1.410	0.3007
actor1	sad	17337.901	1	17337.901	1.887	0.2415
actor1	strong	57378.459	1	57378.459	6.243	0.0669
actor1	tired	6559.653	1	6559.653	0.714	0.4458
actor1	weak	5053.821	1	5053.821	0.550	0.4995
actor2	afraid	42.584	1	42.584	0.005	0.9490
actor2	angry	66330.240	1	66330.240	7.218	0.0549
actor2	excited	390.172	1	390.172	0.042	0.8468
actor2	happy	3808.582	1	3808.582	0.414	0.5548
actor2	neutral	58510.995	1	58510.995	6.367	0.0651
actor2	relaxed	105730.729	1	105730.729	11.505	0.0275
actor2	sad	51390.461	1	51390.461	5.592	0.0773
actor2	strong	63700.997	1	63700.997	6.931	0.0580
actor2	tired	46208.124	1	46208.124	5.028	0.0884
actor2	weak	116060.082	1	116060.082	12.629	0.0237
Error Term	36760.570	4	9190.142			
affect at						
actor1	knock	680622.640	9	75624.738	20.527	0.0000
actor1	drink	409966.420	9	45551.824	12.364	0.0000
actor2	knock	444776.526	9	49419.614	13.414	0.0000
actor2	drink	305908.331	9	33989.815	9.226	0.0000
Error Term	132628.094	36	3684.114			

