The Role of Familiarity and Movement in Face Recognition

Rachel Jo Bennetts
BA(Psych) Hons
Macquarie University, Australia

A thesis submitted for the degree of Doctor of Philosophy

MARCS Institute
University of Western Sydney, Australia
April 2012

© R J Bennetts, 2012
ACKNOWLEDGEMENTS

I wish to thank the many people who have been instrumental in helping me write this thesis. First, my primary supervisor, Doctor Rachel Robbins, has been invaluable. Thank you for your support, guidance, patience, and good humour over the past few years. My associate supervisors, Doctor Darren Burke, Doctor Kevin Brooks, and Associate Professor Jeesun Kim have also provided support, guidance and excellent advice whenever I have asked for it, at all hours of the day and night. Thanks are also due to my collaborators at the CSIRO, Doctor Simon Lucey and Doctor Jason Saragih, who allowed me access to their amazing face-tracking program.

I have been fortunate enough to receive the help of an army of administrative and technical staff, across four departments in two universities. Special thanks are due to Sonya O’Shanna, Darlene Williams, Colin Schoknecht, Johnson Chen, and Donovan Govan. I also wish to thank all the staff and students at the MARCS Institute, for providing an exciting research environment, answering (sometimes foolish) questions, and generously sharing their knowledge with me over the past two years. In particular, thanks to Professor Denis Burnham, Associate Professor Kate Stevens, Professor Roger Dean, Doctor Freya Bailes, Doctor Michael Tyler, Laura Bishop, Ben Schultz, Staci Parlato-Harris, Erin Cvejic, Bronson Harry, Tim Paris, Michael Fitzpatrick, and Amanda Reid, for your comments, advice, and friendship.

Finally, I would like to thank my family: my Mum, Dad, and my brother Sam, who have been tireless in their support of my education, and my husband Timothy, without whom I could never have completed this thesis.
STATEMENT OF AUTHENTICATION

I hereby declare that this thesis is my own work. The work presented in this thesis is original and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due acknowledgement is made.

No material within this thesis has been submitted, in whole or part, for a degree at this or any other institution.

I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project’s design and conception is acknowledged.

_______________________________
Rachel Jo Bennetts
April 2012
# CONTENTS

ACKNOWLEDGEMENTS........................................................................................................II

STATEMENT OF AUTHENTICATION..............................................................................III

CONTENTS..................................................................................................................... IV

LIST OF TABLES ............................................................................................................. X

LIST OF FIGURES .......................................................................................................... XI

ABSTRACT....................................................................................................................... XIII

CHAPTER 1: INTRODUCTION......................................................................................... 3
  1.1 Familiarity, Movement, and Face Recognition ....................................................... 3
  1.2 Outline of the Thesis ............................................................................................. 6

CHAPTER 2: FACE RECOGNITION AND FAMILIARITY: ARE FAMILIAR AND UNFAMILIAR FACES DIFFERENT? ............................................................... 9
  2.1 What is Familiarity? ............................................................................................... 10
  2.2 Models of Face Recognition.................................................................................. 11
  2.3 Familiar and Unfamiliar Faces: Behavioural Research ....................................... 18
    2.3.1 Viewpoint ...................................................................................................... 18
    2.3.2 Expression .................................................................................................... 19
    2.3.3 Internal and External Features .................................................................... 21
    2.3.4 Degraded and Inverted Faces ..................................................................... 22
    2.3.5 Conclusions from the Behavioural Literature .............................................. 27
  2.4 Familiar and Unfamiliar Faces: Neuroimaging and Neuropsychological Research ......................................................................................................................... 28
  2.5 Familiar and Unfamiliar Faces: Conclusions ....................................................... 32

CHAPTER 3: MOVEMENT AND FACE RECOGNITION ............................................. 35
  3.1 Biological Motion .................................................................................................. 36
6.10.3 Eyes and Mouths 221
6.10.4 Advantages and Disadvantages of the Sorting Task 222
6.10.5 Conclusions from Experiments 5, 6, and 7 224

CHAPTER 7: THE ROLE OF MOVEMENT IN PERSONALLY FAMILIAR
AND SELF-FACE RECOGNITION ................................................................. 227

7.1 Recognition of Personally Familiar and Self-Faces ............................................. 227
   7.1.1 Personally Familiar Faces 227
   7.1.2 Self Recognition 230

7.2 Experiment 8: Self, Personally Familiar and Unfamiliar Face Matching
From Rigid and Non-Rigid Movements ............................................................ 232
   7.2.1 Methods of Experiment 8 235
   7.2.2 Results for Experiment 8 238
   7.2.3 Discussion of Experiment 8 244

7.3 Experiment 9: Natural Exaggeration and Temporal Variation in Self,
Personally Familiar and Unfamiliar Face Matching .......................................... 248
   7.3.1 Methods of Experiment 9 254
   7.3.2 Results of Experiment 9 258
   7.3.3 Discussion of Experiment 9 266

7.4 Conclusions from Experiments 8 and 9 ......................................................... 270

CHAPTER 8: MOVEMENT AND FAMILIARITY IN FACE RECOGNITION:
CONCLUSIONS ............................................................................................ 275

8.1 Familiarity and the Movement Advantage: Structure-from-Motion and
Characteristic Movement Patterns .................................................................... 276
   8.1.1 Structure-from-motion 279
   8.1.2 Characteristic Movement Patterns 282
   8.1.3 Familiarity and the Movement Advantage: Other Findings 285

8.2 Exceptions to the Movement Advantage .................................................... 288

8.3 Methodology and the Movement Advantage ............................................. 291
   8.3.1 Does the Task Matter? 292
   8.3.2 Do the Stimuli Matter? 294
8.4 Conclusions: Qualitative and Quantitative Differences in the Movement Advantage for Familiar and Unfamiliar Faces .................................................................297

REFERENCES .............................................................................................................301

APPENDIX A: STIMULI ...........................................................................................316
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>54</td>
</tr>
<tr>
<td>Table 2</td>
<td>62</td>
</tr>
<tr>
<td>Table 3</td>
<td>66</td>
</tr>
<tr>
<td>Table 4</td>
<td>76</td>
</tr>
<tr>
<td>Table 5</td>
<td>96</td>
</tr>
<tr>
<td>Table 6</td>
<td>102</td>
</tr>
<tr>
<td>Table 7</td>
<td>105</td>
</tr>
<tr>
<td>Table 8</td>
<td>110</td>
</tr>
<tr>
<td>Table 9</td>
<td>114</td>
</tr>
<tr>
<td>Table 10</td>
<td>150</td>
</tr>
<tr>
<td>Table 11</td>
<td>156</td>
</tr>
<tr>
<td>Table 12</td>
<td>185</td>
</tr>
<tr>
<td>Table 13</td>
<td>193</td>
</tr>
<tr>
<td>Table 14</td>
<td>198</td>
</tr>
<tr>
<td>Table 15</td>
<td>204</td>
</tr>
<tr>
<td>Table 16</td>
<td>208</td>
</tr>
<tr>
<td>Table 17</td>
<td>277</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1: The Bruce and Young (1986) model of face processing ......................... 12

Figure 2: Neural models of face recognition.......................................................... 15

Figure 3: Our internal representation of a face may incorporate someone's average appearance and how their appearance varies in our experience. From Burton et al. (2011). ...................................................................................................................... 17

Figure 4: Examples of common image manipulations................................................. 23

Figure 5: Commonly used stimuli in biological motion research .......................... 37

Figure 6: Roark et al.’s (2003) neural systems model of movement and face recognition ....................................................................................................................... 41

Figure 7: Placement of dots used when creating point-light-displays (left) and a point-light-display (right). ................................................................................................................. 98

Figure 8: The timeline of a match-to-sample trial in Experiment 1......................... 100

Figure 9: The movement advantage for different clip lengths in Experiment 1a ....... 103

Figure 10: The movement advantage for different clip lengths in Experiment 1b ...... 106

Figure 11: The movement advantage for different clip lengths in Experiment 2a....... 111

Figure 12: The movement advantage for different clip lengths in Experiment 2b ...... 115

Figure 13: Positions of the points tracked in the shape-normalised avatar stimuli ...... 147

Figure 14: The movement advantage for different movement conditions in Experiment 3 ......................................................................................................................... 151

Figure 15: The movement advantage for different movement conditions in Experiment 4 ......................................................................................................................... 158

Figure 16: An illustration of the sorting task.............................................................. 186
Figure 17: An example of the stimuli used in the pilot sorting task and Experiments 5, 6, and 7.................................................................189

Figure 18: Accuracy-per-view results for different movement types in Experiment 5 .............................................................................................195

Figure 19: Accuracy-per-view results for different movement types in Experiment 6 ...........................................................................................201

Figure 20: Accuracy-per-view results for different movement types in Experiment 7 ...........................................................................................210

Figure 21: The configuration of markers used to create the PLDs in Experiment 8, and an example of the PLD stimuli ........................................................236

Figure 22: d' results from the matching task in Experiment 8........................240

Figure 23: Proportion correct in the recognition task in Experiment 8................242

Figure 24: Response frequency in the recognition task of Experiment 8........245

Figure 25: Accuracy for the Experiment 9 sorting task ................................260

Figure 26: Accuracy-per-view for the Experiment 9 sorting task ..................262

Figure 27: Proportion correct in the recognition task in Experiment 9.............264

Figure 28: Response frequency in the recognition task of Experiment 9.........265
ABSTRACT

Research on static faces has found that familiarity has a strong influence on face recognition. To this point, though, few studies have compared familiar and unfamiliar face recognition using moving stimuli. This thesis examined the effects of familiarity and movement on face recognition, to determine whether movement conveys the same benefits for both familiar and unfamiliar faces.

There are both empirical and theoretical reasons to suppose that familiar and unfamiliar faces would derive different benefits from movement. In general, studies using familiar faces have found a reliable advantage for moving stimuli, whereas the movement advantage for learning or matching unfamiliar faces is less consistent. Furthermore, a recent model of movement and face recognition (Roark et al., 2003) suggests that familiar faces should benefit from movement in two ways: first, movement may increase the amount of structural information available in a face; second, people may be able to identify someone’s characteristic movement patterns. Unfamiliar faces, on the other hand, should primarily benefit from enhanced structural information, with little (if any) role for characteristic movement patterns. This implies that familiarity may result in a quantitative, and possibly qualitative, difference in the movement advantage for face recognition. Alternatively, the apparent differences found between familiar and unfamiliar faces may be purely methodological: studies of movement-based face recognition have often used significantly different tasks and stimuli when testing familiar and unfamiliar faces.

This thesis presents nine experiments that assess the role of movement in familiar and unfamiliar faces, with a particular focus on examining whether the difference between familiar and unfamiliar faces is qualitative or quantitative, and on the influence of methodological factors on the movement advantage. Experiments 1 – 4 assess the role of movement in famous and unfamiliar face matching, and examine methodological factors such as stimulus duration, task, and overlapping movement sequences (Experiments 1 and 2); or the type of stimulus (Experiments 3 and 4). Experiments 5, 6 and 7 investigate whether the movement advantage for famous and unfamiliar faces varies depending on the type of movement (rigid, non-rigid, or both) and whether adding movement cues from the eyes and mouth increases the movement advantage. The final
set of experiments (Experiments 8 and 9) examines the effect of two different types of familiarity, using personally familiar and self-faces.

Overall, there appears to be little qualitative difference between the movement advantage for familiar and unfamiliar faces, but the size of the movement advantage (i.e., the quantitative effect of familiarity) varies depending on the task. When participants match or sort two faces (Experiments 1-4, 8-9), unfamiliar faces show a larger movement advantage than familiar faces, primarily because static matching of familiar faces is highly accurate. However, when participants sort four faces at once (Experiments 5-7), famous faces show a larger movement advantage than unfamiliar faces. These results are noteworthy for two reasons: they are the first results to suggest that unfamiliar faces may benefit more from movement than familiar faces under some circumstances; and the first to show that the movement advantage for both familiar and unfamiliar faces relies on a combination of structure-from-motion and characteristic movement patterns. Overall, the results point to a flexible use of static and dynamic cues to identity. The relative contribution of static and dynamic cues varies depending on the familiarity of the face, the amount of static information in the stimuli, and the number of comparisons required to complete the task.
Chapter 1

Introduction
CHAPTER 1: INTRODUCTION

1.1 Familiarity, Movement, and Face Recognition

Human faces are some of the most important visual stimuli we encounter. They contain information about personal characteristics such as age, gender and racial background. Faces can also communicate social information, such as an individual’s emotions, where or to whom they are attending, and what they are saying. One of the most important and highly researched pieces of information carried in the face is identity – people are remarkably good at using facial information to distinguish between the many hundreds of people they encounter over the course of their lives (Bruce & Young, 1998). Many factors affect face recognition. This thesis will consider two – familiarity with a face, and the movement of the face. The effects of familiarity and movement on face recognition have been examined separately, but few studies have attempted to combine these areas. This thesis will examine the interaction between familiarity and movement in a variety of face recognition tasks, to determine whether the use of movement information is consistent across all levels of familiarity. Put another way, this thesis will address whether familiarity with a face opens up extra avenues for person recognition, such as the identification of characteristic movements, or enhances static recognition, via the use of extra three-dimensional cues to face shape; or whether these cues are available and useful for all faces, regardless of prior exposure. This is an important question, with implications for theoretical models of face recognition and real-world face recognition scenarios such as eyewitness testimony.

People are generally considered to be highly accurate at face recognition, but not all faces are recognised equally well. If we are familiar with a person, we can recognise them even from very low quality images (Burton, Wilson, Cowan, & Bruce, 1999), and despite changes in viewpoint, expression, and context (Johnston & Edmonds, 2009). However, people are surprisingly poor when asked to perform identification-based tasks – for example, matching or old/new recognition tasks – on unfamiliar faces (Hancock, Bruce, & Burton, 2000). Recognition of unfamiliar faces is reduced even further when
changes are made to viewpoint, expression, lighting, or even the camera used to take the photographs. (Bruce, 2004; Jenkins & Burton, 2011).

Why are we so much better at identifying familiar than unfamiliar faces? It is possible that the difference is purely quantitative: we have more experience with familiar faces, and we have seen them from different viewpoints, displaying different expressions, and with different hairstyles or lighting conditions, whereas we have only been exposed to an unfamiliar face briefly, from one viewpoint and with one expression. Alternatively, it is possible that the difference between familiar and unfamiliar faces is qualitative. Recent reviews of the behavioural and neuropsychological literature suggest that familiarisation with a face fundamentally changes not only our subsequent identification performance, but also the mental representation of the face and the cues we use to identify a person (Burton, Jenkins, & Schweinberger, 2011; Burton, Jenkins, Hancock, & White, 2005; Johnston & Edmonds, 2009). Notably, these authors suggest that our representations of familiar faces average across variations in viewpoint, lighting, size and other extraneous factors (Burton et al., 2005; Burton et al., 2011), whereas representations of unfamiliar faces are heavily influenced by these same, image-based factors (Hancock et al., 2000; Johnston & Edmonds, 2009). The effect of familiarity on static face representations and recognition will be reviewed in detail in chapter 2.

Movement is another important factor in face recognition. In the real world, we see faces moving – rotating, nodding, speaking and expressing. A recent line of research has found that seeing a face in movement can often improve recognition performance (O’Toole, Roark & Abdi, 2002). Familiar faces are often recognised more accurately from moving than static images (Knight & Johnston, 1997; Lander & Bruce, 2000; Lander, Bruce & Hill, 2001; Lander, Christie & Bruce, 1999). Furthermore, unfamiliar faces may also benefit from movement: in some cases, faces learnt in movement are subsequently recognised more accurately than faces learnt from static images (Pike, Kemp, Towell & Philips, 1997), and unfamiliar faces can be matched based purely on their movement patterns (Hill & Johnston, 2001). Throughout the current series of experiments, any case where moving images result in better recognition, learning or matching performance than static images will be referred to as a “movement advantage”. Evidence for a movement advantage in familiar and unfamiliar faces will be reviewed in
chapter 3.

The most recent model of movement and face recognition (Roark, Barrett, Spence, Abdi & O’Toole, 2003) identified three ways that movement could influence face recognition: first, movement may increase the amount of structural information available in a face; second, people may be able to identify a person’s characteristic movement patterns; and third, social cues inherent in movement may attract attention to (or detract attention from) the processing of identity. Roark et al.’s model offers a comprehensive account of how movement can impact on face recognition. However, one factor that has not been thoroughly examined by the model or the empirical research supporting the model is the effect of familiarity on the movement advantage. There are both theoretical and empirical reasons to suppose that familiar and unfamiliar faces derive different benefits from movement. However, few studies have directly compared the effect of movement on face recognition for familiar and unfamiliar faces. Those that have focused on whether familiar and unfamiliar faces showed a similar quantitative benefit from movement (e.g. Bruce, Henderson, Newman & Burton, 2001), and did not examine whether the movement advantage for familiar and unfamiliar face recognition used the same mechanisms, (i.e. enhanced structural information or characteristic movement patterns), or whether they used them to the same extent. Studies that have examined movement and face recognition for familiar and unfamiliar face separately have used very different procedures and stimuli, making it difficult to draw conclusions about the effect of movement and face recognition for familiar and unfamiliar faces. The aim of this thesis, therefore, was to systematically examine the relationship between familiarity and movement in face recognition.

Specifically, the thesis aimed to answer the question: does the effect of movement in face recognition depend on our level of familiarity with the face? This question was addressed from three angles. First, the thesis examined whether familiar faces show a greater movement advantage than unfamiliar faces – in other words, does familiarity result in a quantitative difference in movement-based face recognition? Second, the thesis examined whether familiar and unfamiliar faces rely on different types of movement information (structural cues or characteristic movement patterns) during movement-based face recognition – in other words, does familiarity result in a
Chapter 1: Introduction

 qualitative difference in movement-based face recognition? Finally, the thesis investigated the effect of methodological and stimulus-based factors on familiar and unfamiliar faces, by examining if the effect of movement for familiar and unfamiliar faces is apparent in different tasks, with different stimuli; or whether the movement advantage varies depending on factors such as the duration of movement, type of movement or the area of the face visible to participants.

The following section provides a brief outline of each chapter of the thesis, and indicates the key concepts that are discussed in each chapter.

1.2 Outline of the Thesis

Chapter 2 provides a brief outline of the concept of familiarity, and reviews current theories of face recognition relating to it.

Chapter 3 reviews research on movement, from biological motion in whole bodies to facial movement. This chapter examines the research on movement in face recognition in relation to Roark et al.’s (2003) model, with a focus on the differences – both empirical and methodological – between studies of familiar and unfamiliar faces.

Chapter 4 presents two experiments that assess the role of movement and familiarity in face recognition. These are the first experiments to directly compare the effect of movement for famous and unfamiliar faces. Experiments 1 and 2 address three methodological questions: the effect of task type on performance; the effect of task difficulty on performance; and the optimum duration of moving stimulus for famous and unfamiliar faces. Chapter 4 also includes data from a pilot experiment that was used to select moving clips and control for the perceived amount and distinctiveness of movement of the stimuli in Chapters 4, 5 and 6.

Chapter 5 presents two experiments that compare the movement advantage for famous and unfamiliar faces, extending on the results presented in Chapter 4. Experiments 3 and 4 examine whether different stimulus types affect movement-based face processing for famous and unfamiliar faces. Using a same/different matching task, these experiments compare shape-normalised stimuli with stimuli that preserve the person’s face shape. Using shape-normalised stimuli impairs the use of structure-from-
motion cues, and forces people to rely on characteristic motion patterns to match faces. This chapter concludes by discussing the contribution of shape, and, by extension, structural cues, to the movement advantage in famous and unfamiliar face matching.

Chapter 6 focuses on the use of characteristic movement patterns in face matching. This chapter presents three experiments that examine the effect of different types of motion – rigid head motion, non-rigid face motion or combined rigid and non-rigid motion – and whether adding movement information to the eye or mouth region of the face improved performance for famous and unfamiliar faces. Experiments 5, 6 and 7 employ a different task to Experiments 1 – 4; participants were required to sort faces into groups rather than match individual pairs (Hill & Johnston, 2001). Chapter 6 also introduces a new dependent measure based on the sorting task, which may offer a different perspective on the benefits of movement for familiar and unfamiliar face matching.

Chapter 7 departs from Chapters 4 - 6, by presenting two studies that examined different types of familiarity. Specifically, Experiments 8 and 9 compare the effect of movement for unfamiliar, personally familiar and self-faces. Both experiments include a naming task, which enables a comparison between matching performance and overt recognition. Experiments 8 and 9 also investigate the effect of movement type, speech type and exaggeration on matching performance – factors that have been generally overlooked in the movement-based face recognition literature.

This thesis finishes by examining the implications of the current study for models of face processing and our understanding of face processing in the real world, and proposing future directions for research in this area. Chapter 8 begins by reviewing the nine experiments described in Chapters 4 – 7. The discussion focuses on the similarities and differences between movement-based face matching for familiar and unfamiliar faces, with particular reference to the use of structure-from-motion and characteristic movement patterns. Chapter 8 also examines the effect of methodological factors on the results. Overall, the results support a quantitative, but not qualitative, difference between movement-based face matching for familiar and unfamiliar faces. The chapter concludes by proposing that people use static and movement-based cues in a hierarchical or strategic manner, limited by attention.
Chapter 2

Face Recognition and Familiarity

Are Familiar and Unfamiliar Faces Different?
CHAPTER 2: FACE RECOGNITION AND FAMILIARITY: ARE FAMILIAR AND UNFAMILIAR FACES DIFFERENT?