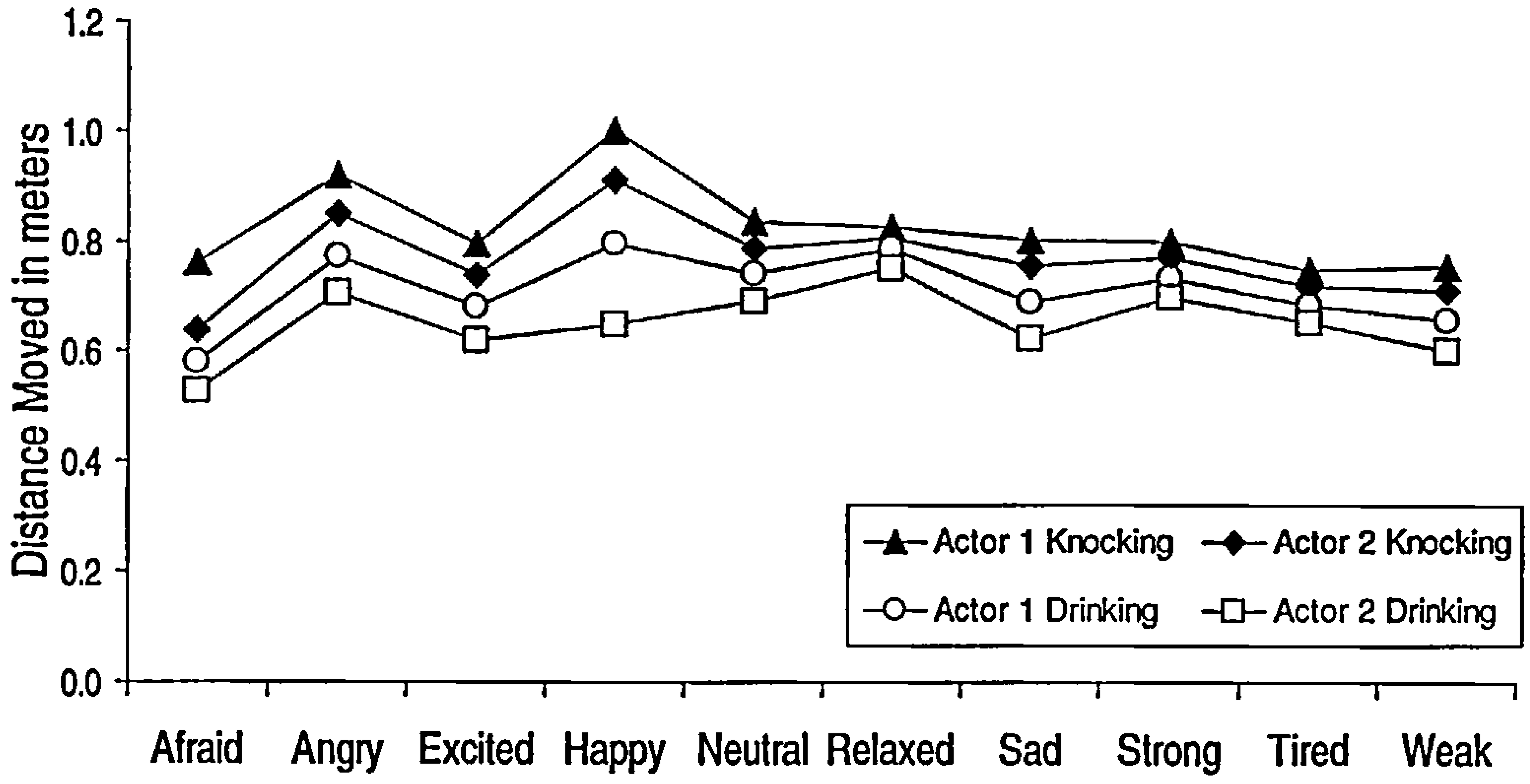

Figure III.3 the distance moved by the hand for different emotions, according to the different actors and actions.

Peak Acceleration

In summary, findings for Peak Acceleration showed significant effects at all levels and interactions as shown below in the reproduced results for the ANOVA. In the graph of the Peak Acceleration, it is clear that there is a consistent pattern in the distance moved by the arm for emotions across different movement types and that the differences in affects is preserved.

Analysis of Variance Summary Table

Data from duration
Mixed Design (alias Split Plot)

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (actor)	541.638	1	541.638	94.778	0.0006
B (action)	1491.468	1	1491.468	292.241	0.0001
C (affect)	1654.587	9	183.843	93.658	0.0000
AB	143.700	1	143.700	28.157	0.0061
AC	168.284	9	18.698	9.526	0.0000
BC	662.453	9	73.606	36.888	0.0000
ABC	61.262	9	6.807	3.411	0.0040
Between Error	22.859	4	5.715		
(Error BxS)	20.414	4	5.104		
(Error CxS)	70.665	36	1.963		
(Error BCxS)	71.834	36	1.995		

Simple Simple Main Effects for Selected Factors

	Source of Variation	Sum of Squares	df	Mean Squares	F	p
actor at						
knock	afraid	64.913	1	64.913	12.001	0.0085
knock	angry	22.718	1	22.718	4.200	0.0746
knock	excited	216.536	1	216.536	40.031	0.0002
knock	happy	136.703	1	136.703	25.272	0.0010
knock	neutral	20.170	1	20.170	3.729	0.0896
knock	relaxed	78.733	1	78.733	14.555	0.0051
knock	sad	29.016	1	29.016	5.364	0.0492
knock	strong	123.677	1	123.677	22.864	0.0014
knock	tired	17.327	1	17.327	3.203	0.1113
knock	weak	31.139	1	31.139	5.757	0.0432
drink	afraid	89.382	1	89.382	16.524	0.0036
drink	angry	9.743	1	9.743	1.801	0.2164
drink	excited	70.794	1	70.794	13.088	0.0068
drink	happy	0.664	1	0.664	0.123	0.7350
drink	neutral	0.014	1	0.014	0.003	0.9610
drink	relaxed	0.236	1	0.236	0.044	0.8399
drink	sad	0.405	1	0.405	0.075	0.7912
drink	strong	1.859	1	1.859	0.344	0.5738
drink	tired	0.023	1	0.023	0.004	0.9499
drink	weak	0.831	1	0.831	0.154	0.7053
Error Term		43.274	8	5.409		
action at						
actor1	afraid	0.355	1	0.355	0.070	0.8050
actor1	angry	500.357	1	500.357	98.041	0.0006
actor1	excited	344.934	1	344.934	67.587	0.0012
actor1	happy	198.081	1	198.081	38.812	0.0034