It has often been said that humans are “experts” at face recognition (e.g., Parr, 2011). Indeed, a large amount of research has established that we are extremely accurate when asked to recognise faces of friends or celebrities, even when we have to generalise across changes in viewpoint, expression, context and age (Bruce & Young, 1998). Familiar faces can be identified from poor quality or degraded images (Burton, Wilson, et al. 1999; Collishaw & Hole, 2000) and from pictures that have been deformed by stretching or caricaturing (Hole, George, Eaves, & Rasek, 2002; Benson & Perrett, 1991). When compared to familiar faces, recognition performance for unfamiliar faces is surprisingly poor – even when photographs are taken on the same day, with the same pose, expression and lighting, people are poor at matching the photographs to other still images, or to live actors (Davis & Valentine, 2009; Megreya & Burton, 2008). When the images to be matched incorporate changes in viewpoint (e.g., front to profile view), expression (e.g., smiling to frowning) or superficial details (such as hairstyle, glasses etc.), performance is even worse (Hancock et al., 2000; Henderson, Bruce, & Burton, 2001).

The focus of this thesis is the role of familiarity in movement-based face recognition. However, only a relatively small amount of research has investigated face recognition using moving stimuli, and very little has examined the effect of familiarity. Therefore, this chapter reviews face recognition research that has used static faces, and examines why performance for familiar and unfamiliar faces is so different. The first section reviews the concept of familiarity, and how the terms “familiar” and “unfamiliar” are used in face recognition research. The second section provides an outline of dominant behavioural and neural models of face recognition, and examines the theoretical differences between familiar and unfamiliar face processing according to these models. The third and fourth sections of this chapter examine behavioural and neural research on face recognition, in an attempt to clarify whether the differences
between familiar and unfamiliar face recognition are quantitative (based purely on different amounts of exposure to a face) or qualitative (different ways of representing or processing familiar and unfamiliar faces).

2.1 What is Familiarity?

The research on face recognition can be divided into four categories of familiarity: unfamiliar, experimentally familiar, personally familiar, and famous face recognition. This section provides a brief definition of each term, and how discusses recognition is defined and measured for different categories of familiar and unfamiliar faces.

Unfamiliar face recognition is a somewhat contradictory term because, by definition, we should not be able to recognise faces that are unfamiliar to us. However, in this thesis the term “unfamiliar face recognition” will be used when participants are asked to complete any identification-based tasks using faces that they have not seen before the experiment, and that have not undergone a specific familiarisation phase. By this definition, unfamiliar face recognition occurs when people are able to identify two images as belonging to the same person – for example, during a sequential or simultaneous same/different identity task, or when people are asked to identify a “suspect” (an unfamiliar face they have only seen once) out of a line-up of other unfamiliar faces (e.g., Megreya & Burton, 2008).

Experimentally familiar face recognition occurs when people are asked to recognise a previously unfamiliar face that has undergone some experimental familiarisation (this is also referred to as “visually familiar” face recognition, e.g., Gobbini & Haxby, 2007; Natu & O’Toole, 2011). Studies using experimentally familiar faces typically involve two phases: a learning phase, where participants are familiarised with the face, and a test phase, where they are asked to recognise the familiarised faces. Recognition may involve naming images in the test phase, or verifying that an image in the test phase shows someone who was familiarised in the learning phase (an old/new recognition experiment). Experimental familiarisation is often used to investigate what factors affect face learning, or to control the amount of time or biographical information
a person is exposed to while becoming familiar with a face (e.g., Bruce & Valentine, 1988, Experiment 1).

Famous and personally familiar face recognition are the most frequently studied forms of familiar face recognition. Like experimentally familiar face recognition, famous and personally familiar face recognition experiments usually involve participants naming or providing other identifying information about a face (e.g., Knight & Johnston, 1997), or identifying a face as famous or familiar (e.g., Collishaw & Hole, 2000). However, the faces in famous and personally familiar face studies are familiar outside the experimental setting – either from exposure in real life (personally familiar), or exposure in movies, television shows or magazines (famous). Unlike experimentally familiar faces, there is no control over how familiar a personally familiar or famous face will be to participants (i.e., how many times a face has been seen, under what condition), or whether the face will carry other associations – for example, autobiographical memories or specific emotions.

In general, the term familiar encompasses famous and personally familiar faces, as well as experimentally familiar faces that have undergone an extensive familiarisation (Johnston & Edmonds, 2009). Throughout this thesis, the term familiar will refer to both famous and personally familiar faces. Experimentally familiar and unfamiliar faces will be identified separately.

### 2.2 Models of Face Recognition

Many models have been proposed to explain human face recognition. This section will outline four influential cognitive and neural models, with a specific focus on the differences between familiar and unfamiliar face recognition. Evidence for these models is discussed in section 2.3.

Bruce and Young (1986) proposed a cognitive model of face recognition that relies on the extraction of various types of information (codes) from a perceived face (Figure 1). The first stage of processing leads to the creation of abstract structural codes, which capture invariant aspects of the face – those that remain unchanged despite alterations in expression, lighting or viewpoint. Bruce and Young (1986) suggested that structural
codes provide information to *face recognition units* (FRUs), which contain stored structural codes (internal representations of a face) for all known faces. According to the model, face recognition then proceeds in a serial manner: a face is identified as familiar when there is a match between the face and a stored structural code in a FRU; FRUs then activate *person identification nodes*, which contain semantic information about a familiar person, such as their occupation and friends; finally, activation of the person identification node allows us to access the *name code* for that person. Obviously, faces contain more than just identity information. Even when we do not know a person, we are quite adept at judging characteristics such as age, gender and race (Bruce & Young, 1998). Bruce and Young (1986) referred to this type of information as visually derived semantic codes. The Bruce and Young model (1986) proposed that these semantic codes might provide a mechanism for recognising unfamiliar faces, as they can be extracted without the need for recognition.

![Figure 1: The Bruce and Young (1986) model of face processing. Adapted from Bruce and Young (1986).](image-url)
In addition to semantic information, we are also able to extract more fleeting cues, such as emotional expressions, eye gaze and speech information from a face. However, Bruce and Young (1986) suggested that expressions, eye gaze and speech information are processed separately from identity – in fact, their model suggests that the structural codes used for identification specifically factor out changeable aspects such as expressions.

Familiarity of faces is an important factor in the Bruce and Young (1986) model – it suggests that familiar and unfamiliar faces are processed in a qualitatively different manner. Familiar faces proceed from structural encoding to FRUs and person identification nodes, as noted above. In contrast, unfamiliar faces are recognised primarily via visually derived semantic codes, and/or selective or strategic attention to particular types of information – referred to as directed visual processing (Bruce & Young, 1986). Bruce and Young (1986) also state that the structural codes derived for familiar and unfamiliar faces are different. Structural codes for familiar faces are refined and elaborated through experience to incorporate within-person variability but factor out extraneous pictorial changes (Burton et al., 2011; Burton et al., 2005; Bruce, 1994), whereas structural codes for unfamiliar faces rely on basic pictorial information, such as viewpoint and lighting (Hancock et al., 2000).

Haxby, Hoffman, and Gobbini (2000) proposed a neural model of face recognition that mirrors many elements of the Bruce and Young (1986) model (Figure 2a). Most importantly, Haxby et al. (2000) suggest that the invariant aspects of a face (i.e., identity information, or structural codes) are processed independently of the changeable aspects of a face, such as expression, eye gaze, and speech movements. In this model, early face processing takes place in the inferior occipital gyri. From there, facial information is passed onto two core areas: invariant aspects of the face are processed in the lateral fusiform gyrus, in an area commonly known as the fusiform face area (FFA); changeable aspects are processed in the posterior superior temporal sulcus (STS). These core areas interact with an extended neural system, which processes the meaning of the information: for example, information from the FFA is passed to the anterior temporal area, which is associated with the retrieval of biographical information and names. Gobbini and Haxby (2007) modified and updated Haxby et al.’s (2000) model,
proposing an extended neural network for familiar face recognition (Figure 2b). The extended network includes a number of areas involved in the retrieval of person knowledge (e.g., the anterior paracingulate, involved in processing personal traits, attitudes and mental states), and areas involved in modulating the emotional response to a face (e.g., the amygdala, insula and striatum).

Haxby et al. (2000) and Gobbini and Haxby (2007) did not specifically highlight the neural differences between familiar and unfamiliar faces – their models were focussed primarily on understanding familiar face recognition. However, several studies have found that familiarity modulates the response of the core and extended neural areas implicated in face recognition, including the FFA (Gobbini, Liebenluft, Santiago, & Haxby, 2004; Liebenluft, Gobbini, Harrison, & Haxby, 2004). These findings will be described in more detail in section 2.4. However, it is important to note that familiarity, or the acquisition of better structural codes and FRUs through higher levels of exposure, does not simply increase neural responsiveness to a face. Instead, familiarity changes the pattern of responsiveness of a network of neural regions, incorporating areas responsible for face perception, person knowledge and emotions.

Although Bruce and Young (1986), Haxby et al. (2000), and Gobbini and Haxby (2007) offered comprehensive and influential models of face recognition, they did not specify what information was contained in structural codes, or how structural codes and FRUs were derived. Burton, Bruce and Hancock (1999) proposed a model of familiar face recognition that incorporated principal components analysis (PCA) as a method of extracting identity information (or structural codes) from a face. Burton, Bruce, et al. (1999) do not suggest that the human perceptual system performs PCA, but they suggest that the information derived from a PCA analysis may overlap with the information used during face recognition. In essence, they proposed a perceptual system that breaks face images down into simple parameters by analysing the variation between different faces, reducing this variation to a small number of dimensions, and then representing each face as a weighted sum of these dimensions. While this model explains the process behind familiar face recognition, Burton, Bruce, et al. (1999) did not specify how unfamiliar faces could be matched, or how familiarity with a face could alter structural codes.
Figure 2: Neural models of face recognition. a) The distributed neural model of face processing, adapted from Haxby et al. (2000). b) A neural model of familiar face processing, adapted from Gobbini and Haxby (2007).
Building on earlier PCA work, Burton et al. (2005) investigated the concept of familiarity and structural codes. Burton et al. (2005) used image averages\(^1\) to approximate the effect of familiarity with a face (i.e. adding more images to the average to simulate extra encounters with the person). PCA performance was poor when only one image was used in the analysis (the “unfamiliar” face condition), possibly because the results were dominated by superficial image characteristics such as lighting. However, performance increased significantly as more images were added to the average (as faces became more “familiar”). Similar effects have been found for human face recognition – average images of individuals result in faster and more accurate recognition performance (Jenkins & Burton, 2011). Jenkins and Burton suggested that image averages create ‘stability from variation’ (see also Bruce, 1994), since averaging reduces or eliminates pictorial elements that are not characteristic of the person’s identity – for example, lighting direction or camera characteristics (Burton et al., 2005). Jenkins and Burton concluded that average images are a relatively close match to a person’s internal representation of the face, or their stored structural code for a person. Burton et al. (2011) recently extended on this concept, suggesting that our visual system may also encode and use within-person variation during recognition, as opposed to simply discarding or averaging out changeable factors such as variation in lighting. Therefore, our internal representation of a familiar person may incorporate information about how their appearance varies, as well as their average appearance, (Figure 3). On the other hand, unfamiliar faces do not have a stored average – any mental representation of them would be derived from a single view, which would preserve superficial variations. Furthermore, we have no experience with the variation of an unfamiliar face to assist our recognition.

In sum, cognitive and neural models of face recognition have generally focused on the extraction of invariant structural codes as a pathway to identification. Familiar faces are thought to have robust structural codes, derived from multiple exposures to the face. These structural codes most likely code information about variation in face shape and shading, similar to the dimensions extracted using PCA.

\(^1\) An image average is derived by standardising the shape of an individuals’ face in a number of images and taking the mean value for each pixel across all the images. See Burton et al. (2005) for more details.
Unfamiliar face recognition, on the other hand, relies primarily on superficial pictorial information.

Although it is possible that the same type of structural information is extracted from familiar and unfamiliar faces, it is difficult to extract reliable identity information from a single exposure to a face (i.e. when you are unfamiliar with a person). Therefore, the lack of a reliable stored representation (FRU) means that unfamiliar face recognition relies on visually derived semantic information or attention to specific visual idiosyncrasies. Therefore, unfamiliar face recognition can be easily disrupted by changes in lighting, viewpoint and expression.

Interestingly, the focus of the behavioural and neural models reviewed above is on extracting and using the consistent, invariant aspects of a face. Movement, like changes in viewpoint or expressions, is considered a changeable aspect to be filtered out or, at most, factored in only to account for the way a face varies (Burton et al., 2011). However, movement may help us build robust familiar face representations, or act as a cue to identity independent of the structure of a face. The contribution of movement to face recognition will be discussed further in Chapter 3.

Figure 3: Our internal representation of a face may incorporate someone's average appearance and how their appearance varies in our experience. a) An image average and b) an example of the results from a PCA of 48 photos of Harrison Ford. This PCA example shows an average texture (i.e. features and skin tone), mapped onto the first five shape components, with a positive weighting (top row) and negative weighting (bottom row). From Burton et al. (2011).
2.3 Familiar and Unfamiliar Faces: Behavioural Research

This section reviews behavioural research on familiar and unfamiliar faces, and the evidence for qualitative differences between familiar and unfamiliar face processing (see Bruce & Young, 1986). Specifically, this section examines differences in recognition across different viewpoints and expressions, and the effects of various image changes (e.g. degrading or caricaturing a face) on familiar and unfamiliar face recognition.

2.3.1 Viewpoint

Results from studies examining viewpoint changes have lead many researchers to suggest that familiar faces are encoded primarily in a viewpoint invariant manner, whereas unfamiliar or experimentally familiar faces are encoded in a more viewpoint dependent manner (Johnston & Edmonds, 2009).

Experimentally familiar and unfamiliar faces are recognised poorly when the viewpoint is changed between learning and test (Bruce, 1982) or between two images in a same-different matching task (Bruce, Valentine, & Baddeley, 1987). On the other hand, familiar faces are less affected by viewpoint changes. Bruce (1982) showed that recognition performance for familiar faces is slower, but just as accurate when the viewpoint is changed between study and test. Other studies have confirmed that familiar faces can be recognised quite well despite quite large viewpoint changes (Bruce et al., 1987; Eger, Schweinberger, Dolan, & Henson, 2005). It is possible that our ability to generalise between viewpoints for familiar faces is a result of viewing the face at many different angles. Hill, Schyns, and Akamatsu (1997) used experimentally familiar faces to investigate viewpoint generalisation, and found that generalisation between viewpoints was generally good when multiple viewpoints had been learned (Experiment 1). However, when only one viewpoint was learned, recognition performance dropped as the angle between learned and tested view increased (Experiment 2) (see also Jiang, Blanz, & O’Toole, 2009; Lee, Habak, & Wilson, 2010).

As well as changing behavioural patterns, recent neural studies also support the theory that viewpoint changes for familiar, experimentally familiar, and unfamiliar faces can evoke different behavioural and neural responses (e.g., Eger et al., 2005; Pourtois,
These studies will be discussed further in section 2.4.

2.3.2 Expression

The Bruce and Young (1986) model of face recognition proposes that identity and expression information are processed independently. However, there is some evidence that expression information has different effects on familiar and unfamiliar faces. For example, changes to expression information can hinder experimentally familiar face recognition (Bruce, 1982), but displaying happy expressions on familiar faces facilitates recognition (Kaufmann & Schweinberger, 2004).

Early studies found that altering expression information impaired recognition of unfamiliar and experimentally familiar faces. Bruce (1982) found that changing an expression between learning and test impaired recognition of experimentally familiar faces to the same extent as changing viewpoint. Similarly, Bruce et al. (1987, Experiment 2) found that participants were better at matching familiar than unfamiliar faces across viewpoint and expression changes, although this trend only reached significance in the analysis of responses to “different” trials. Conversely, recent studies using identity classification tasks found no reaction time disadvantage for unfamiliar faces showing different expressions (Ganel & Goshen-Gottstein, 2004; Kaufmann & Schweinberger, 2004; Schweinberger & Soukup, 1998; Wild-Wall, Dimigen, & Sommer, 2008). The results from these two sets of studies appear contradictory, but this may be due to different task demands. Bruce (1982) and Bruce and Valentine (1987) asked people to compare two images of unfamiliar faces, and found significant declines when the two exemplars differed in expression. This supports the idea that unfamiliar face representations are dominated by superficial image information (including expression). However, the later studies (Ganel & Goshen-Gottstein, 2004; Kaufmann & Schweinberger, 2004; Schweinberger & Soukup, 1998; Wild-Wall et al, 2008) used a familiarity decision task (i.e., asking participants whether a face belongs to a famous person or not), which did not require people to compare the unfamiliar face to an earlier exemplar. Overall, this evidence leads to the conclusion that learning or matching an unfamiliar face is affected by expression, but deciding a face is unfamiliar is largely
expression independent. In other words, expression information may hinder unfamiliar face recognition, but there is no evidence that it can help it.

While there is evidence that expression information has a negative or minimal effect on unfamiliar face recognition, a different pattern of effects has been found for familiar faces. In old/new recognition tests and matching tasks, changes in expression have a minimal effect on familiar faces. Bruce (1982) found that changes in viewpoint and expression for familiar faces slowed recognition down, but did not affect recognition accuracy, and Bruce et al. (1987) found faster responses in a matching task for familiar than unfamiliar faces whose expressions were changed. However, in familiarity decision tasks (where unfamiliar faces show no effect of expressions) there is evidence that some expressions facilitate familiar face recognition. Endo, Endo, Kirita, and Maruyama (1992) found that familiarity decisions for famous faces were faster for smiling than neutral faces (although the opposite pattern was true for personally familiar faces). Similar studies have found that familiarity decisions for famous or personally familiar faces are fastest for happy expressions, compared to neutral, angry or disgusted facial expressions (Kaufmann & Schweinberger, 2004; Wild-Wall et al., 2008). A smiling expression can also facilitate naming performance for famous faces (Gallegos & Tranel, 2005).

Some authors have proposed that a smiling face increases perceived familiarity (Baudouin, Gilibert, Sansone, & Tiberghien, 2000). However, this explanation does not account for the fact that unfamiliar faces are processed equally well from smiling and non-smiling faces (Kaufmann & Schweinberger, 2004; Wild-Wall et al., 2008). Another possible explanation is that our internal representations of familiar faces preserve information about typical facial expressions. Famous people are often photographed smiling, and it is possible that average expression information is captured in the structural encoding process. Unfamiliar faces have no stored structural representation, which could explain why unfamiliar faces are recognised relatively independently from expression information, when compared to familiar faces (Ganel & Goshen-Gottstein, 2004). The typicality explanation can also account for the fact that Endo et al. (1992) found no advantage for smiling personally familiar faces – we are likely to have seen personally familiar faces displaying a range of different expressions, which averages out
to a neutral pose. Wild-Wall et al. (2008) did find an advantage for personally familiar faces pictured smiling or with neutral expressions, but only when compared to disgusted expressions. Disgust is a relatively rare expression compared to a smile or neutral expression, therefore it may be somewhat easier to recognise personally familiar faces from more typical expressions.

A recent study by Carbon (2008) confirms the importance of typical images in familiar face recognition. Carbon showed participants ‘iconic’ images of famous faces (e.g., John Lennon, Cindy Crawford, and Marilyn Monroe), alongside modified or uncommon images of the same people. Carbon also tested performance with similar images of personally familiar faces. He found that participants were impressively good at naming famous people from iconic images (roughly 75% accuracy), but when the images were modified or uncommon, performance dropped to around 25%. On the other hand, personally familiar faces were recognised equally well from original, modified or uncommon images (around 40% accuracy). These findings suggest that our face recognition skills, particularly for famous faces, are strongly influenced by frequent exposure to typical or iconic images of a person. Therefore, it is quite conceivable that our internal representation of a famous person (our structural code for that person) includes some expression information, based on our history of exposure and our experience of how that face varies (Burton et al., 2011). By contrast, our skills at identifying personally familiar people are much more flexible, because we often see personally familiar people in different situations, and we experience a wider range of variation of expression and poses.

2.3.3 Internal and External Features

Several studies have found a dissociation between familiar and unfamiliar faces based on the ability to recognise a face from the internal or external features. In general, familiar faces are recognised better from internal facial features (the eyes, nose, and mouth area) than external facial features (the face outline, hair, and ears) (Ellis, Shepherd, & Davies, 1979; Young, Hay, McWeeny, Flude, & Ellis, 1985). In comparison, unfamiliar or experimentally familiar faces are recognised equally well from internal and external features (Ellis et al., 1979; Young et al., 1985), or sometimes
better from external than internal features (Bonner, Burton, & Bruce, 2003; Bruce et al., 1999). People are also more sensitive to changes in internal features (specifically the eye region) of familiar than unfamiliar faces (Brooks & Kemp, 2007; O’Donnell & Bruce, 2001).

Young et al. (1985) found an exception to the internal/external feature difference: when the task was altered so the pictures to be matched were derived from the same photograph (rather than two different photographs), the effect of familiarity disappeared. This finding indicates that the internal feature advantage for familiar faces is based on the extraction of invariant structural information about a face, leading to enhanced sensitivity to internal features. Once again, this supports the contention that our internal representation of a familiar face is comparable to an image average, constructed from all our experiences with the person. This image average would smooth out changes to external features such as hairstyle, hair colour or face shape (for example, variations in shape due to age or weight changes), but preserve the less variable inner features of the face.

### 2.3.4 Degraded and Inverted Faces

To this point, the research that has been reviewed has focussed on facial transformations that can and do occur in everyday life (with the exception of some internal/external feature manipulations). However, in order to investigate the concept of structural codes, and the effect of removing certain types of visual information, many researchers have systematically degraded images of faces. Figure 4 shows some common methods of image degradation: blurring, negation, thresholding, pixilating, and inversion.

Blurring usually involves removing the high spatial frequency information from a face. It is thought to impair the encoding of information about individual features (Costen, Parker, & Craw, 1996). Negation involves reversing the brightness contrast in an image, so that it appears like a photographic negative (Knight & Johnston, 1997). Negation makes it difficult to extract shape-from-shading information (Bruce & Young, 1988) and information about the spacing between different facial features (Kemp, McManus, & Pigott, 1990). This spacing information is also known as second-order
configural information (Maurer, Le Grand, & Mondloch, 2002). Inversion involves rotating an image 180 degrees in the picture plane, leaving it upside-down. Inversion impairs the extraction of second-order configural information and interrupts holistic processing \(^2\) (Maurer et al., 2002; Valentine, 1988). However, inversion preserves some feature processing (Searcy & Bartlett, 1996).

![Figure 4: Examples of common image manipulations: a) original image; b) blurring; c) negation; d) thresholding; e) pixilating; and f) inversion.](image)

The effects of blurring, negation and inversion have been thoroughly reviewed in several articles (Bruce, 1994; Collishaw & Hole, 2000; Johnston & Edmonds, 2009), and in general, these reviews have concluded that degrading faces in this manner impairs recognition performance for both familiar and unfamiliar faces. Furthermore, studies

\(^2\) Holistic processing refers to a process where faces are seen as a ‘gestalt’, with little or no part decomposition (Farah, Wilson, Drain & Tanaka, 1998; Maurer et al., 2002)
comparing the effect of these degradations on familiar and unfamiliar faces provide some evidence that the effect of familiarity is purely quantitative.

In one of the few experiments to directly compare famous and experimentally familiar face processing, Collishaw and Hole (2000) used blurring, inversion and scrambling (cutting out and rearranging facial features) to investigate the relative contribution of features and configurations in face recognition. Famous faces were recognised better overall, but the effect of blurring, inverting and scrambling the faces was equal for famous and experimentally familiar faces. This suggests that the stored structural codes for famous faces and the information we extract when viewing a new face are somewhat similar, and include information about individual facial features and their spacing.

Research investigating holistic processing leads to a similar conclusion. The composite effect (Young, Hellawell, & Hay, 1987) is a common measure of holistic processing. Aligning the top half of one face with the bottom half of another to create a ‘composite’ leads to significant delays in identifying the owner of one half of the face, when compared to non-aligned stimuli. Importantly, the composite effect has been demonstrated with famous faces (Young et al., 1987) and unfamiliar faces (Robbins & McKone, 2007), suggesting that holistic processing occurs regardless of familiarity. Similarly, both experimentally familiar (Tanaka & Farah, 1993) and unfamiliar faces (Ramon, Busigny, & Rossion, 2010) undergo show a part-whole effect: individual features are matched or identified more accurately when presented in the context of a whole face than in isolation. Overall, then, it appears that faces are processed similarly regardless of familiarity – features and their configurations are important for recognition, and there is strong evidence for integration of information from across the whole face (holistic processing).

Although there is a large amount of evidence that image degradations such as negation and inversion have a similar negative impact on familiar and unfamiliar faces (see Johnston & Edmonds, 2009), some studies suggest that the exact source of the impairment differs between familiar and unfamiliar faces. Kemp, Pike, White, and Musselman (1996) observed that negation involves two effects – a change to the shading information, and a change to the apparent pigmentation of the face. Kemp et al. (1996)
showed that changes to shading, but not pigmentation information, disrupted familiar face recognition, whereas changes to pigmentation information disrupted unfamiliar face recognition. These findings imply that shape-from-shading information may be more important for familiar than unfamiliar face recognition, whereas surface characteristics such as pigmentation dominate the matching of unfamiliar faces.

Megreya and Burton (2006, 2007) also provided evidence that familiar and unfamiliar face recognition may be qualitatively different. As mentioned above, inversion impairs face recognition for familiar and unfamiliar faces (Johnston & Edmonds, 2009; Valentine, 1988). Inversion is thought to impair the extraction of second-order configuration and holistic information about a face, leaving feature-based processing relatively unaffected (Searcy & Bartlett, 1996; Valentine, 1988; Yin, 1969; Young et al., 1987; but see McKone & Yovel, 2009). Megreya and Burton (2006) correlated participants’ performance across a series of identification tasks, and found that the ability to match upright, unfamiliar faces was positively correlated with the ability to match inverted faces (famous, experimentally familiar, and unfamiliar). However, there was no significant correlation between the ability to match upright unfamiliar faces and upright famous or experimentally familiar faces. Megreya and Burton (2006) concluded that upright unfamiliar faces are processed in a similar manner to inverted faces (i.e., with less contribution from configuration information or holistic processing), and both unfamiliar and inverted faces are processed in a different manner to familiar faces. Megreya and Burton (2007) compared the pattern of responding for familiar and unfamiliar faces, and found that the pattern of hits and false positives in a matching task for unfamiliar faces resembled that for inverted, but not upright, experimentally familiar faces. Once again, they concluded that upright unfamiliar and inverted faces are processed in a similar manner (in this case, resulting in a similar pattern of responses), whereas familiar faces are processed in a qualitatively different manner. One of the advantages of the Megreya and Burton (2006, 2007) studies is that the majority of their conditions involved no or minimal, image degradation. The results of previous studies (e.g., Collishaw & Hole, 2000) may reflect the fact that participants had no choice in what cues to use – when the faces were blurred, for example, participants may have relied on configuration information because they could not match
individual features. This could have lead to abnormal processing strategies, or the use of directed visual processing (Bruce & Young, 1986) for all faces, regardless of familiarity, or what cues would normally be used during recognition. On the other hand, participants in Megreya and Burton’s face recognition tasks (2006, 2007) were able to use all the available cues, since the images were not degraded in the upright condition. This could have resulted in a more natural processing style, which better reflects how unfamiliar faces are matched or recognised in the real world.

The image manipulations shown in Figure 4 are designed to impair performance in recognition tasks. A separate line of research has sought to investigate our internal representation of faces by examining an image manipulation that helps recognition: caricatures. In general, faces that are more distinctive (i.e., further from average) are recognised and learned better than less distinctive faces (Johnston & Edmonds, 2009; Kaufmann & Schweinberger, 2012). Caricatures exaggerate the differences between an average face and an individual, increasing their distinctive properties (Benson & Perrett, 1991). Studies have found that caricatures of famous and personally familiar faces are seen as a better likeness than original images (Benson & Perrett, 1991; 1994; Rhodes, Brennan, & Carey, 1987), and are recognised faster and more accurately than original images (Benson & Perrett, 1994; Rhodes et al., 1987). However, the effect of caricatures requires a level of familiarity with a face: Rhodes et al. (1987) found that unfamiliar and experimentally familiar caricatures were considered poorer likenesses of a face than veridical line drawings.

The differential effect of caricatures on familiar and unfamiliar faces suggests that it takes time to learn distinctive elements of a face. As we become more familiar with a person, our mental representation of that person may also incorporate information about deviation from the average face. This effect is not simply due to the original distinctiveness of an image, which helps recognition of both familiar and unfamiliar faces (Johnston & Edmonds, 2009).

In conclusion, a short review of research on degraded and inverted faces suggests that people use information from features, configurations and holistic processing during familiar and unfamiliar face recognition. Based on this evidence, it appears that the differences between familiar and unfamiliar faces recognition are primarily quantitative.
However, while familiar and unfamiliar face recognition can be disrupted by similar image manipulations (Collishaw & Hole, 2000; Johnston & Edmonds, 2009), a closer examination of the effects of negation suggests that the source of the disruption may be different for familiar and unfamiliar faces (Kemp et al., 1996). Furthermore, unfamiliar faces are processed more similarly to inverted than familiar faces, which suggests that there is a qualitative shift in processing style as a face becomes familiar. Caricature effects for familiar (but not unfamiliar) faces imply that familiarity with a face may increase our knowledge of how a face is different or distinctive compared to other faces (Rhodes et al., 1987), and it is possible that this information is incorporated into our mental representation of a familiar face.

2.3.5 Conclusions from the Behavioural Literature

Overall, evidence from behavioural studies suggests that familiar and unfamiliar faces are represented and recognised in both quantitatively and qualitatively different ways. Viewpoint studies indicate that familiar faces are represented in a viewpoint invariant manner, whereas unfamiliar face recognition is viewpoint dependent. Research on expressions and the internal feature advantage for familiar faces suggests that our experience shapes our stored structural code of a face – as we become more familiar with a person, it is easier to recognise their face in a ‘typical’ expression, and we rely more on the consistent internal features, rather than the changeable external features. Interestingly, personally familiar faces are less susceptible than famous faces to the effects of expression changes and “iconic” images, possibly because our exposure to famous faces is more limited and involves more stereotypical images than our exposure to friends, family and colleagues. Experience with a face may also make us more sensitive to the way that face deviates from the average.

Research using degraded face images confirms that it is possible to extract and use similar types of information for familiar and unfamiliar faces, especially when all other cues are degraded or ambiguous. However, studies on unfamiliar and inverted faces suggest that, in the absence of obvious image degradations, unfamiliar face processing relies on different processes than familiar face processing. This conforms with the Bruce and Young (1986) model of face recognition, which suggests that we extract structural
codes for both familiar and unfamiliar faces, but the codes themselves and the way we recognise a face may differ significantly depending on the familiarity of the face.

2.4 Familiar and Unfamiliar Faces: Neuroimaging and Neuropsychological Research

Behavioural research supports the theory that familiar and unfamiliar faces are processed in a qualitatively different manner. There is also a large amount of neural and neuropsychological evidence that familiar and unfamiliar faces are processed differently. Since there is a recent, comprehensive review of neuroimaging studies comparing familiar and unfamiliar faces (Natu & O’Toole, 2011), and the focus of this thesis is behavioural research, this section will only discuss a small number of studies that highlight the neural distinction between familiar and unfamiliar faces.

Some of the most compelling evidence for separate processing of familiar and unfamiliar faces comes from research on prosopagnosic patients – that is, people who have lost the ability to recognise faces (Steede, Tree, & Hole, 2007). Malone, Morris, Kay, and Levin (1982) presented two cases of prosopagnosia: one patient recovered the ability to process familiar, but not unfamiliar faces; whereas the second patient showed the opposite pattern. This dissociation suggests that the neural areas associated with familiar and unfamiliar face recognition differ. It is possible the differences are hemispheric: Benton (1980) found that deficits in unfamiliar face processing are associated primarily with right hemisphere, particularly posterior, damage; whereas an inability to recognise familiar faces is associated with bilateral damage to the occipito-temporal regions. Mohr, Landgrebe, and Schweinberger (2002) reported a similar finding in healthy patients: a bilateral advantage (indicating interhemispheric cooperation) was present for familiarity decisions when viewing familiar, but not unfamiliar faces.

Studies using positron emission tomography (PET) paint a more complex picture of the neural areas involved in familiar and unfamiliar face recognition. Rossion, Schiltz, and Crommelinck (2003) reanalysed two earlier PET studies and found that the right fusiform face area (FFA) and right occipital face area (OFA) both exhibited higher
activity when viewing unfamiliar, compared to experimentally familiar faces. This finding indicates that familiarity can modulate neural responding even within face-specific areas.

Functional magnetic resonance imaging (fMRI) studies confirm that familiarity can modulate the response of extensive neural areas. Gobbini et al. (2004) showed participants pictures of unfamiliar, personally familiar and famous faces, and found widespread differences in activation for all three comparisons. Areas associated with person knowledge (e.g., the anterior paracingulate cortex), emotion (e.g., the amygdala), and face processing (e.g., the fusiform cortex) were more active when viewing personally familiar than famous faces (see also Leibenluft et al., 2004). Interestingly, the association between activation in the fusiform cortex and familiarity was not linear: famous faces elicited a weaker response than either personally familiar or unfamiliar faces. In a similar study, Leveroni et al. (2000) compared neural responses to unfamiliar, experimentally familiar and famous faces. Famous faces lead to widespread activation in the frontal, temporal, and parietal lobes compared to experimentally familiar or unfamiliar faces, and familiarisation with a face increased responding in the frontal and parietal regions. Like Gobbini et al. (2004), familiarisation did not have a linear effect on responses in the fusiform cortex, but Leveroni et al. found a different pattern of activation to Gobbini et al. (2004): there was no difference between famous and unfamiliar faces, but unfamiliar and famous faces both elicited greater responses from the right fusiform area than experimentally familiar faces.

As mentioned above, studies examining the effect of viewpoint have established that familiar and experimentally familiar faces also elicit different responses during adaptation and priming. Ewbank and Andrews (2008) found adaptation\(^3\) in the FFA was viewpoint invariant for familiar, but not unfamiliar faces. Eger et al. (2005) also found neural differences arising from familiarity. Participants were primed with identical or

---

\(^3\) Neural adaptation refers to a decrease in responding (habituation) after repeated presentations of the same stimulus. Adaptation studies are widely used to determine whether the brain region being studied is selective for a particular attribute: if responding returns to initial levels when a new stimulus (e.g. another facial expression or viewpoint) is presented, it suggests the brain area is sensitive to that attribute. See Natu and O’Toole, 2011, for more details.
different images of famous and unfamiliar people. Priming resulted in decreased activity in the middle fusiform cortex for both familiar and unfamiliar faces, but image-independent priming (i.e., using different prime and test images of the same person) resulted in larger decreases in the anterior than middle fusiform cortex for familiar, but not unfamiliar faces. These results indicate that some regions of the fusiform cortex may be involved in viewpoint-invariant representations, which could support our robust, viewpoint invariant recognition skills for familiar faces.

PET and fMRI studies confirm that familiarisation with a face can change the response of a wide variety of neural areas, but the temporal resolution of these techniques is poor (de Haan & Thomas, 2002). Studies using event-related potentials (ERPs) have investigated whether the time-course of face recognition also depends on familiarity. There are two prominent ERPs related to face recognition – a negative wave around 170 msec after the presentation of a face (N170) (Schweinberger & Burton, 2003) and another negative wave around 200-300 msec after the presentation of a face, which is strongly modulated by the repeated presentation of the same face (N250r) (Schweinberger & Burton, 2003; Tanaka, Curran, Porterfield, & Collins, 2006).

Schweinberger and Burton (2003) and Burton et al. (2011) reviewed the role of the N170 and N250r in face recognition, and concluded that the N170 represents perceptual analysis and structural encoding of a face, whereas the N250r represents the activation of an individual face representation. Interestingly, the N250r, but not the N170, is reduced for unfamiliar faces compared to familiar faces (Schweinberger & Burton, 2003; Tanaka et al., 2006; cf., Marzi & Viggiano, 2007).

Neuroimaging, ERP and magnetoencephalography (MEG) studies have also been used to investigate the contribution of familiarity to face processing. Herzmann, Schweinberger, Sommer, and Jentzsch (2004) examined the topography of the N250r for famous, personally familiar and unfamiliar faces, and found that the N250r for unfamiliar faces peaked in the midfrontal regions, whereas the N250r for familiar faces peaked around the posterior central regions – they interpreted this as further support for a qualitatively different processing mechanisms for familiar and unfamiliar faces. As in the behavioural literature, the effect of different image manipulations has also been examined to determine whether the differences due to familiarity are quantitative or
Chapter 2: Face Recognition and Familiarity

qualitative. Marzi and Viggiano (2007) found that familiarity modulated the N170 for upright faces, and the P250 for inverted faces – this suggests that familiarity can interact with the perceptual encoding and retrieval of faces. However, there was no indication that familiar and unfamiliar faces differed qualitatively, and no evidence that upright unfamiliar faces were processed similarly to inverted familiar faces (cf., Megreya & Burton, 2006, 2007). Harris and Aguirre (2008a, b) investigated the role of holistic and feature-based processing in famous and unfamiliar face recognition. Results from an fMRI analysis indicated that holistic processing resulted in greater adaptation in the right fusiform gyrus for famous, but not unfamiliar faces during a familiarity judgement task (Harris & Aguirre, 2008b). In a subsequent MEG study, Harris and Aguirre (2008a) found that a late (250-400 ms) face-selective component, M400, showed a similar pattern – holistic processing increased the M400 response for famous, but not unfamiliar faces. Harris and Aguirre (2008a, b) concluded that face selective regions of the brain engage in both holistic and feature-based processing, but the use of these processing styles is modulated by familiarity. This lends some support to the idea that the familiarity of a face can lead to qualitatively different processing styles (Megreya & Burton, 2006, 2007).

In conclusion, neuropsychological and neuroimaging evidence support the concept that familiar and unfamiliar face recognition often involve different neural areas (Gobbini et al., 2004; Leveroni et al., 2000). The extra areas activated during familiar face recognition probably reflect person information (e.g., biographical information) and emotional responses, as stated in Gobbini and Haxby’s (2007) model of familiar face processing. When the same neural areas are activated (e.g., the fusiform areas or the OFA), familiarity may modulate the level of activation, but not in a linear fashion (see Natu & O’Toole, 2011, for a more detailed analysis). Research using ERPs and MEG supports the concept that familiar and unfamiliar faces elicit different patterns of responding, but it is unclear whether the difference is purely quantitative (Marzi & Viggiano, 2007) or whether these findings indicate the use of different processing styles for familiar and unfamiliar faces (Harris & Aguirre, 2008a, b). Overall, despite significant behavioural differences, there is little decisive evidence that familiar and unfamiliar faces are presented in qualitatively different ways in the brain.
2.5 Familiar and Unfamiliar Faces: Conclusions

Familiar and unfamiliar faces show important behavioural and neural differences. Evidence from behavioural studies suggests both qualitatively and quantitatively different processing styles for familiar and unfamiliar faces – familiar face recognition is relatively viewpoint invariant, and reflects our aggregated experience with a face, whereas unfamiliar face recognition is viewpoint dependent, and easily disrupted by superficial changes. On the other hand, neuroimaging research reflects primarily quantitative differences. Familiarity can modulate the responses of brain regions associated with face processing, emotions, and person knowledge, and may also influence the response of areas responsible for viewpoint-invariant and holistic processing.

One of the most important conclusions that can be drawn from face recognition models and research is that invariant, viewpoint-independent information plays an important role in face recognition (Bruce & Young, 1986; Haxby et al., 2000). As we become familiar with a face, we are better able to extract or use information that is diagnostic of identity (Bruce, 1994; Jenkins & Burton, 2011). Our internal representation of a face becomes more viewpoint-independent, and may encode information about typical expressions or other visual cues. We are also able to predict which information is likely to vary (i.e. what parts of our representations are due to changeable or image-based characteristics) and which information will be consistent (i.e., what cues are truly diagnostic of identity, and are likely to appear frequently in our encounters of a person) (Burton et al., 2011). This raises the important question: can we use the way a face changes as a cue to identity? Specifically, does the way a person moves form a part of our internal face representation? Furthermore, if viewpoint invariance is important for recognition (as many studies suggest), does head movement (which can provide multiple viewpoints) improve our recognition of familiar faces, or make it easier to acquire or match a representation of an unfamiliar face? The next chapter summarises and analyses research that has used moving faces as stimuli, in order
to examine whether movement can act as a cue to recognition for familiar and unfamiliar faces.
Chapter 3

Movement and Face Recognition
CHAPTER 3: MOVEMENT AND FACE RECOGNITION

Chapter 2 reviewed the role of familiarity in face recognition. Another important, but under-researched factor is movement – in everyday life, we see people in motion. They may be talking, smiling or frowning, nodding and shaking their heads, or simply moving around us. Research on whole-body movement has shown that we are remarkably sensitive to the social cues carried in motion – for example, we can judge someone’s gender or emotions, and even identify a person, based on their walk alone (Blake & Shiffrar, 2007). Faces, too, can convey important information via motion – moving faces can supplement speech information (Rosenblum & Johnson, & Saldana, 1996), help emotion judgements (Bassili, 1979), and support gender classification (Hill, Jinno, & Johnston, 2003). The focus of this chapter is on how we can use movement information to help face recognition, and what it is, precisely, that movement “adds” to face recognition.

The literature on movement and face processing has generally focused, quite reasonably, on research involving face recognition or identification. However, the first half of this chapter adopts a wider perspective, and examines the research on whole-body biological motion, and facial motion in non-recognition tasks (e.g., gender or emotion classification), to highlight the different types of information that movement can provide. The first section of this chapter (section 3.1) provides a broad overview of the role of movement (both whole-body and face-only movements) in the perception of actions and other social information. Section 3.2 reviews an influential model of movement in face recognition, with a focus on supporting evidence from whole-body biological motion research and the facial emotion and speech-processing literature. The second half of the chapter adopts a narrower perspective, and summarises the body of research that has investigated whether seeing a face in motion improves face recognition. Section 3.3 will provide an overview of current research on movement-based recognition using famous, personally familiar, experimentally familiar and unfamiliar faces. The final section (section 3.4) will examine the relationship between
familiarity and movement in face recognition, and emphasise the importance of methodological factors in movement-based face research.