actor1	neutral	80.629	1	80.629	15.799	0.0165
actor1	relaxed	144.024	1	144.024	28.220	0.0060
actor1	sad	59.018	1	59.018	11.564	0.0273
actor1	strong	266.309	1	266.309	52.181	0.0019
actor1	tired	41.762	1	41.762	8.183	0.0459
actor1	weak	37.281	1	37.281	7.305	0.0539
actor2	afraid	3.972	1	3.972	0.778	0.4275
actor2	angry	429.472	1	429.472	84.152	0.0008
actor2	excited	150.582	1	150.582	29.505	0.0056
actor2	happy	10.222	1	10.222	2.003	0.2299
actor2	neutral	19.106	1	19.106	3.744	0.1251
actor2	relaxed	13.055	1	13.055	2.558	0.1850
actor2	sad	8.598	1	8.598	1.685	0.2641
actor2	strong	43.054	1	43.054	8.436	0.0439
actor2	tired	6.005	1	6.005	1.177	0.3390
actor2	weak	2.066	1	2.066	0.405	0.5592
Error Term		20.414	4	5.104		
affect at						
actor1	knock	1316.931	9	146.326	74.545	0.0000
actor1	drink	319.245	9	35.472	18.071	0.0000
actor2	knock	880.632	9	97.848	49.848	0.0000
actor2	drink	29.776	9	3.308	1.685	0.1287
Error Term		70.665	36	1.963		

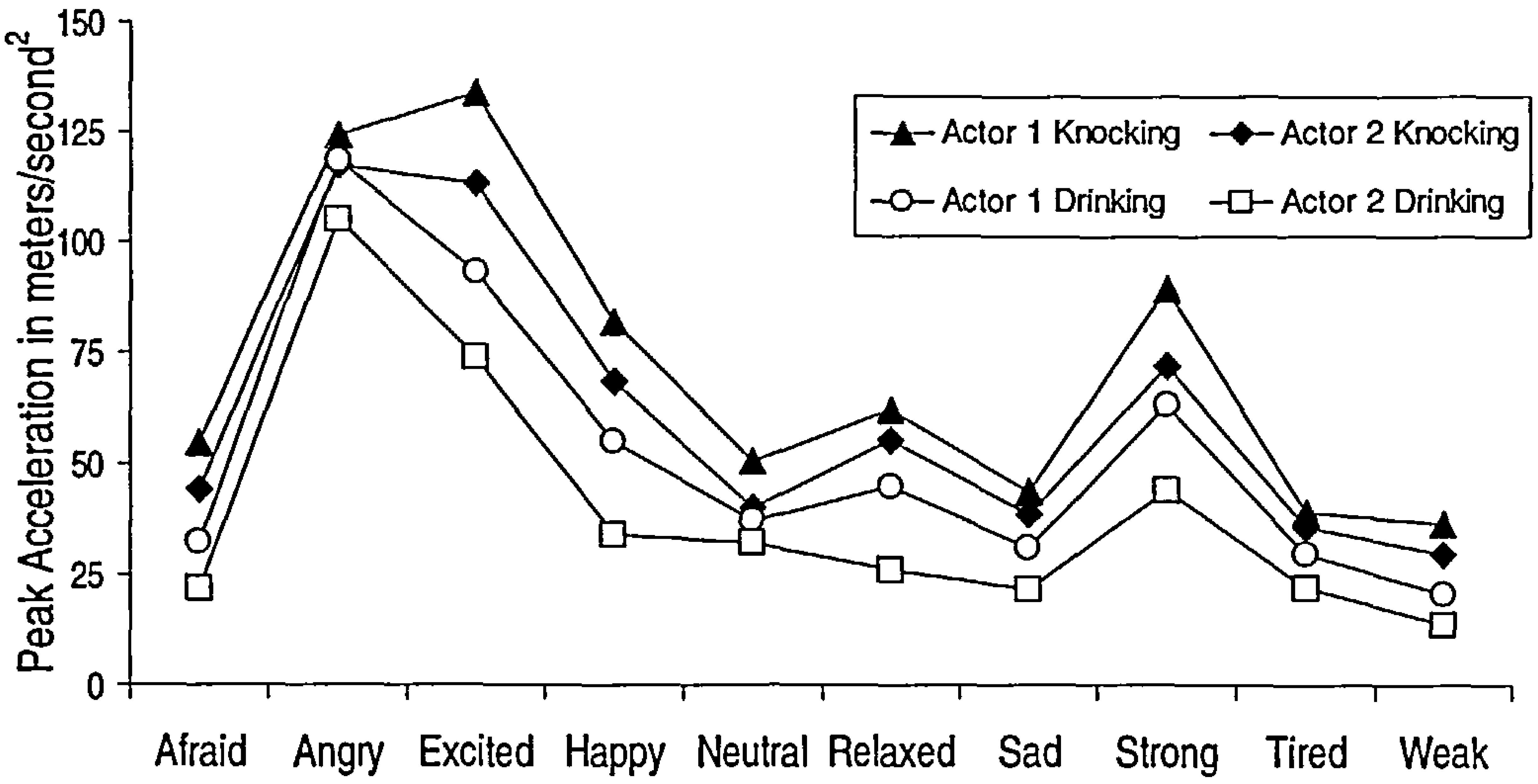

Figure III.4 the distance moved by the hand for different emotions, according to the different actors and actions.