3.1 Biological Motion

Face motion is an example of class of stimuli known as biological motion, which has been defined as “the visual perception of a biological entity engaged in recognisable activity” (Pelphrey & Morris, 2006, p. 136). Although biological motion encompasses animal and human movements, and the movements of individual body parts (such as faces or arms), the majority of research on biological motion has focused on whole-body human motion. Traditionally, much of this research has used point-light-display (PLD) stimuli (Figure 5). PLDs are simplified visual stimuli that show the movement of small dots attached to various parts of the body – most commonly, the limbs and joints (Johansson, 1973). Generally, a static frame from a PLD will appear to be a random display of dots, whereas a sequence of frames will convey the impression of biological motion.

Biological motion is highly salient– even when the dots on a PLD are spatially scrambled, extremely brief, or embedded in randomly moving “noise” dots, perceivers are quite good at recognising biological motion (see Blake & Shiffrar, 2007, for a review). Developmental and cross-species studies also support the idea that the detection of biological motion perception is an important skill. Human infants as young as two days old show a preference for a display containing an upright, walking hen over a display containing random motion (Simion, Regolin, & Bulf, 2008), and four-month-old infants look preferentially to an upright human walker over an identical inverted display (Fox & McDaniel, 1982). Newly hatched chicks also show a preference to approach biological motion (Vallortigara, Regolin, & Marconato, 2005), and cats can discriminate between a point-light cat and scrambled dot sequences (Blake, 1993).
3.1.1 Whole-body Biological Motion

The perception of biological motion is an important social skill, because motion can carry information about a person’s actions, emotions, gender, and speech. In the real world, we may use biological motion cues to avoid walking near an angry stranger, judge the weight of a box someone is lifting, or supplement poor hearing in a noisy environment. In experimental settings, studies on whole-body biological motion have shown that people can recognise locomotory movement (e.g., walking, jumping), instrumental actions (interacting with objects; e.g., hammering), and social actions (interacting with people; e.g., dancing), purely from movement patterns (Dittrich, 1993). It is also possible to discriminate between different styles of the same basic movement, such as a tennis serve (Pollick, Fidiopiastis, & Braden, 2001). Action categorisation is possible even with extremely short clips (Norman, Payton, Long, & Hawkes, 2004, Experiment 1), when part of the PLD is occluded (Norman et al., 2004, Experiment 2), and from a number of different viewpoints (Prasad & Shiffrar, 2009). Observers can use biological motion to infer information about the world around the actor: for example, how far an object is thrown, or the expected weight of a box to be lifted (Runeson & Frykholm, 1983). In addition to action and object perception, research has established that the information carried in whole-body PLDs is sufficient to allow people to classify...
emotions and detect deception (Dittrich, Troscianko, Lea, & Morgan 1996; Runeson & Frykholm, 1983).

It is clear that biological motion carries important social information, akin to what Haxby et al. (2000) described as the changeable aspects of a face – that is, information that varies over time, such as emotion and actions. However, biological motion cues can also convey information about invariant aspects of a person. Observers are well above chance levels when asked to classify the gender of a PLD (Pollick, Kay, Heim & Stringer, 2005), and surprisingly accurate at identifying familiar people from the way they move. Cutting and Kozlowski (1977) asked seven undergraduate participants to identify their friends from whole-body PLDs and found that participants were remarkably accurate, correctly identifying their friends on 38% of trials (chance was 16.7%). When asked to report how they made a decision, most participants reported using movement-based cues, such as speed or rhythm of the walker.

Subsequent studies have confirmed that participants can identify or discriminate between familiar people from their walking patterns (Jokisch, Daum, & Troje, 2006) and from a variety of actions such as jumping, playing ping-pong, and dancing (Loula, Prasad, Harber, & Shiffrar, 2005; Prasad & Shiffrar, 2009). In general, participants were poorer when asked to identify or discriminate strangers than familiar people (Loula et al., 2005; Prasad & Shiffrar, 2009), which suggests that some level of familiarity is important for the task. However, there is evidence that familiarity with a person’s walk is acquired relatively rapidly. Even using relatively short stimuli (2 secs), participants can learn to name six PLD walkers with 100% accuracy after an average of 8.7 trials (Stevenage, Nixon, & Vince, 1999). Furthermore, learning a person’s walk does not require exposure to the full body – several studies have successfully trained and tested participants using PLDs alone (Stevenage et al., 1999; Troje, Westhoff, & Lavrov, 2005; Westhoff & Troje, 2007). Furthermore, Stevenage et al.’s study found no advantage for learning fully illuminated bodies compared to PLDs, which suggests that participants were not relying on information such as size, height or body contours.

In summary, whole-body biological motion can convey a number of social cues. Most importantly for this thesis, whole-body biological motion carries enough information to identify familiar people, and to learn new people. In other words,
biological motion can contribute to both familiar and experimentally familiar person recognition. These results will be examined further in section 3.2.

3.1.2 Facial Motion

3.1.2.1 Types of facial movement. Although the majority of research on biological motion perception has focused on whole-body stimuli, there is also evidence that movement is important for a variety of face processing tasks. As mentioned above, faces can move in a number of ways: people move their faces when they talk, laugh or frown, or when they shake and nod their head. We also experience movement when people move around us, or when we move around other people. All of these movement cues can be divided into two basic categories: rigid and non-rigid movement. Movements of the head, such as nodding or shaking, or the relative motion that occurs as people move around each other, are referred to as rigid movements. Rigid head movement provides constantly changing perspectives or viewpoints of the face and head, but the face itself does not move. Movements of the face, such as speech and expressions, are referred to as non-rigid movements. Non-rigid movements involve temporary changes to the facial features, but they do not change our viewpoint of the head. Most everyday encounters include multiple types of rigid and non-rigid head movement – for example, someone might frown while shaking their head, or smile while talking and tilting their head to the side.

3.1.2.2 Research on facial movement. The point-light technique (Figure 5), along with many other stimulus creation techniques, have been applied to faces to examine the contribution of motion information to face detection (Bassili, 1978), emotion recognition (Bassili, 1978; 1979; Bruce & Valentine, 1988), speech perception (Rosenblum et al., 1996), gender classification (Hill et al., 2003) and identification (Bruce & Valentine, 1988; Rosenblum, Niehus, & Smith, 2007; Rosenblum et al., 2002).

Early research established that facial motion is sufficient to categorise emotions. Bassili (1978) scattered 100 white dots onto the faces of actors and filmed them while they expressed six basic emotions (happy, sadness, disgust, surprise, fear, anger, and interest). Participants were accurate on 33.3% of the trials, and a subsequent experiment
established that participants were significantly more accurate when identifying emotions from moving displays (both full face and PLD) than static images (Bassili, 1979). Other facial expressions, for example, conversational expressions, also show a movement advantage. Expressions such as confusion, agreement, and thinking, are perceived more accurately when shown in motion, even when the amount of static information is equated across conditions (Cunningham & Wallraven, 2009). In some instances, the motion of a single facial area is sufficient to carry conversational information – for example, ‘clueless’ and ‘thinking’ expressions can be identified from the eye region (Nusseck, Cunningham, Wallraven, & Bulthoff, 2008).

Facial motion also contributes to speech perception. In the McGurk effect (McGurk & MacDonald, 1976), participants’ auditory perception of a syllable is influenced by their visual perception of a discrepant syllable (e.g. auditory “ba” and visual “ga” perceived as “da”). Seeing a face in motion can also enhance perception of speech-in-noise, even when the face is represented as a PLD. Rosenblum et al. (1996) demonstrated that minimal PLDs, which only presented movement information from the lips, were still able to enhance speech-in-noise thresholds by up to 7.3 dB compared to an auditory only condition. Interestingly, it is not just the movement of the mouth that affects speech perception. Viewing speech with natural rigid head motion also improves speech intelligibility – exaggerating rigid motion has been shown to impair performance on a speech perception task (Munhall, Jones, Callan, Kuratate, & Vatikiotis-Bateson, 2004).

Finally, like whole-body biological motion, face motion can also be used to extract “invariant” aspects of a face, such as gender and identity. Observers can classify the gender of a PLD or animated face, even when shape information is averaged (Hill & Johnston, 2001; Hill et al., 2003; Morrison, Gralewski, Campbell & Penton-Voak, 2007). Several studies have also found that faces can be named or matched based on identity when movement cues are isolated using PLDs or facial animation (Bruce & Valentine, 1988; Hill & Johnston, 2001; Rosenblum et al., 2007). The ability to use facial motion as a cue to identity will be the focus of the remainder of the chapter.
3.2 A Model of Movement and Face Recognition

The research reviewed to this point has established that whole-body and facial biological motion carry social and identity information. However, it has not established why movement is useful when trying to identify a person or their actions or emotions. Roark et al. (2003) proposed a model of movement-based face recognition (Figure 6), incorporating three ways that motion can contribute to face identification. First, movement can help a viewer extract detailed structural information about the face— that is, movement may provide extra information about the shape of the face and head. Second, people may be able to identify a person based on the characteristic way they move— movement contains dynamic cues, such as timing information, which may be useful for identification. Third, movement can act as a social cue, which may attract attention towards the face or distract people from encoding identity information.

Figure 6: Roark et al.’s (2003) neural systems model of movement and face recognition. Adapted from Roark et al. (2003).
3.2.1 Structural Information

Movement can help face recognition by providing access to static structural information about the face and head. It is possible for an observer to use structure-from-motion processes to extract three-dimensional information about complex shapes and objects, such as a moving head. These cues could help build a robust, accurate representation of the face and head, which could enhance subsequent recognition performance. Roark et al. (2003) refer to this as the representation enhancement hypothesis.

3.2.1.1 What is structure-from-motion? When we see an object move (or when we move around an object), it can help us construct a better mental representation of the object. In general, this rigid rotational movement of an object provides additional viewpoints, which can make it easier to subsequently remember or match that object (see Tarr & Cheng, 2003, and Ullman, 1998 for a review of viewpoint dependent and viewpoint specific object recognition). In addition to providing additional viewpoints, movement information allows a perceiver to extract information about the three-dimensional characteristics of an object – a process referred to as structure-from-motion.

Structure-from-motion occurs when an object rotates in depth, inducing a perception of three-dimensional shape. The structure-from-motion process uses movement information over a number of spatial locations to construct and update a mental representation of the 3D surface of an object (Andersen & Bradley, 1998). Early research examined structure-from-motion by projecting shadows from solid objects and wire shapes onto a translucent screen (Wallach & O’Connell, 1953). The projections were only perceived as three-dimensional when they were shown in movement, but the information was sufficient for participants to match shapes or reconstruct wire figures with high levels of accuracy. Subsequent research has established that structure-from-motion does not require solid shapes. Sperling, Landy, Dosher, and Perkins (1989) used a shape identification task (identifying patterns of hills and valleys on a flat surface), and found that participants were able to accurately identify a large number of complex shapes from displays of dots, even when a single frame contained no shape cues.
3.2.1.2 Structure-from-motion in body and face perception. Structural cues are important in biological motion perception as well as object perception. In fact, several studies have suggested that structure-from-motion information plays a strong role in whole-body gender discrimination tasks. Cutting, Proffitt, and Kozlowski (1978) proposed that gender categorisation of point-light walkers uses structural information – specifically the ratio of the shoulder width to the sum of the shoulder and hip widths. Although this cue is apparent from static information when viewing the body from the front or back, the extraction of shoulder and hip ratios from a side view requires structure-from-motion. There is strong evidence that people use structure-from-motion information when making gender judgements: in synthetic walkers, varying this structural cue from 0.2 to 0.8 changes perceivers’ judgements of the “maleness” of a walker from 20% to 80% (Cutting, 1978). There is also evidence that structural cues are important in identification tasks. Troje et al. (2005) trained participants to identify point-light walkers, and then tested identification performance for stimuli that had been size-, shape- or walking frequency-normalised. Size normalisation did not affect identification, but both shape and frequency normalisation decreased performance. Overall, these findings suggest that the structural information derived from biological motion may be sufficient, but not necessary, in gender classification and identification tasks. Given that participants were able to identify PLDs at above-chance levels even after they were shape-normalised (Troje et al., 2005), it is unlikely that structural information was the only movement-based cue that contributed to identification.

Recent studies have confirmed that structure-from-motion information is also sufficient to support face perception. Farivar, Blanke, and Chaudhuri (2009) tested matching performance for purely motion-defined faces, which consisted of a uniform density random-dot texture. When the image was presented as a static frame, no facial information was available, but when the image rotated, the displacement of the random-dot texture resulted in the perception of structure-from-motion. Even though the faces were unfamiliar, participants were able to match a motion-defined face to a static, shaded face at well above chance levels. The motion-defined faces were also matched significantly more accurately than static images. Farivar et al.’s findings confirm that structure-from-motion cues can be useful in face matching tasks.
3.2.1.3 The neural basis of structure-from-motion. Currently, the neural basis of structure-from-motion processing is unclear, however it appears that area MT is involved (Andersen & Bradley, 1998), along with area MST (see Farivar, 2009, for a review). Roark et al. (2003) proposed that structure-from-motion information for faces is processed in MT and then fed back to the fusiform face area (FFA), to facilitate static face identification (Figure 6). The fact that motion-defined faces can be matched to static face images (Farivar et al., 2009) supports the idea that motion-defined shape information can be used by static face processing areas in identity-based tasks. Structure-from-motion defined faces also activate the FFA, whereas structure-from-motion defined random objects do not elicit an FFA response (Kriegeskorte et al., 2003), which supports the contention that the neural areas responsible for structure-from-motion in faces are linked in some way to areas that process static face images.

3.2.2 Characteristic Dynamic Information

The second way movement may help face recognition is via “dynamic facial signatures” (Roark et al., 2003 p. 20). Put simply, it may be possible to identify someone based on the characteristic or idiosyncratic way they move. For example, a person may frequently shake their head as they talk, or they may have a peculiar way of moving their eyebrows or frowning, and it is possible that someone could use that face and head motion as a cue to identity. Roark et al. (2003) refer to this as the supplemental information hypothesis.

3.2.2.1 Dynamic information in body perception. The research on whole-body biological motion perception supports the hypothesis that people are highly sensitive to characteristic dynamic information, and can use it to make subtle discriminations such as emotion categorisation (Dittrich, et al., 1996). However, it is possible that participants in Dittrich et al.’s study were able to categorise the emotions purely because the PLDs contained structural cues, which conveyed emotional body positions. To address the

---

4 In the context of emotion, gender and speech-based tasks, the “characteristic” motion is typical of the whole category (e.g., “happy”), not an individual.
contribution of characteristic dynamic cues (as opposed to structural information) in emotion recognition, Pollick, Paterson, Bruderlin, and Sanford (2001) showed participants a PLD of an arm making “knocking” movements. Participants were able to discriminate between 10 emotions, and were above chance even when the points were phase-shifted (each point started at a different, random stage of the movement) and inverted (presented upside-down). By analysing the confusions and movement properties of the actions, Pollick et al. (2001) were able to conclude that motion cues could be sufficient to support emotion categorisation, even in the absence of structural information.

Characteristic dynamic information may also be important in whole-body gender discrimination tasks. In contrast to Cutting et al. (1978), Mather and Murdoch (1994) suggested that lateral body sway (sway of the hips and shoulders) is the dominant cue for gender discrimination. Mather and Murdoch’s study found that when body sway was available as a cue (i.e., frontal views), participants were 79% accurate at gender classification. When structural cues were mismatched with motion cues, gender discrimination performance only dropped 6%, indicating that participants were relying more heavily on the movement cues than the structural information. Furthermore, androgynous body shapes can be classified using body sway alone. Troje (2002) used linear analysis to examine the gait patterns of 20 male and 20 female walkers, then asked participants to perform a gender discrimination task on shape-averaged walkers (with individual walking characteristics) or motion-averaged walkers (with individual structural information). In line with Mather and Murdoch’s results, Troje found that depriving participants of individual structural information did not significantly reduce performance compared to the original PLDs, but depriving them of individual motion patterns significantly impaired gender discrimination.

The studies listed above confirm that there are patterns of movement that are characteristic of different categories, such as emotions and gender. However, Roark et al.’s (2003) hypothesis suggests that we are also sensitive to characteristic movements on an individual level. Sensitivity to individual characteristic motion patterns has been demonstrated for whole-body biological motion displays: as mentioned above, Troje et al. (2005) found that normalising the walking frequency of a PLD (i.e., removing one
characteristic dynamic cue) decreased performance in an identification task. However, even when walking frequency and shape were normalised, recognition rates remained almost six times higher than chance, implying that participants were able to use other, non-frequency based motion cues to identify the walkers.

Westhoff and Troje (2007) identified a number of these cues. They subjected the shape- and frequency-normalised walking patterns of 20 individuals to a temporal frequency Fourier analysis, and tested whether participants could discriminate between a subset of individuals when the majority of frequency components were eliminated. Participants were able to discriminate between individuals when only the first or second harmonics were present (despite the fact that the second harmonic only captures an average of 6.4% of the variance in walking pattern): these coded primarily for horizontal movements of the legs and arms, and vertical movement in the gait pattern respectively. Westhoff and Troje also examined identification when phase (timing-based) or amplitude (spatial displacement) information was disrupted. Participants could identify walkers when either the timing-based or the spatial displacement information were normalised, but stimuli that contained spatial displacement information were identified significantly more accurately than stimuli that contained timing-based information.

Further evidence for the importance of characteristic dynamic cues comes from studies of exaggeration. Hill and Pollick (2000) trained participants to recognise individuals’ arm movements, and then tested them on temporally exaggerated movements made by the same actors. Recognition levels were higher for increasing levels of exaggeration, suggesting that time-based cues were important for identification. However, there was no evidence that absolute duration played a role in identification. Hill and Pollick suggest that relative duration (which may capture the rhythm of the movement) could be a critical factor in biological motion recognition.

3.2.2.2 Dynamic information in face perception. The research on whole-body biological motion provides compelling evidence that we are sensitive to very subtle dynamic cues. However, while frequency and phase information may be important for a rhythmic activity such as walking, it is more difficult to apply these concepts to facial movement. Research on dynamic cues in facial movement, therefore, has generally
concentrated on absolute timing cues – for example, comparing emotion perception when the movement cues are slowed or speeded. Kamachi et al. (2001) investigated whether changing the duration of a morphing sequence between emotions improved emotion categorisation. Morphs from neutral to happy and surprised expressions were categorised poorly from slower clips, whereas morphs from neutral to sad faces showed the opposite pattern. A similar relationship between time and emotion was found when participants were asked to rate the emotional intensity of moving images, but no such effect appeared for static images that were displayed for different durations. Subsequent studies have established that temporal characteristics play an important role in emotion perception: presenting slowed, speeded and disrupted motion impairs categorisation performance (Bould, Morris, & Wink, 2008; although Pollick, Hill, Calder, & Paterson, 2003 did not find equivalent effects for slowed and speeded PLDs) and ratings of “naturalness” (Sato & Yoshikawa, 2004). These findings offer support to the theory that dynamic properties, such as speed, are important for face perception as well as whole-body motion perception.

3.2.2.3 The neural basis of dynamic cues. Haxby et al.’s (2000) neural model of face processing proposed that the changeable aspects of the face, such as lip movement and eye gaze, are processed in the superior temporal sulcus (STS). There are many other studies that support the contention that the STS plays an important role in social perception, particularly biological motion processing (Allison, Puce, & McCarthy, 2000; Puce & Perrett, 2003). Based on this evidence, Roark et al. (2003), and later Gobbini and Haxby (2007), suggested that the STS might also encode characteristic face and head movements, and individuals may use this information as an alternate route to identification when static cues are unavailable or ambiguous (Figure 6).

Neuropsychological evidence supports the theory that characteristic motion patterns offer an alternate route to identification. Lander, Humphreys, and Bruce (2004) tested a prosopagnosic patient, H.J.A., and found that showing a face in motion (either rigid or non-rigid) improved matching performance. However, there was no benefit if only one of the images was moving, suggesting that H.J.A. was matching movement patterns, not extracting structural cues. In a similar experiment, Steede, Tree, and Hole
(2007) found that prosopagnosic patient C.S. could match and learn moving faces as well as control participants. This result is particularly interesting because the movement was projected onto a shape-averaged animated head, which did not preserve any individual structural cues. The authors concluded that C.S. must have been learning movement patterns, although it is possible he was focusing on different cues to the matched control participants. Both of these studies support the contention that characteristic motion patterns can be processed somewhat separately from the normal face processing system, which was profoundly impaired in both patients.

### 3.2.3 Social Information

Roark et al.’s (2003) third hypothesis explored the possibility that social signals carried in motion could affect identification both positively and negatively – the “motion as a social signal hypothesis” (p.21). For example, if an observer was using the movement of a face to lip-read they may not pay attention to the identity of the face, which could hinder subsequent recognition. On the other hand, it is also possible that movement can attract attention to a face or encourage a person to engage in deeper processing of the face, which may improve subsequent recognition performance (Bower & Karlin, 1974).