Jerk Index

In summary, findings for Jerk Index showed significant effects at all levels and interactions as shown below in the reproduced results for the ANOVA. In the graph of the Jerk Index, it is clear that there is a consistent pattern in the distance moved by the arm for emotions across different movement types and that the differences in affects is preserved.

Analysis of Variance Summary Table

Data from Jerk Index
Mixed Design (alias Split Plot)

Source of Variation	Sum of Squares	df	Mean Squares	F	p
A (actor)	262.540	1	262.540	110.855	0.0005
B (action)	807.669	1	807.669	343.486	0.0000
C (affect)	589.514	9	65.502	105.055	0.0000
AB	182.936	1	182.936	77.799	0.0009
AC	154.005	9	17.112	27.444	0.0000
BC	404.668	9	44.963	66.236	0.0000
ABC	99.377	9	11.042	16.266	0.0000
Between Error	9.473	4	2.368		
(Error BxS)	9.406	4	2.351		
(Error CxS)	22.446	36	0.623		
(Error BCxS)	24.438	36	0.679		

Simple Simple Main Effects for Selected Factors

	Source of Variation	Sum of Squares	df	Mean Squares	F	p
actor at						
knock	afraid	15.534	1	15.534	6.583	0.0334
knock	angry	104.616	1	104.616	44.331	0.0002
knock	excited	306.902	1	306.902	130.051	0.0000
knock	happy	88.005	1	88.005	37.293	0.0003
knock	neutral	3.705	1	3.705	1.570	0.2456
knock	relaxed	22.910	1	22.910	9.708	0.0143
knock	sad	8.805	1	8.805	3.731	0.0895
knock	strong	130.281	1	130.281	55.207	0.0001
knock	tired	5.654	1	5.654	2.396	0.1603
knock	weak	3.745	1	3.745	1.587	0.2433
drink	afraid	1.770	1	1.770	0.750	0.4116
drink	angry	2.466	1	2.466	1.045	0.3366
drink	excited	3.160	1	3.160	1.339	0.2806
drink	happy	0.127	1	0.127	0.054	0.8223
drink	neutral	0.017	1	0.017	0.007	0.9343
drink	relaxed	0.038	1	0.038	0.016	0.9021
drink	sad	0.005	1	0.005	0.002	0.9630
drink	strong	1.111	1	1.111	0.471	0.5121
drink	tired	0.003	1	0.003	0.001	0.9723
drink	weak	0.003	1	0.003	0.001	0.9723
Error Term		18.879	8	2.360		
action at						
actor1	afraid	15.371	1	15.371	6.537	0.0629
actor1	angry	374.022	1	374.022	159.064	0.0002
actor1	excited	467.557	1	467.557	198.843	0.0001
actor1	happy	121.197	1	121.197	51.543	0.0020
actor1	neutral	23.221	1	23.221	9.875	0.0348
actor1	relaxed	40.876	1	40.876	17.384	0.0140
actor1	sad	22.430	1	22.430	9.539	0.0366

actor1	strong	210.876	1	210.876	89.682	0.0007
actor1	tired	18.087	1	18.087	7.692	0.0502
actor1	weak	10.090	1	10.090	4.291	0.1070
actor2	afraid	1.716	1	1.716	0.730	0.4411
actor2	angry	114.098	1	114.098	48.524	0.0022
actor2	excited	34.598	1	34.598	14.714	0.0185
actor2	happy	3.937	1	3.937	1.674	0.2653
actor2	neutral	9.148	1	9.148	3.890	0.1198
actor2	relaxed	1.993	1	1.993	0.848	0.4093
actor2	sad	3.393	1	3.393	1.443	0.2959
actor2	strong	17.317	1	17.317	7.365	0.0533
actor2	tired	3.313	1	3.313	1.409	0.3009
actor2	weak	1.408	1	1.408	0.599	0.4823
Error Term		9.406	4	2.351		
affect at						
actor1	knock	1057.046	9	117.450	188.372	0.0000
actor1	drink	15.444	9	1.716	2.752	0.0148
actor2	knock	174.082	9	19.342	31.022	0.0000
actor2	drink	0.990	9	0.110	0.176	0.9954
Error Term		22.446	36	0.623		