The idea that movement can enhance attention to important features of the face and increase general arousal has been explored in the emotion processing literature. Ambadar, Schooler, and Cohen (2005) found that participants were better at identifying subtle emotions from dynamic than static or multi-static displays. However, participants were equally accurate when shown the first and last frames of the emotion sequence. Ambadar et al. suggested that the dynamic unfolding of the emotion itself was not important – instead, motion enhanced the perception of change from one expression to another, which in turn improved emotion perception. It is important to note that these results do not discount the possibility that dynamic cues such as duration (Kamachi et al., 2001) also play an important part in facial emotion processing. However, Ambadar et al.’s results highlight the fact that motion can also act to enhance attention to social cues, which can in turn improve processing of a stimulus. This conclusion is supported
by fMRI research that has found higher general arousal for moving than static facial expressions (Trautmann, Fehr, & Herrmann, 2009).

Despite the importance of motion for extracting a number of social cues, many researchers have overlooked the social signals hypothesis when explaining the presence (or absence) of a movement advantage for face recognition. While many studies have found that moving faces are better recognised, matched or learnt than static images, some studies have failed to find any benefit of viewing a face in motion (see section 3.3 for a review). It is possible that a lack of attention to identity (in favour of attending to social cues) could account for some studies that have failed to find a movement advantage (this will be discussed further in the section 3.3.2, which reviews studies on unfamiliar faces). Conversely, the experiments that have found a movement advantage may have engaged the participants more effectively – for example, some experiments have used incidental learning paradigms, which encourage participants to engage with a face by making social judgments about it (e.g., Knappmeyer, Thornton, & Bulthoff, 2003; Pilz, Thornton, & Bulthoff, 2006, Pilz, Bulthoff & Vuong, 2009). Pilz et al. (2009) found direct support for this hypothesis: participants in their study showed a movement advantage for faces learned in motion when they were asked to complete an in-depth questionnaire about the personality of the face, but not when they were simply asked to rate faces along a number of personality dimensions.

In some cases, participants may not need to engage specifically with the target face to benefit from social signals. Bruce et al. (2001, Experiment 2) found that a simple 30 or 60 second exposure to a moving face did not increase identification performance when compared to a static control condition. However, if participants were encouraged to socialize while they viewed the moving faces (Experiment 3), their performance was significantly better than groups that had not been familiarized with the faces, or who had not socialized during familiarization. These findings support the hypothesis that the movement advantage may depend on the manner of exposure to the face. It is important to note, though, that Bruce et al.’s (2001) Experiment 3 did not include a static control condition – it is unclear whether social interaction would have had the same benefit for static faces, or whether the effect is movement-specific.
3.2.4 Combining Structural and Dynamic Information

Roark et al.’s (2003) model of movement and face recognition presents the representation enhancement hypothesis, the supplemental information hypothesis and the social signals hypothesis as separate but non-exclusive hypotheses. This means that when we see a face in motion, it is possible to benefit from all three mechanisms simultaneously. However, it is unclear whether one mechanism is more influential overall, or whether the contribution of structural, dynamic and social cues varies depending on the task. Previous studies have rarely examined the effect of social cues in tandem with the use of structure-from-motion or characteristic motion patterns, so this section will focus on the relative contribution of structural and dynamic cues in gender and identity classification tasks.

As reviewed above, studies on whole-body gender discrimination have found that both structural and movement-based cues can be used to make accurate gender judgements. Some studies have also sought to determine which of these cues is more perceptually dominant. In a meta-analysis of gender discrimination studies, Pollick et al. (2005) found results that were consistent with observers using structural cues, as suggested by Cutting et al. (1978). However, they were careful to note that their results do not confirm that people are actually using this cue in gender categorisation tasks. Studies that have tried to isolate dynamic and structural information (Troje, 2002) or have displayed competing structural and dynamic cues (Mather & Murdoch, 1994) suggest that dynamic cues are the dominant source of information for gender classification, whereas structural information has “only a comparatively weak effect on both human and artificial gender classification” (Troje, 2002, p. 382).

The relative contribution of structural and dynamic information has also been investigated in whole-body biological motion identification tasks. Superficial structural cues, such as size and body contour, do not appear to play a large role in identification. As mentioned above, participants learn to identify people from their walk equally quickly when they are shown fully illuminated videos, or when contour information is eliminated through the use of PLDs (Stevenage et al., 1999). Troje et al. (2005) also found that identification performance of PLDs was not affected by size normalisation. However, structural information does play some role in identification – Troje et al. also
found that normalising structure or walking frequency resulted in an equal drop in performance. Importantly, though, when both structure and walking frequency were normalised, performance was still well above chance, suggesting that the majority of information used for identification is dynamic, not structural.

There is little research on the relationship between the neural bases of structure-from-motion and characteristic motion patterns. Several studies have established that moving faces elicit enhanced neural activity (Sato, Kochiyama, Yoshikawa, Naito, & Matsumura, 2004), and one recent study suggested that moving face stimuli are more reliable at activating the core face processing areas identified in Haxby et al.’s (2000) model of face processing, particularly the STS (Fox, Iaria & Barton, 2009). Unfortunately, these studies did not specifically study identification using structure-from-motion or characteristic motion patterns. There are some neuropsychological studies, however, that indicate the structure-from-motion and characteristic motion patterns are processed by distinct neural pathways. Lander et al. (2004) and Steede et al. (2007) found evidence that prosopagnosic patients were using facial motion patterns to match or identify unfamiliar and experimentally familiar faces (see section 3.2.2.3). A third neuropsychological study suggests that structure-from-motion cues are processed separately, in a system most likely linked to static face processing. Farivar et al. (2009) found that prosopagnosic patient P.S. was unable to match motion-defined faces (rigid rotational motion only) in an identification task, despite being highly accurate when performing the same task with chairs as stimuli. Taken together, these studies support the theory that characteristic motion patterns are processed separately from static faces (at least initially), likely in the STS, whereas structure-from-motion information supplements static face processing.

3.2.5 Conclusions From the Model

There is a large amount of empirical support for Roark et al.’s (2003) model of movement and face recognition. Studies using biological motion confirm that we are capable of using structure-from-motion information in a range of tasks, including face perception, and we are also highly sensitive to dynamic information such as the frequency, amplitude and phase of movement in a point-light walker; and the speed and
duration of emotional facial expressions. Social cues have generally been overlooked in
the research on movement and face recognition, but there is evidence that movement can
attract attention to a face, and simple social interaction during a learning experiment
may improve subsequent recognition performance.

One outstanding question is whether structure-from-motion and dynamic cues
have equal weight during movement-based face recognition. Investigations on whole-
body biological motion indicate that characteristic motion patterns (e.g., body sway or
individual walking amplitude) are more influential than structural information for gender
classification and identification, but to this point no studies have used analogous
methods to investigate the relative contribution of structural and dynamic cues to face
perception. In fact, relatively little research has been conducted into the independent
contribution of structural and dynamic information in movement-based face recognition,
possibly because the effect of movement has proven to be somewhat inconsistent across
different studies.

The following section reviews the evidence for a movement advantage in face
recognition, specifically looking at familiar and unfamiliar faces. Where a movement
advantage is present, the contribution of structure-from-motion, characteristic movement
patterns and social signals is assessed.

3.3 Recognising Moving Faces: The Movement Advantage in Familiar and
Unfamiliar Face Recognition

There is clear evidence that people can use full-body biological motion as a cue to
identity, and a number of studies have examined whether people can use facial motion in
the same way. This section reviews evidence for a movement advantage in face
recognition, and attempts to relate the numerous, sometimes conflicting, findings to
Roark et al.’s (2003) model of movement and face recognition. This section also
reintroduces the concept of familiarity – studies of movement in face recognition are
examined separately for famous, personally familiar, experimentally familiar and
unfamiliar faces. To highlight the potential differences between the use of movement for
familiar, experimentally familiar and unfamiliar face recognition, sections 3.3.1.3 and
3.3.2.3 examine the origins of the movement advantage for familiar face recognition (section 3.3.1.3) and experimentally familiar and unfamiliar face recognition (3.3.2.3). In order to facilitate comparisons between familiar and unfamiliar face research, Tables 1 – 4 present a simplified overview of the studies reviewed in this section, including details of their stimuli, task and findings.

3.3.1 Familiar Faces

3.3.1.1 Famous faces. The majority of studies examining movement in familiar face recognition have used famous faces as stimuli (Table 1). As can be seen in Table 1, the stimuli used in famous face studies have usually been derived from television shows. They tend to show short (up to 5 secs) clips of natural face and head movements, which incorporate both rigid and non-rigid motion. Due to the similarities between the stimuli, it is quite easy to compare results across the different studies.

In general, famous face studies support the idea that motion can improve recognition. Knight and Johnston (1997) tested recognition of moving and still famous faces that were presented either non-degraded or negated, and either upright or inverted. They found that moving images were recognised significantly better than still images, but only when the faces were presented both upright and negated. They proposed that movement could provide an important cue to identity when static cues were degraded. Knight and Johnston’s findings also suggest that the extraction of characteristic movement patterns or structure-from-motion information requires the face to be presented upright – indicating that the movement advantage may be related to the processing of whole faces and their configurations, rather than recognising the movement or structure of individual facial features (Valentine, 1988). However, a subsequent study by Lander, Christie, and Bruce (1999) found a movement advantage for inverted famous faces, which suggests that movement may help the recognition of highly familiar faces that have been degraded in any way, not just via negation.
### Table 1:
*An overview of previous research involving familiar faces and motion*

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental/image manipulation</th>
<th>Movement</th>
<th>Length of clip</th>
<th>Major findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bruce &amp; Valentine, 1988, Experiment 2</td>
<td>Recognition of friends from moving and static PLDs.</td>
<td>Rigid and non-rigid scripted</td>
<td>5 s</td>
<td>Participants could identify friends from both rigid and non-rigid moving displays. Moving PLDs recognised better than static PLDs.</td>
</tr>
<tr>
<td>Knight &amp; Johnston, 1997</td>
<td>Naming famous faces from moving/static images. Inversion, negation.</td>
<td>Natural: from television</td>
<td>5 s</td>
<td>Moving, upright negated faces recognised better than still, upright negated faces. No effect of movement when faces were inverted/positive.</td>
</tr>
<tr>
<td>Lander et al., 1999</td>
<td>Naming famous faces from moving/static images. Inversion, negation, thresholding. Movement slowed/staggered.</td>
<td>Natural: from television</td>
<td>2.5 s</td>
<td>Significant advantage for moving over multiple-static images in all conditions. Slowed/staggered movement recognised worse than normal movement (only when normal movement shown 4 times).</td>
</tr>
<tr>
<td>Lander &amp; Bruce, 2000</td>
<td>Naming famous faces from moving/static images. Thresholding. Movement jumbled/speeded/reversed.</td>
<td>Natural: from television</td>
<td>1 – 2.5 s</td>
<td>Moving images in all conditions recognised better than jumbled static images. Normal speed, forward movement recognised better than speeded or backwards movement.</td>
</tr>
<tr>
<td>Lander et al., 2001</td>
<td>Naming famous faces from moving/static images. Pixelating, blurring.</td>
<td>Natural: from television</td>
<td>2.5 s</td>
<td>Moving faces were recognised significantly better than static images, particularly for higher levels of blur.</td>
</tr>
<tr>
<td>Lander &amp; Bruce, 2004</td>
<td>Priming study. Famous faces. Static and moving primes and test images. Slowed movement.</td>
<td>Natural: from television</td>
<td>2.5 s</td>
<td>Moving images prime more effectively than static images to moving and static test images (including identical static and different moving images). Slowing movement results in equal priming to static images.</td>
</tr>
<tr>
<td>Source</td>
<td>Task Description</td>
<td>Conditions</td>
<td>Time (s)</td>
<td>Results</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-----------------------------------------</td>
<td>--------------------------</td>
<td>----------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Lander &amp; Chuang, 2005, Experiment 2</td>
<td>Naming famous faces from moving/static images. Thresholding.</td>
<td>Natural: from television</td>
<td>2.5</td>
<td>Moving famous faces were recognised better than static images. Distinctive motion increased performance.</td>
</tr>
<tr>
<td>Layton &amp; Rochat, 2007</td>
<td>Infant study. Measured habituation recovery to moving/static mother or stranger. Non-degraded, negation.</td>
<td>Natural: speech and expressive</td>
<td>20</td>
<td>No effect of movement when faces were not negated. Movement advantage for 8 month-old infants viewing their mother in negated condition. No discrimination for stranger in any degraded condition.</td>
</tr>
<tr>
<td>Rosenblum et al., 2007</td>
<td>Recognition of friends from moving/static PLDs.</td>
<td>Non-rigid: speech movements</td>
<td>3</td>
<td>For 4/7 of the target identities, movement supported facial recognition. Chance performance with static PLDs.</td>
</tr>
</tbody>
</table>
Research using thresholded, pixilated and blurred famous faces (see Figure 4, Chapter 2) supports the idea that movement advantage is not isolated to negated faces (Lander et al., 1999; Lander & Bruce, 2000; Lander, Bruce, & Hill, 2001). Furthermore, these studies have established that the movement advantage does not arise simply because moving images contain a number of static frames. Lander et al. (1999) and Lander and Bruce (2000) found that the movement advantage for thresholded faces persisted even when multiple static images were used as a control condition.

All of the famous face studies mentioned above used naming tasks to assess recognition. However, Lander and Bruce (2004) demonstrated that moving faces also act as better primes than static faces. Participants were significantly faster at making speeded familiarity judgments for faces that had been primed with a moving image, compared to a static frame. If the movement in the prime and test clip was different, a reaction time advantage for moving primes was still found for moving and static test images, but the movement advantage was reduced (although still significant). The evidence from Lander and Bruce’s (2004) priming experiments suggests that the movement advantage is not linked specifically to the naming task, and it can arise even if participants are not consciously looking for characteristic patterns of movement.

3.3.1.2 Personally familiar faces. Overall, studies have confirmed that movement can provide an important cue to recognition of famous faces. Research using personally familiar faces supports the same conclusion. Like famous face studies, research on personally familiar faces has generally used short (1.5 – 5 s) clips of degraded faces, and most studies have used straightforward recognition tasks (see Table 1 for more details). Bruce and Valentine (1988) and Rosenblum et al. (2007) presented facial PLDs of participants’ friends and acquaintances, and asked participants to name them. While static PLDs of personally familiar faces were identified at chance levels, both studies found that the majority of moving displays were recognised more accurately than their static counterparts. Likewise, Lander and

---

2 Thresholding changes all pixel values to pure black or white, based on a nominal threshold. Pixelation reduces the number of pixels in an image, by replacing each adjacent cluster of pixels in the higher resolution image (minimally a 2 x 2 grid) with a luminance average of those pixels. See Figure 4 (Chapter 2) for an example; or Lander and Bruce (2000) and Lander et al. (2001) for more details.
Chuang (2005, Experiment 1) found that participants were more accurate at naming degraded images of lecturers and teaching staff when they were shown talking and expressing than when they were shown static images.

The movement advantage for personally familiar faces is not isolated to adults. Layton and Rochat (2007) habituated 8-month-old infants to moving or static unfamiliar faces, and tested their level of dishabituation to their mother or to a second unfamiliar face (also moving or static). Overall, the infants showed more dishabituation to images of their mother than unfamiliar faces. Furthermore, when the images were photographic negatives, infants only showed significant dishabituation to moving, not static images of their mother’s face. This demonstrates that motion can help people compensate for degraded static information even at an early stage of development.

Unlike famous face studies, using personally familiar faces as stimuli has allowed researchers to control the face and head movements displayed in the clips. Consequently, several studies have investigated what type of movement is important for recognition, by comparing rigid head movements and non-rigid face movements. Bruce and Valentine (1988) found that participants were equally accurate at identifying their acquaintances from rigid (nodding, shaking and rocking) and non-rigid (smiling, frowning and surprise) movements. However, participants in Lander and Chuang’s (2005) experiment were more accurate at identifying personally familiar faces from non-rigid than rigid movements (in fact, they showed no movement advantage in the rigid condition), a surprising result given that both studies used similar types of movement in their rigid and non-rigid conditions. It is possible that the type of stimulus was an important factor in this result. Bruce and Valentine (1988) used PLDs, which contain very little form-based information (picture-based information, such as eye or nose shape). Participants may have used the rigid head movements to extract structural information, which could have helped with subsequent recognition. On the other hand, Lander and Chuang (2005) used images that were degraded by reducing contrast, altering brightness and adding blur. Although this reduced static recognition rates, these stimuli preserved some form-based information, which could have eliminated any structure-from-motion benefit that the rigid head movements conveyed. The role of stimuli will be discussed further in section 3.4.
3.3.1.3 The origin of the movement advantage in familiar faces. One important question that studies of moving faces have sought to address is the origin of the movement advantage. Studies that show naturalistic movement (i.e., all of the famous face studies mentioned above) demonstrate that movement is useful for identification, but they do not show why. It is possible that participants can use motion to recover basic structural cues that are obscured by degradations such as negation or the transformation into a PLD. The fact that basic rigid motions (such as nodding or shaking the head) can lead to identification supports this idea (Bruce & Valentine, 1988). It is also possible that participants are using individuals’ characteristic motion patterns (either rigid or non-rigid motions) to boost recognition.

Several studies have sought to distinguish between these possibilities by altering the temporal characteristics of the facial movement. If the movement advantage is based purely on characteristic dynamic cues then changing the absolute or relative timing of the motion should eliminate the movement advantage (i.e. moving and static clips should be identified equally well). However, if some of the movement advantage is derived from characteristic dynamic information and some is derived from structure-from-motion, then altering the temporal properties of the motion should reduce, but not eliminate, the movement advantage (i.e. original movement identified better than altered movement, which is identified better than static movement). Finally, if the movement advantage is derived purely from structure-from-motion, the movement advantage for temporally altered clips should be the same size as for original timing clips (i.e. all movement, original or altered, recognised equally well, and better than static clips).

Overall, there is evidence that characteristic motion patterns are important for famous faces in recognition tasks (Lander et al., 1999; Lander & Bruce, 2000), but not priming tasks (Lander & Bruce, 2004). Lander et al. (1999, Experiment 4) presented motion that had been altered by slowing or staggering the clip. Participants were better at naming famous faces when the original timing of the motion was preserved, but only if the original clip was repeated four times. If the original timing clip was only shown once, participants were no better than in the

3 Staggered motion preserves the order of the frames in a video, but alters how long each one is shown. This results in a jerky perception of motion.
slowed or staggered motion conditions – perhaps because the original timing clip was one quarter the length of the other clips, so any characteristic movement advantage was outweighed by the advantages of seeing the face for longer. Subsequent studies confirmed that the timing of the movement is important for famous faces: participants were significantly better at recognising famous faces when the timing and order of frames was preserved than if the order of the frames was jumbled,\(^4\) (Lander & Bruce, 2000, Experiment 1).

The results from Lander et al. (1999, Experiment 4) and Lander and Bruce (2000, Experiment 1) suggest that characteristic motion patterns play an important role in movement-based face recognition. However, these experiments did not include a static control condition, making it impossible to tell whether participants were using structure-from-motion information in the temporally altered conditions. Lander and Bruce (2000, Experiment 2) compared natural movement to reversed and speeded-up clips of famous faces. Clips showing natural movement were recognised significantly more accurately than clips showing reversed or speeded up motion, reinforcing the role of characteristic motion in famous face recognition tasks. However, all the moving conditions were recognised better than the static condition, which suggests that participants may also have been using structure-from-motion processes to extract structure from the moving clips. Lander and Bruce’s findings also suggest that absolute temporal relationships are important to identification. If participants were basing their identity judgements on relative timing cues, speeding up or slowing down the clip should not impair recognition. The fact that speeded and slowed clips were recognised at the same level as staggered or jumbled motion implies that characteristic motion patterns include some representation of absolute timing or speed.

Personally familiar faces also appear to use characteristic motion patterns as a cue to identity. Lander, Chuang, and Wickham (2006) asked participants to identify personally familiar faces from natural and morphed smiles, or a single static frame of a smile (see Table 4). The natural smile, which preserved characteristic motion patterns, was recognised more accurately than either the static or morphed smile. However, when the smiles were presented at double speed, the advantage for the

---

\(^4\) Jumbled motion changes the order of presentation of the frames in a video clip. Unlike staggered motion, the duration of each frame is held constant.
natural smile disappeared. Since there was no possibility of structure-from-motion cues (the stimuli showed no rigid head motion), Lander et al. (2006) interpreted these findings as showing that the movement advantage arises because of characteristic motion patterns, particularly the temporal characteristics of the movement.

Further evidence for the use of characteristic motion patterns comes from a study that looked at the relationship between distinctiveness of movement and the movement advantage. Highly distinctive static faces are easier to recognise (Valentine & Bruce, 1986), and there is some evidence from studies using exaggeration that arm movements and facial expressions are also easier to recognise when they are further from the average value (Hill & Pollick, 2000; Hill et al., 2005; Pollick et al., 2003). Lander and Chuang (2005, Experiment 2) asked participants to rate a number of famous faces based on the distinctiveness of their movements. A second group of participants attempted to name the famous faces from moving clips or still images. The faces that had been rated as having highly distinctive motion were recognised significantly more accurately than the static images. However, faces with non-distinctive motion showed no movement advantage. These results provide further support for the idea that the movement advantage for familiar faces is derived primarily from recognition of characteristic motion.

There is an important caveat to these conclusions. The movement used in all of the famous face studies was derived from television shows, and displayed limited amounts of rigid head motion. The only personally familiar face study that has investigated the origins of the movement advantage (Lander et al., 2006) presented limited non-rigid movements (smiling). It is possible that structure-from-motion processes require more extensive rigid head motion to give rise to a movement advantage. Therefore, it is difficult to determine whether familiar faces derive any benefit from structure-from-motion processes, or whether the movement advantage is based primarily on characteristic motion patterns. This idea will be discussed further in relation to unfamiliar faces.

3.3.2 Experimentally Familiar and Unfamiliar Faces

3.3.2.1 Experimentally familiar faces. Although studies of famous and personally familiar faces have regularly found some benefit for movement, studies using unfamiliar faces have been less consistent. Many studies have investigated whether motion helps people to learn new faces. Experimental familiarity studies
using movement tend to use short familiarisation phases (few exposures with short clips), with non-degraded facial images in both the study and test phases (see Table 2). Some studies have found an advantage for learning or testing face in motion, but the exact details of the movement advantage – for example, the type of motion, or an advantage for movement at learning or test – are highly inconsistent. Pike, Kemp, Towell, and Philips (1997) found that faces learnt in rigid rotational movement (i.e., rotating in horizontal view) were recognised significantly more accurately than faces learnt from static images, even when the number of frames was equated. Lander and Bruce (2003) also found that moving faces were learnt better than static images, but unlike Pike et al. they only found a movement advantage for non-rigid face movements (smiling, talking), and not rigid head movements (nodding and shaking the head).

Shiff, Banka, and de Bordes-Galdi (1986) also investigated the effect of dynamic and static learning and test phases, in the context of an eyewitness experiment. They found no movement advantage for the learning phase, possible because the “robbery” scene distracted participants from attending to the faces. However, dynamic cues in the test phase (consisting of a similar rigid rotational movement to Pike et al., 1997) resulted in better discrimination performance. In a similar eyewitness paradigm, Shepherd, Ellis, and Davies (1982) investigated the effect of seeing a live target. They found that having a live actor in the learning phase resulted in significantly better recognition than videos or still images. However, unlike Shiff et al. (1986), there was no benefit for dynamic presentation (live or video) during the test phase.

Although the majority of studies have investigated the effect of movement on recognition accuracy, Pilz, Thornton, and Bulthoff (2006) investigated the effect of movement on reaction times in priming (Experiment 1) and visual search (Experiment 2). Dynamic primes resulted in significantly faster reaction times than static primes, despite changes in expression and viewpoint between the prime and target images. Pilz et al. (2006) also found a significant movement advantage in a visual search task: reaction times were significantly faster for faces learnt in movement, compared to static images (also see Pilz et al., 2009). Infant studies using familiarisation and novelty preference also show a movement advantage – infants between 3 and 4 months old ‘recognised’ a face (demonstrated a novelty preference for a new face) after 30 s familiarisation with a moving face. However, static faces
Table 2:
An overview of previous research involving experimentally familiar faces and motion

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental/image manipulation</th>
<th>Movement</th>
<th>Length of clip</th>
<th>Major findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shepherd et al., 1982</td>
<td>“Eyewitness” experiment. No</td>
<td>Natural</td>
<td>Learning: 2 mins</td>
<td>Live, but not video learning phase resulted in better identification than static images. No effect of live/video presentation at test.</td>
</tr>
<tr>
<td></td>
<td>image manip.</td>
<td>Test: unlimited</td>
<td>Test: unlimited</td>
<td></td>
</tr>
<tr>
<td></td>
<td>and static learning and test</td>
<td>Test: rigid, rotational</td>
<td>Test: 10 s</td>
<td></td>
</tr>
<tr>
<td></td>
<td>phases. No image manip.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schiff et al., 1986</td>
<td>Moving and static learning and</td>
<td>Rigid and non-rigid scripted</td>
<td>Learning: 10 / 20</td>
<td>No benefit for moving sequence in learning phase. Longer exposure time lead to better recognition overall.</td>
</tr>
<tr>
<td></td>
<td>test phases. No image manip.</td>
<td>Learning: 10 s</td>
<td>s Test: static</td>
<td></td>
</tr>
<tr>
<td>Bruce &amp; Valentine, 1988,</td>
<td>Moving and static learning</td>
<td>Rigid, rotational</td>
<td>Learning: 3 s</td>
<td>Participants consistently recognised faces learned as moving images better than those learned as multiple or single static images.</td>
</tr>
<tr>
<td>Experiment 1</td>
<td>phases. No image manip.</td>
<td>Test: 3 secs</td>
<td>Test: 3 secs</td>
<td></td>
</tr>
<tr>
<td>Pike et al., 1997</td>
<td>Moving and static learning</td>
<td>Rigid and non-rigid scripted</td>
<td>Learning: 10 s</td>
<td>No benefit at study or test for moving images compared to multiple static frames.</td>
</tr>
<tr>
<td></td>
<td>phases. No image manip.</td>
<td>Learning: 10 s</td>
<td>Test: static</td>
<td></td>
</tr>
<tr>
<td>Christie &amp; Bruce, 1998</td>
<td>Moving and static learning and</td>
<td>Rigid and non-rigid scripted</td>
<td>Learning: 30 / 60 s</td>
<td>No advantage for simple familiarisation by video, regardless of length.</td>
</tr>
<tr>
<td></td>
<td>test phases. No image manip.</td>
<td>Learning: 3 s</td>
<td>Test: static</td>
<td></td>
</tr>
<tr>
<td>Bruce et al., 2001,</td>
<td>Familiarisation study. Matching</td>
<td>Rigid and non-rigid scripted</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 2</td>
<td>static low/high quality images</td>
<td>Learning: 30 / 60 s</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>of faces learned in motion.</td>
<td>Test: static</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Description</td>
<td>Type</td>
<td>Learning:</td>
<td>Test:</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>---------------</td>
<td>-----------</td>
<td>---------</td>
</tr>
<tr>
<td>Knappmeyer et al, 2003</td>
<td>Facial morphs animated with individual movements.</td>
<td>Non-rigid</td>
<td>Learning: 8-10 s</td>
<td>Test: 8-10 s</td>
</tr>
<tr>
<td>Lander &amp; Bruce, 2003</td>
<td>Rigid/non-rigid/reversed non-rigid/static/multiple static learning phases. No image manip.</td>
<td>Rigid and non-rigid</td>
<td>Learning: 7 s</td>
<td>Test: static</td>
</tr>
<tr>
<td>Lander &amp; Davies, 2007, Experiment 1</td>
<td>Moving and static learning and test phases; degraded images.</td>
<td>Non-rigid</td>
<td>Learning: 5 s clips, viewed unlimited times</td>
<td>Test: 5 s</td>
</tr>
<tr>
<td>Otsuka et al., 2009</td>
<td>Infant familiarisation study (3-4 months), moving or static faces. No image manip.</td>
<td>Non-rigid</td>
<td>Learning: 30 / 90 s</td>
<td>Test: static</td>
</tr>
<tr>
<td>Pilz et al., 2009</td>
<td>Delayed visual search task with moving and static; high and low demand familiarisation phases.</td>
<td>Non-rigid</td>
<td>Learning: 100 s</td>
<td>Test: static</td>
</tr>
</tbody>
</table>
required 90 s of familiarisation to reach similar levels of recognition (Otsuka et al., 2009).

One of the most intriguing and innovative studies in the area of movement and face recognition examined the interaction between form and motion information in experimentally familiar faces. Knappmeyer, Thornton, & Bulthoff (2003) familiarised their participants with two animated heads, which contained individual structural information and motion patterns. Participants subsequently viewed morph faces, which varied the amount of structural information from each face. However, the faces were animated with motion patterns from either of the original faces. Knappmeyer et al. found that animating a morphed face with experimentally familiar motion significantly biased identity judgements – the point at which participants judged a face to belong to person A 25%, 50% and 75% of the time changed by between 14% and 25% if the face was animated with person B’s motion. This effect was reduced to 3.9-11.4% when the faces had individual skin textures, which shows that form cues do play a strong role in influencing identity decisions. However, even when strong form cues were present, motion still significantly influenced identity decisions.

Not all studies have found a movement advantage for experimentally familiar faces. Bruce and Valentine (1988) and Christie and Bruce (1998) found no recognition benefit for faces learnt in motion, regardless of whether rigid head movements or non-rigid face movements were presented to participants. Similarly, Bruce et al. (2001, Experiment 2) failed to find a movement advantage in a matching task, even when the familiarisation phase was 1 min long (although when participants were encouraged to socialise in Experiment 3, a movement advantage emerged). Possible reasons for these conflicting results will be analysed in section 3.3.2.3.

In sum, the evidence for a movement advantage for experimentally familiar faces is inconclusive. In some cases, but not all, movement can help us to learn new faces. Where it is present, it is unclear whether rigid or non-rigid movement is responsible for the movement advantage, and whether the advantage arises for learning or testing images in motion.

3.3.2.2 Unfamiliar faces. Based on the results reviewed above, it is unclear whether movement plays any significant role in acquiring robust representations of new faces. However, it is possible that motion is still a valid cue to identity when
viewing unfamiliar faces: several studies have found a movement advantage for unfamiliar faces that have not undergone any familiarization phase (see Table 3). Bruce et al. (1999) showed a limited movement advantage (only present when participants could view the videos for an unlimited time) when matching from a video or still image to a static array, and Rosenblum et al. (2002) found a movement advantage when matching unfamiliar faces to PLDs.

Studies of unfamiliar face recognition have also shown that motion alone can support recognition. Hill and Johnston (2001) presented participants with shape-averaged animated heads, which depicted a number of unfamiliar people telling a short joke (i.e., only the movement differed between stimuli). Participants were able to sort the unfamiliar faces into groups according to identity, and complete a match-to-sample task, despite the elimination of any identifying form or structure cues. In a similar vein, an experiment using the same shape-normalised animated heads demonstrated that infants between 4 and 8 months old could discriminate between actors, even when the infants were habituated and tested using different motion sequences from the same actor (Spencer, O’Brien, Johnston, & Hill, 2006). In other words, both adults and infants can discriminate between individuals based solely on characteristic movement patterns.

As in personally and experimentally familiar face recognition experiments, several studies have investigated what type of motion underpins the movement advantage for unfamiliar faces. Hill and Johnston (2001) found that participants performed at above-chance levels when only the rigid motion was shown, and marginally above chance when only the non-rigid information was shown. This does not mean that non-rigid motion is unhelpful for unfamiliar face matching – Rosenblum et al. (2002) found a movement advantage when participants were asked to match a fully illuminated, unfamiliar face to a PLD, despite the fact that the actor’s head was restrained when the clips were being produced. Non-rigid motion is also more resilient to viewpoint changes than rigid motion (Watson, Johnston, Hill, & Troje, 2005).

Moving unfamiliar faces can also elicit a priming effect, similar to that found with famous and experimentally familiar faces (Lander & Bruce, 2004; Pilz et al., 2006). Unfamiliar faces were matched faster in a same/different task when the first “prime” image was moving than when it was static (Thornton & Kourtzi, 2002).
<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental/image manipulation</th>
<th>Movement</th>
<th>Length of clip</th>
<th>Major findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bruce et al., 1999, Experiment 3</td>
<td>“Eyewitness” experiment with 10-choice array. Static/moving/unlimited moving conditions. No image manip.</td>
<td>Rigid and non-rigid scripted</td>
<td>5 s/ unlimited</td>
<td>No benefit for 5 s video over still image, but limited benefit for unlimited video viewing over 5 s video and unlimited still image conditions.</td>
</tr>
<tr>
<td>Rosenblum et al., 2002</td>
<td>Matching articulating PLDs to full-face videos. Normal, static, reduced frames, jumbled and staggered.</td>
<td>Non-rigid: speech movements</td>
<td>3 s</td>
<td>Dynamic and reduced frame presentation matched better than static, jumbled and staggered conditions.</td>
</tr>
<tr>
<td>Thornton &amp; Kourtzi, 2002</td>
<td>Priming task: Comparing dynamic and static primes. Test images upright or inverted.</td>
<td>Non-rigid: scripted expressive</td>
<td>540 ms</td>
<td>RT advantage for moving primes, regardless of test face orientation. No movement advantage when task was expression judgement.</td>
</tr>
<tr>
<td>Kamachi et al., 2003</td>
<td>Cross-modal study. Face-to-voice/voice-to-face matching. Forward/backward motion.</td>
<td>Natural: speech movements</td>
<td>2.8 – 4 s</td>
<td>Participants could match faces to voices (and vice versa), despite different sentence content. Backwards motion/sentences prevented matching.</td>
</tr>
<tr>
<td>Watson et al., 2005</td>
<td>Matching animated heads from rigid or non-rigid motion, across different viewpoints.</td>
<td>Natural. Isolated rigid/non-rigid/combination</td>
<td>6.75 s</td>
<td>Overall, combined &gt; rigid &gt; non-rigid. Matching rigid motion is more viewpoint dependent than matching non-rigid motion.</td>
</tr>
</tbody>
</table>
Inverted motion also shows some viewpoint invariance.

<table>
<thead>
<tr>
<th>Study (year)</th>
<th>Experiment Description</th>
<th>Stimulus</th>
<th>Condition</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spencer et al., 2006</td>
<td>Infant familiarisation study (4-8 mths). Averaged animated head.</td>
<td>Natural.</td>
<td>30 s</td>
<td>Infants familiarised with motion sequences from one actor showed a preference for a new actor. Infants could generalise to different sequences from the same actor.</td>
</tr>
<tr>
<td>Lander et al., 2007</td>
<td>Cross-modal study. Face-to-voice and voice-to face matching. Language, manner of speech, speeded/slowed speech.</td>
<td>Natural: speech movements</td>
<td>2-2.5 s</td>
<td>Participants could match faces to voices and vice versa despite changes to content and language. Artificial changes to speed did not impair performance. Changing manner of speech impaired matching.</td>
</tr>
<tr>
<td>Davis &amp; Valentine, 2009</td>
<td>“Eyewitness” experiment with live actors and photographs.</td>
<td>Natural: walking or rigid rotational</td>
<td>Unlimited</td>
<td>Live actors were recognised equally as well as photographs in all conditions.</td>
</tr>
<tr>
<td>Farivar et al., 2009</td>
<td>Matching a motion-defined face to a static shaded face</td>
<td>Rigid rotational</td>
<td>Unlimited</td>
<td>Significantly better matching for unfamiliar motion-defined faces than static frames.</td>
</tr>
</tbody>
</table>
A reaction-time advantage for moving images remained when matching identity across different expressions, but disappeared when participants were asked to match expression across identity in each trial, suggesting that movement was not conveying a general arousal or orienting advantage, but some benefit specific to identification.

While moving clips can facilitate matching and priming, the movement advantage for unfamiliar faces does not appear to extend to live presentation – no studies to date have found a benefit for matching live, unfamiliar actors to videos or photographs. Davis and Valentine (2009) investigated whether live actors could be matched to high quality video or CCTV footage, and found that participants were no better at matching high quality video footage to live actors than to photographs. More worryingly, even when participants can view the actors and videos simultaneously, they still made identification errors between 15% and 44% of the time (Davis & Valentine, 2009; Megreya & Burton, 2008). Similar error rates have been obtained when participants have been asked to match live actors or photographs to an array of photographs, regardless of whether the images are presented sequentially or simultaneously (Megreya & Burton, 2008).

Overall, these studies suggest that unfamiliar faces can be matched based purely on their movement. However, error rates for simple identification tasks are quite high for unfamiliar faces, even when the stimuli are high-quality images (Davis & Valentine, 2009; Megreya & Burton, 2008). Although movement can facilitate unfamiliar face matching in some circumstances, overall matching performance still remains remarkably poor in comparison to famous and personally familiar faces. One question that remains, and will be addressed throughout this thesis, is whether the overall recognition benefit for famous and personally familiar faces extends to movement-based face recognition (i.e., do familiar faces also show a higher movement advantage than unfamiliar faces), or whether people can use movement-based cues to improve their performance with unfamiliar faces (i.e., does movement compensate for poor recognition of unfamiliar faces). The theoretical and empirical evidence for each of these possibilities will be discussed further in section 3.4.

3.3.2.3 The origin of the movement advantage in experimentally familiar and unfamiliar faces. It is possible that unfamiliar and experimentally familiar faces show unreliable motion effects because there is only one way that they can benefit from motion – that is, via structure-from-motion mechanisms. Unfamiliar and
experimentally familiar faces are less likely to use characteristic motion patterns as a cue to identity due to the amount of time it takes to build these pathways (O’Toole, Roark, & Abdi, 2002; Roark et al., 2003). However, evidence that participants can match unfamiliar, shape-normalised faces (Hill & Johnston, 2001) and that motion patterns bias identity judgements of facial morphs (Knappmeyer et al., 2003) suggests that it is still possible to extract and compare characteristic motion patterns with minimal prior exposure or training with a face.

3.3.2.3.1 Experimentally familiar faces. The role of structure-from-motion and characteristic motion patterns in experimentally familiar faces is unclear because most of the studies in this area have not investigated whether people acquire characteristic motion patterns during experimental familiarization with a face. Many studies on experimentally familiar faces have used static test images, which do not test whether characteristic motion patterns are being encoded, but focus almost exclusively on whether motion builds a better representation of a static face (e.g. Bruce & Valentine, 1988, Experiment 1; Bruce et al. 2001, Experiments 2 and 3; Lander & Bruce, 2003; Pike et al., 1997; Pilz et al., 2006). Pike et al.’s results suggest that rigid head motion can help subsequent static recognition, which supports the hypothesis that unfamiliar faces can use structure-from-motion cues to enhance recognition. However, other studies using static test images have failed to find a beneficial effect of rigid motion (Bruce & Valentine, 1988; Christie & Bruce, 1998, Experiments 1A and 2A), or have only found a movement advantage when comparing performance with single static images, not multiple static images (Lander & Bruce, 2003).

It is possible that the conflicting results between Pike et al.’s (1997) and other studies arise because of the amount of rigid head movement shown. Pike et al. showed a full 360-degree rotation around the head, which offered participants every possible viewing angle of the face and head, whereas Christie and Bruce (1998, Experiment 2), Bruce and Valentine (1988, Experiment 1) and Lander and Bruce (2003, Experiments 1 and 2) used head nodding and shaking movements, providing relatively few viewing angles of the head. Alternatively, Roark et al. (2003) posited that the conflicting results may have arisen due to differing social content in the head movements. In Pike et al’s study, the face remained neutral throughout, so the participants could not have been distracted by any social cues present in the videos. On the other hand, nodding and shaking contain many social cues, such as agreement.
and disagreement, and positive and negative valence. The presence of social cues in the rigid motion may have distracted participants or encouraged them to pay attention to non-identity based cues, thus impairing their performance on the task.

The effect of social cues may not always be detrimental to identification. As mentioned above, Pilz et al. (2009) only found a movement advantage for faces learned in motion when participants had answered an in-depth personality questionnaire about the faces. Furthermore, Lander and Bruce (2003, Experiments 1, 3 and 4) found that faces learnt in non-rigid motion (even reversed non-rigid motion) were subsequently recognised better than faces learnt from still images. Since the non-rigid motion could not have conveyed structure-from-motion information, and participants could not match characteristic motion patterns (all test images were static), the authors concluded that the moving images were most likely remembered better because they were more interesting to participants.

When motion is present at both learning and test, the role of structure-from-motion and characteristic motion patterns for experimentally familiar faces remains unclear: Christie and Bruce (1998) showed no benefit of rigid or non-rigid motion during learning or test phases, while Lander and Davies (2007, Experiment 1) found a motion advantage only if movement (in this case, non-rigid motion) was present during both learning and test phases. Knappmeyer et al. (2003) also found a strong effect of non-rigid motion on identity decisions, using facial morphs. However, evidence from eyewitness studies is inconclusive: Shiff et al. (1986) found a benefit for motion at test, but not learning, while Shepherd et al. (1982) found a benefit for live actors during the learning, but not test phases.

Lander and Davies’ (2007) and Knappmeyer et al.’s (2003) results clearly imply that their participants were learning characteristic motion patterns. However, it is unclear why the same effect did not appear in Christie and Bruce’s (1998) results. It is possible that the differences between the findings arose due to different stimuli in the test phase – Lander and Davies presented degraded test images, and Knappmeyer et al. presented ambiguous identity information by morphing the test faces, whereas Christie and Bruce used non-degraded videos. Degraded or ambiguous images may have encouraged participants to focus on movement by depriving them of clear static cues. Participants in Lander and Davies’ study also had unlimited viewing time in the learning phase, which may have allowed them to learn characteristic movement patterns. Similarly, Knappmeyer et al. presented 30 second
videos, which presented the facial motion multiple times. In contrast, Christie and Bruce only allowed participants 3 seconds to view each face, which may not have been enough time to learn characteristic motion patterns.

Unlike Lander and Davies (2007) and Knappmeyer et al. (2003), the eyewitness experiment results do not indicate that characteristic motion patterns are useful in experimentally familiar face recognition. Schiff et al.’s findings may reflect a contribution of structure-from-motion. Schiff et al. used a dynamic test phase that resembled the learning phase from Pike et al (1997), with 180 degrees of rigid rotational movement. It is possible that this mode of presentation facilitates recognition, as it supplies multiple viewpoints of the face and head, and allows the extraction of structure-from-motion information. The eyewitness experiments also suggest that social signals have a large impact on whether or not participants benefit from the presence of movement during the learning phase. Live actors may have attracted more attention than videos or photographs (Shepherd et al., 1982), but presenting an entire “crime” could have pulled attention away from identity (Schiff et al., 1986).