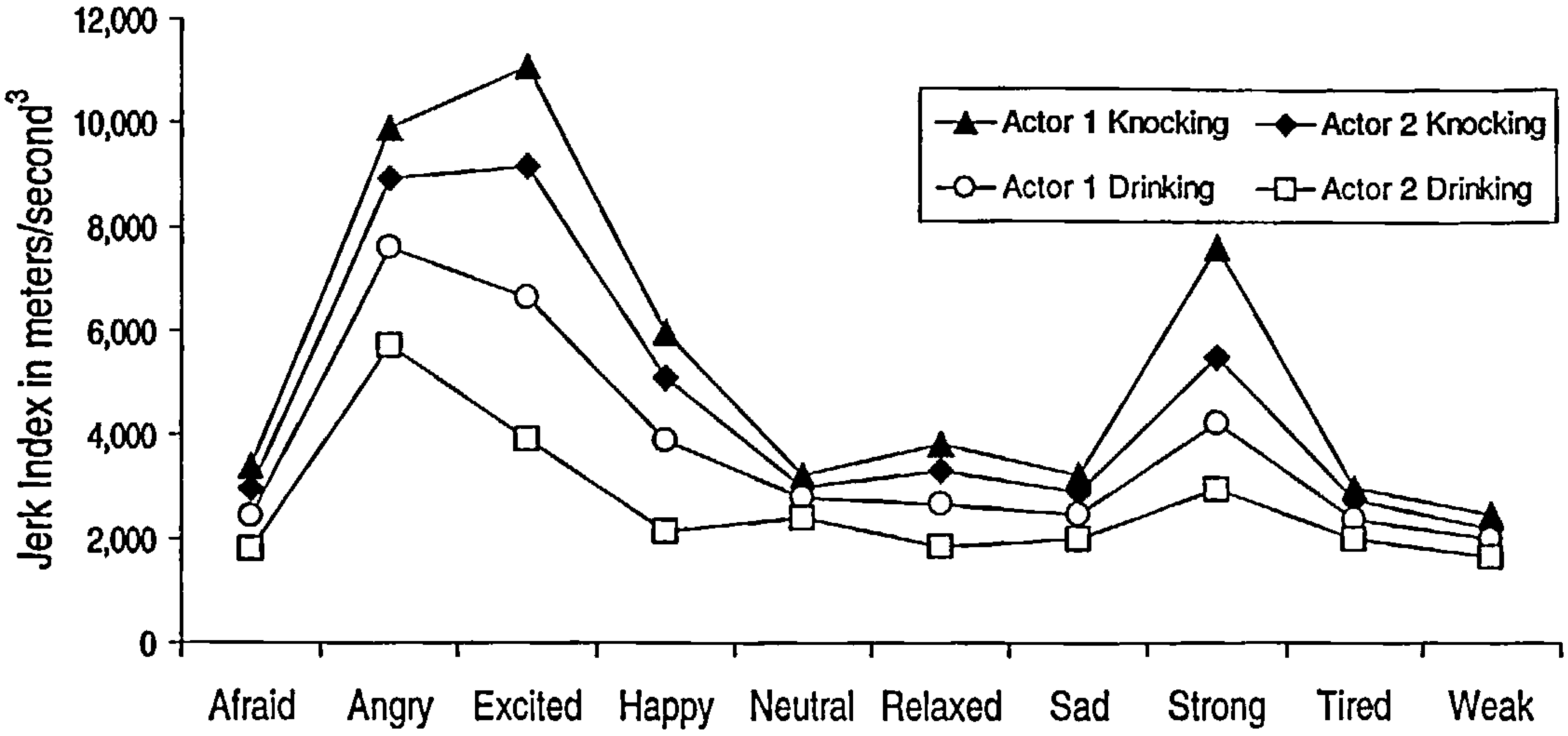

Figure III.5 the distance moved by the hand for different emotions, according to the different actors and actions.