Overall, the inconsistency of the results and the incredibly varied stimuli and methodology employed by experimentally familiar face recognition studies make it difficult to conclude whether the presence of a movement advantage results from characteristic motion patterns, structure-from-motion, social signals or a mix of all three mechanisms.

3.3.2.3.2 Unfamiliar faces. The mechanisms underlying a movement advantage are unclear for experimentally familiar faces. However, there is clear evidence that characteristic motion patterns can be used to match unfamiliar faces. As mentioned above, Hill and Johnston (2001) showed that shape-averaged heads could be sorted on the basis of rigid or non-rigid motion alone. On the surface, these results suggest that participants were able to extract characteristic motion patterns from the stimuli, and were not relying on speech-reading or isolated idiosyncratic movements to complete the tasks. However, participants could also discriminate between animations equally well when they were played forwards and backwards, and performed better than chance when discriminating inverted animations. It is possible that backwards or inverted motion preserves characteristic motion patterns, or produces equally distinctive new motion patterns; it is equally possible that the stimuli contained some characteristic static cues that participants were using
to identify the faces. As Hill and Johnston did not test static sorting or discrimination performance, it is difficult to say whether some proportion of their results was due to idiosyncratic head poses or characteristic facial expressions.

Like Hill and Johnston (2001), Rosenblum et al. (2002) also found that moving, unfamiliar faces could be matched at above-chance levels, even though they did not depict any rigid head motion. Rosenblum et al. (2002) also provided evidence that dynamic information was responsible for the matching performance: staggering or jumbling the facial speech information in the PLDs reduced matching performance to chance levels for the majority of actors. Based on these results, they suggested that visual speech information carries sufficient characteristic information to support identification. In fact, auditory speech alone provides participants with sufficient information to perform a cross-modal identity-matching task (Kamachi, Hill, Lander & Vatikiotis-Bateson, 2003). Several studies have explored cross-modal effects by asking people to match an auditory speech token to a moving face (or vice versa) in a delayed match-to-sample task. In general, people are quite accurate at cross-modal identity matching. Studies have found that participants can match single words to PLDs (Lachs and Pisoni, 2004a) and videos (Lachs & Pisoni, 2004b; Lachs & Pisoni, 2004c); and can match short sentences to PLDs (Rosenblum, Smith, Nicols, Hale & Lee, 2006) and videos (Kamachi et al., 2003; Lander, Hill, Kamachi & Vatikiotis-Bateson, 2007), even when the sentence is different in the audio and visual modalities. Cross-modal matching is possible even when the speech is in a foreign language (Kamachi et al., 2003) or when the fundamental frequency is removed (Lachs & Pisoni, 2004b; Kamachi et al., 2003), suggesting that the ability is not based on language-specific speech or prosodic cues. Interestingly, participants perform well above chance levels when the speech is artificially slowed or speeded (Lachs & Pisoni, 2004c; Lander et al., 2007). This suggests that, unlike famous face recognition (Lander & Bruce, 2000; Lander et al., 1999), participants in the cross-modal tasks are not using basic physical characteristics or absolute timing information as a cue to identity. However, there is strong evidence that people use the relative timing of speech as a cue to identity. Identification performance is significantly impaired when the auditory information is reversed (Lachs & Pisoni, 2004b) or transformed in a non-linear fashion (Lachs & Pisoni, 2004c); or when the visual information is presented as a static image (Kamachi et al., 2003; Rosenblum et al., 2006), jumbled or staggered (Rosenblum et al., 2006). Cross-modal matching
performance is also impaired when the manner of speech is changed between the auditory and visual cues – for example, if participants see casual speech but hear a question (Lander et al., 2007).

Overall, these studies demonstrate that the natural timing information carried in speech, whether auditory or visual, is highly useful in identification. Although these studies do not test whether people use the same timing cues in purely visual comparisons, they provide evidence that characteristic movement patterns may consist (at least partly) of characteristic timing patterns, which are expressed concurrently in the auditory and visual domains during speech. More importantly, it is possible to match this characteristic information even when we are unfamiliar with the face and voice.

As well as characteristic movement or timing information, there is evidence that unfamiliar face matching can benefit from structure-from-motion. As mentioned in section 3.2.1.2, Farivar et al. (2009) generated motion-defined faces and tested whether participants could match the rigidly rotating head to one of eight static shaded faces. Participants performed at chance level (12.5%) when a single static frame was extracted from the dynamic stimuli. Performance increased to roughly 25% accuracy (significantly above chance) when participants saw incongruent texture and motion information, suggesting that the texture gradient carried some residual cues to identity. However, when structure-from-motion cues were included (i.e., the texture and motion were congruent), participants were able to match the unfamiliar faces with almost 50% accuracy – significantly higher than either the static or incongruent conditions – confirming that structure-from-motion can support unfamiliar face matching.

Unfortunately, while there is evidence that structure-from-motion and characteristic movement information are available for unfamiliar faces, none of the studies reviewed above can establish the relative contribution of the two in unfamiliar face matching. Like familiar faces, it is unclear whether one mechanism is primarily responsible for the movement advantage in face recognition. Furthermore, it is unclear whether the movement advantage itself differs between familiar and unfamiliar faces. The following section examines this question more closely.
3.4 Differences Due to Familiarity or Differences Due to Methodology?

3.4.1. Familiarity and Movement in Face Recognition

Research on static faces has established that familiarity with a person greatly affects our ability to recognise their face, but to date there is very little empirical research directly comparing the effect of familiarity on moving faces. This chapter reviews the theoretical and empirical evidence directly comparing the role of movement in familiar and unfamiliar faces, and concludes by examining whether some of the differences in results across the literature can be explained by methodological factors.

3.4.1.1 Theoretical differences between familiar and unfamiliar faces.

Theoretically, familiarity is an important factor in movement-based face recognition. The supplemental information hypothesis suggests that we should be better at recognising familiar faces than unfamiliar faces from their motion (O’Toole et al., 2002; Roark et al., 2003). If a person is familiar with a face, they may have already encoded characteristic motion patterns that can provide an alternate pathway to recognition when static information is degraded or ambiguous. Likewise, familiarity with a face should lead to a more robust structural representation of the face and head (Burton et al., 2011), which would be beneficial when attempting to match a stored facial representation with information derived from structure-from-motion. Therefore, when we view a moving face, familiarity should bolster recognition via both of the proposed movement-based mechanisms: characteristic motion patterns, processed in the pSTS; and structure-from-motion information, processed in MT and the FFA (Roark et al., 2003).

Compared to familiar faces, unfamiliar faces should receive a relatively small benefit from motion. Unfamiliar faces can benefit from structure-from-motion (Farivar et al., 2009), and it is possible that structure-from-motion cues help people learn a face – Pike et al. (1997) suggested that the benefit of motion for experimentally familiar faces arose from large, rigid motions that helped people derive a viewpoint-independent representation of the face and head. It is also possible that minimal exposure to a moving face allows people to learn and/or compare dynamic movements (Lander & Davies, 2007; Knappmeyer et al., 2003;
Roark, O’Toole, Abdi & Barrett, 2006), but the inconsistent results in the experimental familiarity literature imply that this is not always the case (Christie & Bruce, 1998). Nonetheless, rapid acquisition of characteristic movement patterns may explain why some studies have found a movement advantage for unfamiliar faces in the absence of large rigid movements (e.g. Hill & Johnston, 2001; Lander & Davies, 2007, Experiment 1; Pilz et al., 2006; Thornton & Kourtzi, 2002). However, it seems likely that increased exposure, leading to more robust moving and static representations of familiar faces, would result in a comparative advantage for recognising moving familiar faces over moving unfamiliar faces.

Currently, it is unclear whether these assumptions about familiarity are correct, and if so, how large the familiarity advantage would be. It is also unclear what effect familiarity should have on the social signals hypothesis. Bower & Karlin (1974) found that static faces judged on “deep” personality characteristics were subsequently recognised better than faces judged on gender, which supports the contention that the social cues carried in movement may help people acquire new face representations – in other words, social cues may help unfamiliar and experimentally familiar face recognition (Pilz et al., 2009). There is also evidence that the processing of social and identity cues becomes less independent as familiarity with a face increases. Research using the Garner speeded classification task found that it is harder to ignore identity information when processing static emotional expressions for familiar faces than unfamiliar faces (Ganel & Goshen-Gottstein, 2004). This could indicate that familiar face identification also benefits from the presence of social cues, as they are processed together. However, it is not clear whether the same effect would apply for moving faces, or whether other social signals, such as speech or conversational expressions, would share the same relationship with familiarity as static emotional faces.

3.4.1.2 Behavioural differences between familiar and unfamiliar faces.
Despite the theoretical differences between familiar and unfamiliar faces, very few studies have directly compared the role of movement in familiar and unfamiliar faces (see Table 4). To the author’s knowledge, no studies have compared moving famous and unfamiliar faces and only four studies have compared recognition performance for personally familiar and unfamiliar faces. Burton, Wilson, et al. (1999) showed participants low quality CCTV footage taken from a university security system, and
### Table 4:
An overview of previous research analysing the effect of familiarity and motion

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental/image manipulation</th>
<th>Familiarity</th>
<th>Movement</th>
<th>Length of clip</th>
<th>Major findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burton, Wilson, et al., 1999</td>
<td>Old/new recognition test, from CCTV to photographs. Body, face or gait obscured.</td>
<td>Personally familiar and unfamiliar</td>
<td>Natural: walking</td>
<td>~ 6 s (E1) 3 s (E2)</td>
<td>Familiar faces matched better than unfamiliar faces. For familiar people, unobscured video &gt; gait-obscured, body-obscured &gt; face obscured. Face is the origin of the familiarity advantage.</td>
</tr>
<tr>
<td>Bruce et al., 2001, Experiment 1</td>
<td>Matching from CCTV (moving/static/multi-static) to photographs.</td>
<td>Personally familiar and unfamiliar</td>
<td>Natural: walking</td>
<td>3 s</td>
<td>Higher hits and correct rejections for familiar than unfamiliar faces, regardless of movement. No movement advantage for familiar or unfamiliar faces.</td>
</tr>
<tr>
<td>Bonner et al., 2003</td>
<td>Learning study. Moving/static learning phases, tested 3 times on internal/external face matching.</td>
<td>Experimentally familiar</td>
<td>Rigid and non-rigid scripted; speech</td>
<td>90 s per session, 3 sessions</td>
<td>Performance for internal, but not external features improved with increased exposure. Slight movement advantage at final test, advantage for matching internal features.</td>
</tr>
<tr>
<td>Lander &amp; Chuang, 2005, Experiment 1</td>
<td>Naming moving/static images. Blurring.</td>
<td>Personally familiar (median familiarity split)</td>
<td>Rigid and non-rigid scripted</td>
<td>1.5 s</td>
<td>Non-rigid movement (expressing/talking) better recognised than rigid movement or static views. “More familiar” faces named better than “less familiar”.</td>
</tr>
<tr>
<td>Lander et al., 2006</td>
<td>Naming degraded images. Natural, morphed and static smiles, normal or fast speed.</td>
<td>Personally familiar (median familiarity split)</td>
<td>Non-rigid: smile</td>
<td>1.1-1.7 s</td>
<td>Natural smiles recognised better than morphed smiles. Speeding natural smiles impaired recognition. “More familiar” faces better than “less familiar”.</td>
</tr>
<tr>
<td>Study</td>
<td>Paradigm</td>
<td>Familiarity</td>
<td>Duration</td>
<td>Moving Advantage</td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>--------------------------------------</td>
<td>----------</td>
<td>------------------</td>
<td></td>
</tr>
<tr>
<td>Roark et al., 2006</td>
<td>Learning study with differing exposure; face-to-gait or gait-to-face. Dynamic and static face images. Front-to-profile. Static test.</td>
<td>Experimentally familiar</td>
<td>9 s (gait); 5 s (face), viewed 1 – 4 times.</td>
<td>Face-to-gait better than gait-to face. Movement advantage increased with extra exposure (face-to-gait). No movement advantage when learning frontal view and testing profile.</td>
<td></td>
</tr>
<tr>
<td>Lander &amp; Davies, 2007, Experiment 2</td>
<td>Learning study with differing exposure; moving and static test phases; degraded images.</td>
<td>Experimentally familiar</td>
<td>1 – 4 episodes, 4.5 - 10.5 mins per episode.</td>
<td>Movement advantage for test images. Extra viewing time increases accuracy, but the beneficial effect of motion does not depend on amount of exposure.</td>
<td></td>
</tr>
</tbody>
</table>
asked them to identify the same people from high quality still images in an old-new recognition task. One third of their participants were personally familiar with the people in the videos, while the other two thirds were not (half of these were students, half were experienced police officers). Familiarity significantly improved performance, but it is unclear whether movement contributed to the result, as there was no static control condition. In a subsequent experiment, participants who were familiar with the targets viewed similar video sequences in four conditions: head obscured, body obscured, multiple static images, and unedited clips. Burton, Wilson, et al. found that the multiple static images were identified less accurately than the unedited video, which suggests that movement may have played a role in the original finding (although “movement” in this case incorporates the whole body, not just the face). However, the multiple static images were recognised equally as well as the body-obscured condition, and far better than the face-obscured condition. Burton, Wilson, et al. concluded that participants were recognising their professors from their faces, and were unable to make use of movement (gait) or body-based cues. In a similar study, Bruce et al. (2001, Experiment 1) asked participants to match low quality CCTV footage (moving or still) with high-quality colour photographs, using participants who were personally familiar or unfamiliar with the targets. Bruce et al. (2001) replicated the earlier familiarity findings – participants who knew the targets showed better overall discrimination and higher confidence than participants who were unfamiliar with the images. However, there was no movement advantage – personally familiar and unfamiliar faces were matched equally well from still shots and moving footage. Once again, these findings support the theory that participants were identifying or matching the images from the face, rather than attempting to use movement-based cues.

Like Burton, Wilson, et al. (1999), Lander and Chuang (2005, Experiment 1) used a naming task to examine the effect of personal familiarity and movement. After the main experiment, participants were asked to rate faces for familiarity, and the faces were split into “more” and “less’ familiar groups for analysis. Unsurprisingly, participants were better at naming faces they rated as more familiar to them. Lander and Chuang’s participants were also better at naming faces that were shown speaking or smiling than still images. Interestingly, the movement advantage was a similar size regardless of whether the faces was considered “more familiar” or “less familiar” to the participant. Lander et al. (2006) carried out a similar
experiment using morphed and natural smiles. Once again, a median split was used to create “more” and “less” familiar groups for analysis, and participants were found to be more accurate when naming faces that had been rated as more familiar. Participants were also more accurate at naming images that showed natural smiles than morphed smiles or static images. Like Lander and Chuang, though, there was no difference in the movement advantage for “more” and “less” familiar faces. However, in both studies the faces that participants rated as “less” familiar could still be identified by participants – they were not unfamiliar, and they may have been associated with characteristic movement patterns.

All four of the studies on personally familiar faces confirm that familiar faces – whether moving or static – are recognised and matched more accurately than unfamiliar or less familiar faces. There is no evidence that the effect of movement changes with increasing familiarity, but this conclusion is tempered by the methodological limitations of the studies – no static control condition (Burton et al., 1999); no movement advantage (Bruce et al., 2001); and “less familiar”, not unfamiliar faces (Lander & Chuang, 2005; Lander et al., 2006). Overall, then, the role of personal familiarity in movement-based face recognition remains unclear.

A separate line of research has attempted to examine the relationship between familiarity and the movement advantage using experimentally familiar faces. Roark et al. (2006) used an old-new recognition task to compare face-to-gait and gait-to-face matching. To test the role of familiarity in movement-based matching, participants viewed each face or gait clip one, two, or four times. Roark et al. (2006) found a significant movement advantage: participants performed more accurately overall when they saw moving faces than static faces. They also found a familiarity advantage: more viewings improved overall performance. Most noteworthy, however, they found that in the face-to-gait condition, the movement advantage increased with familiarity: moving faces led to higher accuracy than static faces in the four viewings condition (i.e. highest familiarity), but not in the one or two viewings conditions. Bonner et al. (2003) used a more extensive familiarisation paradigm (90 s per day over three days), and tested matching performance for familiarised and unfamiliar faces over three sessions. They found that faces learned in motion were matched significantly better at the third testing session (after all familiarisation) than the first testing session (prior to familiarisation), whereas faces learned from static images and unfamiliar faces did not improve across the three
testing sessions. This strongly resembles Roark et al.’s (2006) results, where a movement advantage only appeared after the face was viewed several times. Bonner et al. also tested the effect of movement on internal and external face parts, and found that internal features were matched better for faces learned in motion than for unfamiliar faces. This suggests that familiarisation helped participants acquire better representations of the individual facial features, rather than simply become more familiar with the images as a whole, the shape of the head, or a specific hairstyle.

In a similar study, Lander and Davies (2007, Experiment 2) also tested the effect of motion for different levels of experimental familiarity. They asked participants to name degraded faces of characters after watching one, two, three or four episodes of an unfamiliar television show. Like Roark et al. (2006) and Bonner et al. (2003), they found that moving faces were recognised more accurately than static faces (in this case, naming performance was measured), and accuracy increased with additional viewing time. However, Lander and Davies found no interaction between the movement advantage and viewing time – increased familiarity with the faces did not increase the movement advantage. This could be because participants in Lander and Davies’ study had more time to learn the characteristic motion patterns than those in Roark et al.’s (2006) or Bonner et al.’s (2003) study: Lander and Davies’ one episode condition presented each face for an average of 4.5-10 mins; whereas Roark et al.’s (2006) study showed the moving faces for a maximum of 20 secs in the 4 viewings condition, and Bonner et al. presented each face for 90 s per day. It is possible that the effect of familiarity had plateaued in Lander and Davies’ study. Alternatively, it is possible that the characteristic motion patterns in Lander and Davies’ study were acquired quickly because they were processed in a more social manner (as in Bruce et al., 2001, Experiment 3). Watching a television show (e.g., Lander & Davies, 2007) generally involves following the plot and interpreting emotions and relationships between characters – a much more complex social process than watching a person walk across the screen or speak to someone off camera (Roark et al., 2006), or making name and personality assessments (Bonner et al., 2003).

It is likely that a combination of factors, including differences in duration and social engagement with the stimuli, contributed to the differing results from Roark et al. (2006), Bonner et al. (2003) and Lander and Davies (2007). Regardless, the results of all three studies suggest that experimental familiarity with a face is acquired rapidly,
and movement can help in the process of acquisition. However, it is still unclear how experimental familiarity compares with real-world experience with a face (e.g. personally familiar or famous faces), or whether “deeper encoding” of a face (Bower & Karlin, 1974) is necessary to maximize the movement advantage.

Taken together, these seven studies suggest that familiarity may play an important role in movement-based face recognition, but it is still unclear what that role is. Most studies agree that familiarity improves general matching performance, and it is possible to find a movement advantage for both familiar and relatively unfamiliar faces. However, these studies leave several questions unanswered. Firstly, it is unknown whether the findings from experimentally familiar faces will generalize to personally familiar or famous faces, which generally have much higher levels of familiarity. Secondly, it is unclear whether the movement advantage (or lack thereof) in some studies was a result of the task or stimuli chosen. The following section addresses some of the methodological differences between the studies of in this area, and analyses how they may have impacted on our understanding of movement-based recognition for familiar and unfamiliar faces.

3.4.2 A Methodological Perspective

Many studies have found a movement advantage for face recognition, particularly for famous faces (e.g., Knight & Johnston, 1997; Lander & Bruce, 2000; Lander et al., 2001), but several studies have failed to find a movement advantage for experimentally familiar or unfamiliar faces (e.g., Bruce et al., 2001; Bruce & Valentine, 1988; Christie & Bruce, 1998; Davis & Valentine, 2009). It is possible that the contribution of movement differs in familiar and unfamiliar face recognition, but it is also possible that the divergent findings from previous studies are a result of the widely varied methodologies each study employed. There are many variations in methodology, which can be broadly divided into procedural and stimulus-based differences. This review will focus on three main procedural differences: task, dynamic and static test images and task difficulty. It will also address three stimulus-based differences: image manipulation, type of face motion and stimulus duration.

3.4.2.1 Procedural differences.

3.4.2.1.1 Task. The choice of task may have a significant impact on the experimental findings in this field, particularly for unfamiliar and experimentally
familiar faces. The majority of famous face studies have used naming tasks (e.g. Knight and Johnston, 1997; Lander et al., 1999 etc.), but this is impractical for unfamiliar faces, and requires extensive training for experimentally familiar faces (e.g. Knappmeyer et al., 2003). Consequently, many studies of experimental familiarization have used old/new recognition or “eyewitness” tasks, with various changes between familiarization and test phases – for example, viewpoint changes in Christie and Bruce, (1998), or the presentation of static or multiple static images in Pike et al. (1997). However, an old/new recognition test or delayed eyewitness task requires participants to remember information over an extended period of time, which may erode any short-term advantages conferred by movement. Studies using methods that eliminate or reduce this memory load have consistently found a significant movement advantage. For example, Pilz et al. (2006) used priming and visual search tasks to investigate the role of movement in experimentally familiar faces, and found a significant reaction time advantage for faces primed or learnt in motion. Likewise, studies on unfamiliar faces have used matching-based tasks such as sorting, odd-one-out or sequential same/different decisions (Hill & Johnston, 2001; Thornton & Kourtzi, 2002), and found surprisingly good movement-based performance. Thornton and Kourtzi (2002) suggested that the movement advantage they observed could have been a result of using a task to tap working memory, rather than (more difficult to establish) long-term memory representations.

It is possible, therefore, that the inconsistent results from experimentally familiar face studies, and the differences between experimentally familiar and familiar face studies, can be explained by differences in the task. Unfamiliar and some experimentally familiar faces might only show a movement advantage for short-term memory tasks, because they do not have robust long-term memories for characteristic movement patterns or the information carried by structure-from-motion. On the other hand, faces that have undergone more extensive familiarization (e.g. famous or personally familiar faces, or experimentally familiar faces with unusually long familiarization phases such as Knappmeyer et al., 2003) have a more robust, long-term memory for structure and characteristic movement patterns, which results in a movement advantage in recognition or old/new memory tasks.

Consequently, it is important to use a standard task, which provides the best chances of obtaining a reliable movement advantage for both familiar and unfamiliar faces (see Chapter 4).
3.4.2.1.2 Static and dynamic test images. Another methodological consideration that may impact studies in this field is the use of static or dynamic images during the test phase of experiments on unfamiliar or experimentally familiar faces. As discussed in section 3.3.2.3, many studies using experimentally familiar faces have presented static images at test (see Table 2), which only examines the possibility that movement can build a better 3D head and face representation via structure-from-motion. To fully test the role of motion, including characteristic motion patterns and structure-from-motion, it is necessary to present both static and moving test images.

3.4.2.1.3 Task difficulty. Finally, the difficulty of the old/new recognition and matching tasks has varied between studies, and this has not always been accounted for in results. For example, Pike et al. (1997) presented static test images that had been extracted from the same viewpoint/movement range that their dynamic images showed, while Christie and Bruce (1998) and Pilz et al. (2006) presented at least half the test images at a different viewpoint or expression than the learning phase. Christie and Bruce (1998) also used identical movement sequences at study and test in half of their trials, and Knappmeyer et al. (2003) trained and tested their participants using identical motion clips (although the static information in the heads was varied). On the other hand, Hill and Johnston (2001) and Lander and Davies (2007, Experiment 1) used different movement sequences within their experiments. It is unclear whether the differences in results across these studies may arise due to the overlap between the stimuli that were used, the presentation of those stimuli as static or dynamic in different phases of the experiment, or simply the task participants were asked to complete.

3.4.2.2 Stimulus differences.

The stimuli used to investigate how motion impacts on face recognition have varied widely across previous studies. Differences in the image degradation, type of motion and duration of the moving clips make it difficult to compare results for familiar and unfamiliar faces, and even harder to determine whether the effect of movement is consistent across different levels of familiarity. Stimulus variation may also have contributed to some of the inconsistent findings in the literature.

3.4.2.2.1 Image degradation. One factor that may play a significant role in recognition is the type of image degradation used in different studies. Knight and
Johnston (1997) compared recognition with negated and unaltered images of famous faces, and suggested that the benefits of motion are only apparent when the static form information is less accessible than in a normal image. Subsequently, studies on famous faces have used many image degradations, including blurring, pixilation and thresholding (Lander & Bruce, 2000; Lander et al., 2001, Lander & Chuang, 2005; see Table 1). However, it is important to note that although these image degradations reduce static recognition away from ceiling levels, they do not eliminate it altogether. For example, participants in Knight and Johnston’s (1997) study could name an average of 42% of the negated static faces. Furthermore, each image manipulation could have impaired static face recognition in a different manner. For example, static recognition of negated faces may have been impaired because participants were unable to access shape-from-shading information (Kemp et al., 1996), whereas blurring faces may have made it harder to extract information about individual facial features (Costen et al., 1996). If these image degradations have different effects on static recognition, they may also have different effects on movement-based recognition. Participants viewing a negated face may focus on the movement of individual features, whereas participants watching a blurred or pixelated face may focus primarily on structural cues, since the features are more difficult to see.

Studies using degraded images of famous faces have established that motion can improve otherwise poor face recognition, but they do not suggest that facial movement on its own can provide a separate, stand-alone cue to identity. On the other hand, studies of personally familiar faces have often used point-light-displays, and unlike the degradations used in the studies of famous faces, the use of PLDs has been shown to reduce static recognition to chance levels (Bruce & Valentine, 1988; Rosenblum et al., 2007). The fact that people can still identify their friends from these highly degraded displays confirms that, for personally familiar faces at least, movement information alone can support identification. At this stage, however, no famous face studies have used point-light displays. Therefore, it is difficult to compare the results across different levels of familiarity, or to assess whether movement alone can support recognition in famous faces.

In comparison to famous and personally familiar face studies, very few studies on experimentally familiar or unfamiliar faces have degraded the images at all (see Tables 2 and 3). This is another factor that may explain why some studies have found a movement advantage for experimentally familiar faces, whereas others have not.
Lander and Davies (2007) used degraded images in the test phase of their experimental familiarization study, and found a motion advantage when participants were trained and tested on moving faces. Hill and Johnston (2001) also used degraded face images (in this case, shape-averaged animated heads) to test the motion recognition of unfamiliar faces, and found participants were able to sort and discriminate motion from shape-averaged heads which contained no individual form information. Based on these results, it is possible that experimental familiarization and unfamiliar face studies using non-degraded stimuli encountered a ceiling effect for static recognition, which eliminated any chance of finding a movement advantage. This explanation may also account for the reaction time advantage in priming and visual search tasks (Pilz et al., 2006) – when using non-degraded images, the effect of movement is smaller, and may not be apparent in simple measures of recognition like accuracy. Consequently, it may be important to test unfamiliar faces that have been degraded in the same way as familiar faces so the contribution of form and motion cues can be properly assessed. Furthermore, when comparing the effect of movement on familiar and unfamiliar faces, it may be necessary to design tasks or include dependent measures that go beyond basic measures of accuracy.

### 3.4.2.2 Type of motion

Another consistent difference in the literature is the type of motion participants are shown. As can be seen in Table 1, the stimuli for famous face studies have generally been drawn from television shows. The famous face stimuli show faces from various viewpoints and distances, with a wide range of motion including talking, expressing and rigid head movements. In general, the amount of movement and distinctiveness of movement have not been controlled, despite findings that distinctive motions result in a movement advantage, whereas more typical motions may not (Lander & Chuang, 2005). On the other hand, experimentally familiar face studies and unfamiliar face studies have often used specific scripted motions – for example, looking up and down; smiling; or making speech sounds, (Bruce & Valentine, 1988; Christie & Bruce, 1998; Knappmeyer et al, 2003; see Tables 2 and 3 for more studies). It is possible that natural speech and expressive movements, as displayed in famous face studies, are more idiosyncratic than scripted motions, and therefore more conducive to finding a movement advantage than scripted movements. This is supported by the fact that Hill and Johnston (2001) found above-chance performance for unfamiliar faces that were
shown telling jokes – a speech task that encourages characteristic movements. Alternatively, the interplay of movement and pauses in natural speech could also contribute to face recognition. Roark et al. (2003) proposed that movement during natural speech acts to capture the observer’s attention, which then ensures that the observer is attending to the face during the pauses. This pattern of movement and pauses may act to provide more static views of the face, which in turn enhances face learning.

In addition to the type of motion, familiar, experimentally familiar, and unfamiliar face studies have differed in other ways – for example, experimentally familiar and unfamiliar face studies have generally controlled extraneous factors such as viewpoint, camera distance and the type of emotion or speech being displayed. Therefore, it is possible that participants viewing famous faces may have used cues to identity that were not related to face motion, such as characteristic head positions, body sway or arm movements.

3.4.2.2.3 Movement duration. Another stimulus factor that has varied greatly between studies is the duration of movement. To this point, the impact of duration has not been systematically tested or compared across familiar and unfamiliar faces. Many famous face studies have used relatively short moving clips (1-2.5 s, e.g. Lander & Bruce, 2000; Lander et al., 2001; Lander, Christie & Bruce, 1999; see Table 1), whereas studies on personally familiar, experimentally familiar and unfamiliar faces have often used longer clips – for example, Hill and Johnston’s (2001) clips had an average length of 7.2 s, Rosenblum et al. (2007) used 2-3 s clips, and Lander and Davies (2007) edited their test clips to be 5 s long (see Tables 2 and 3). In general, longer exposure to a moving face in the learning phase of a familiarisation study has lead to better recognition (Bruce & Valentine, 1988; Experiment 1; Lander & Davies, 2007, Experiment 2), however relatively little research has varied exposure time when testing famous, personally familiar, or unfamiliar faces; or when using a task other than overt or old/new recognition. Bruce et al. (1999, Experiment 2) found that allowing participants unlimited viewing of an unfamiliar moving face in a matching task gave rise to a movement advantage, whereas a single 5 s clip was recognised equally as well as a static comparison. This suggests that the duration of exposure to a face may play a role in face matching, but it is unclear what the optimum duration is and whether it differs for familiar and unfamiliar faces. It is possible that familiar or famous faces only require a short
excerpt of movement in order for characteristic movements and 3D structure to be extracted, whereas unfamiliar faces may require significantly longer exposure to efficiently extract structural information or to acquire a mental representation of “characteristic” motion patterns. Once again, without systematic study it is difficult to say whether the different pattern of results for familiar and experimentally familiar or unfamiliar faces arose due to familiarity, or because the studies used different stimuli, with different types of motion, presented for different lengths.

3.4.3 Recognising Moving Faces – Do Familiarity and Methodology Matter?

There are compelling reasons to believe that familiar and unfamiliar faces benefit from movement in different ways. Familiar faces appear to be able to use characteristic motion patterns as a pathway to recognition. However, it is unclear what role, if any, structure-from-motion plays in movement-based familiar face identification. On the other hand, evidence for a movement advantage in experimentally familiar and unfamiliar faces is mixed. Some studies have found that movement facilitates face learning or matching, and there is evidence that characteristic motion (or timing) patterns and structure-from-motion can both support unfamiliar face matching. However, the exact details – such as whether motion is helpful at learning or test, and what type of motion is useful – are extremely unclear. Direct comparisons of moving familiar and unfamiliar faces are rare, so at this stage it is unclear whether familiarity results in a larger movement advantage, or whether familiar faces benefit from structure-from-motion or characteristic motion patterns more or less than unfamiliar faces. A comparison of movement in familiar and unfamiliar faces is further hindered by differences in the procedure and stimuli of various studies. In order to clarify the role of movement in familiar and unfamiliar face recognition, it is necessary to control and investigate these methodological factors – for example, the effect of task difficulty and stimulus type – and to directly compare matching and recognition performance for moving familiar and unfamiliar faces across a range of tasks. The remainder of the thesis presents a series of experiments that examine the role of familiarity in movement-based face recognition, using standard tasks and stimuli, with the aim of determining whether familiar and unfamiliar faces show a movement advantage that is qualitatively different, quantitatively different, or both.
Chapter 4

The Role of Movement in Famous and Unfamiliar Face Matching

Time, Task and Task Difficulty
CHAPTER 4: THE ROLE OF MOVEMENT IN FAMOUS AND UNFAMILIAR FACE MATCHING: TIME, TASK AND TASK DIFFICULTY

As reviewed in Chapter 3, there have been a number of studies addressing the role of motion in face processing, but it is unclear whether familiarity with a face moderates the movement advantage for face recognition. One of the reasons that the role of familiarity in movement-based face recognition is unclear is that very few studies in this area have directly compared familiar and unfamiliar faces. Another reason the role of familiarity is unclear is that the methodology used in studies of famous, personally familiar, experimentally familiar and unfamiliar faces has varied widely, which makes comparisons across studies difficult.

Experiments 1 and 2 were designed test the effect of familiarity in movement-based face matching, and to address some of the methodological concerns that were raised in the previous sections. Primarily, these experiments aimed to clarify whether participants show a movement advantage for both familiar (in this case famous) and unfamiliar faces, and if so, whether the size of the movement advantage is comparable for both types of face. To assess the movement advantage, participants completed either a match-to-sample or a same/different identity task with moving and static sample images (unmanipulated, non-degraded faces) and moving and static test images (point-light-displays, or PLDs). Using both moving and static images in the sample and test positions maximizes the possibility of finding a movement advantage, and facilitates comparisons with previous studies that have not included all possible variations of moving and static images – for example, Hill and Johnston (2001), which only contained moving clips; or Thornton and Kourtzi (2002), which only tested matching from a moving clip to a static image.

The choice of stimuli in Experiments 1 and 2 was designed to maximise the possibility of finding a movement and familiarity advantage, whilst limiting the contribution of static facial information. Non-degraded images were chosen as the sample stimuli for two reasons. The primary reason for presenting non-degraded images in each trial was to maximize the effect of familiarity. It is possible that viewing a familiar face allows us to access memories of stored characteristic motion patterns – what Roark et al. (2003) refer to as the supplemental information...
hypothesis – which would facilitate comparison between remembered motion patterns and the motion presented in the PLDs. Non-degraded images also present a large amount of structural information, which may be important if the movement advantage relies on structure-from-motion cues. As well as maximizing the effect of familiarity and presenting structural cues, using non-degraded images might also encourage participants to approach the task as a face processing task, rather than an exercise in matching abstract dot patterns.

PLDs were chosen as the test stimuli because they should limit static recognition of faces (e.g., Rosenblum et al., 2002, 2007). Consequently, identity judgments should be based on movement alone. One caveat to this assumption is that famous people, who are generally highly familiar with cameras and appearing in the media, may have more consistent or predictable poses and expressions. It is possible that we are better at recognising famous faces when they are in typical poses, holding a typical expression (or “iconic” images; Carbon, 2008). This is particularly true in a matching task, where participants are provided with a second image for comparison. Previous studies of famous faces have often used static images with a relatively neutral expression and uniform pose (e.g., Knight & Johnston, 1997; Lander et al., 2001; Lander et al., 1999, Experiment 1; Lander & Chuang, 2005) – for example, the criteria for several studies was that static frames “avoided any unusual, momentary expression, as well as unusual head movement or angle” (Lander et al., 1999, p. 976). This could have artificially inflated the movement advantage in those studies in two ways – firstly, by depriving participants of an extra static identity cue (e.g., characteristic head and face poses), and secondly, by reducing and/or controlling the diversity of static images without applying the same constraints to the dynamic images. Some studies have attempted to address this issue by presenting equal numbers of static frames in an array or out of order (Lander & Bruce, 2000; Lander et al., 1999). While it is true that this approach preserves characteristic static information, the mode of presentation (with multiple face images shown at once) or the unnatural movement style may have been distracting, or may have prevented participants from processing all the frames and/or concentrating on identifying the face. In the current studies participants viewed single static frames, extracted from the centre of the moving clip, in order to preserve any characteristic static information that may be present in a single frame and to make the task as simple as possible.
Research to date has not systematically investigated the duration of movement that is necessary to support identification or matching, or whether the amount is different for familiar and unfamiliar faces. Although some studies have found an effect of movement with clips as short as 1 s (Pilz et al., 2006) or as long as 7 – 8 s (Hill & Johnston, 2001; Knappmeyer et al., 2003), it is still unclear whether the duration of a moving clip affects the movement advantage. Several studies using experimentally familiar faces found that the duration of the learning phase does not have any impact on movement advantage: Bruce and Valentine (1988) and Bruce et al. (2001) found no benefit of learning a face in motion, regardless of the duration. On the other hand, Bruce et al. (1999) found that unfamiliar faces were matched more accurately from a moving clip when participants could watch it indefinitely than when they were restricted to a single 5 s viewing. To date, no studies on famous faces have investigated the effect of clip duration on the movement advantage. Therefore, in order to investigate the optimum duration of test stimuli, the PLDs in Experiments 1 and 2 were presented for 1, 2, 4 and 8 s. The sample images (moving and static) and the static PLDs were 2 s long.5

In addition to examining the effect of familiarity on the movement advantage, Experiments 1 and 2 also examine whether the type of task and task difficulty has an impact on the overall pattern of results. Given that a naming task is impractical for unfamiliar faces, this chapter compares two matching tasks: a match-to-sample task (Experiment 1) and a same/different task (Experiment 2). Task difficulty was varied by manipulating whether or not each “same identity” trial contained overlapping samples of movement. Half of the participants matched clips that contained overlapping movements sequences (Experiments 1a and 2a), and half the participants matched clips that featured different motion sequences (Experiments 1b and 2b).

Given that several other studies have found a movement advantage for famous faces (Knight & Johnston, 1997; Lander et al., 1999; Lander & Bruce, 2000; Lander et al., 2001), and matching performance is consistently better for familiar than unfamiliar faces regardless of movement (Bruce et al., 2001; Burton, Wilson et al., 1999), it was expected that participants would perform best when watching famous faces.

5 Although it would have been desirable to vary the duration of the sample image, the duration of the experiment would have been impractical, possibly leading to fatigue effects. As the sample images were not degraded, it was assumed that 2 s was enough time for participants to make a familiarity judgement and/or recognise the famous faces.
faces with dynamic sample and test sequences, particularly in the experiments with overlapping motion. Under these conditions, participants should be able to recognise and/or extract characteristic motion patterns and structural information in the first (non-degraded) sequence, then recognise and/or match these to motion patterns or structural cues in the second (PLD) sequence. Similar predictions were made for famous faces with a static sample image. If participants are able to access and match characteristic motion patterns from a moving famous face, it may be possible to access the same representations from a still image. Participants should also be able to compare structural information from a static sample image to information derived from structure-from-motion in the test image (Farivar et al., 2009), particularly if they have robust static representations of the face stored in memory. Therefore, participants were expected to perform well in all famous face conditions that contained moving test images.

For unfamiliar faces with moving sample and moving test images, previous research suggested several possibilities. If matching requires stored, familiar characteristic motions for each identity, participants should perform poorly in the unfamiliar face conditions, particularly in trials with non-overlapping movement sequences. However, if participants familiarise rapidly with individual motion patterns, or if they are able to use structure-from-motion information, it would be possible for participants to perform well when given dynamic non-degraded sample sequences.

In contrast to famous faces, unfamiliar faces should not have any pre-existing motion signatures to be accessed from a static image. Performance for unfamiliar faces with static sample images would therefore rely solely on extracting structure-from-motion cues from the PLDs. If participants can extract structure-from-motion cues efficiently and match them to unfamiliar faces, they should be able to match a static sample image to a moving PLD. There is evidence that structure-from-motion can support matching (Farivar et al., 2009), but Farivar et al.’s stimuli showed 45 degrees of rigid rotational movement. The current experiments used clips from television shows, and typically only showed a small amount of rigid head movement. Consequently, it was expected that performance in the unfamiliar, static sample image condition would be the poorest overall.

Since the primary rationale behind using PLDs rather than degraded videos was to minimise form cues in the test images, all conditions in which the PLD test
image was static were expected to result in chance or near-chance-level performance, regardless of familiarity.

4.1 Familiarity, Distinctiveness and Amount of Movement Pre-test Ratings

To ensure that stimuli were well matched on extraneous variables and that they were suitably famous or unfamiliar, a pilot study was conducted. A number of video clips showing natural movements during speech were selected and rated for familiarity, distinctiveness and amount of movement. The results of this pilot experiment were then used to create stimuli for Experiments 1 and 2 (in this chapter), Experiments 3 and 4 (Chapter 5) and Experiments 5, 6 and 7 (Chapter 6).

4.1.2 Methods

4.1.2.1 Participants. Thirty-four undergraduate students (27 female) from the University of Western Sydney, aged between 18 and 47 years (mean age 24 years) participated in this experiment in return for course credit. All reported normal or corrected-to-normal vision.

4.1.2.2 Stimuli and Materials. A set of video images of 21 adult males was obtained from the online content of two talk shows. The people chosen included actors, television personalities, politicians, and members of the general public. For each identity, three clips of eight seconds each were created using iMovie (Apple). All clips showed the person speaking, facing towards the camera in an interview situation. The clips were selected to show each face from approximately the same angle (an approximately frontal viewpoint) and distance from the camera, and to exclude extreme facial movements. If a clip of eight seconds could not be found, two shorter clips were edited together, and 3 black frames were inserted between them. Three black frames were inserted at the end of each clip.

The resulting images were divided into three sets. Each set contained one video of each of the 21 identities. All videos measured 960 x 540 pixels, and were presented on a white background. All experiments were run on a MacBook Pro using Superlab 4.0.3, and images were presented on a BENQ E2200 HD 22-inch monitor, with resolution set to 1920 x 1080 pixels.
4.1.2.3 Procedure. Participants completed three rating tasks: familiarity, distinctiveness, and amount of movement, always in that order. One set of 21 videos was used for each rating task, and the set used for each task was counterbalanced across participants. Within each set, the order of presentation of the videos was randomised.

First, participants were asked to rate how familiar the person in each video was to them, on a scale of 1 (completely unfamiliar) to 7 (very familiar). After they had entered their rating, participants were asked to provide the person’s name or some other uniquely identifying information about them, if known. If a face was completely unfamiliar, participants could press enter to continue without entering any information. After completing the familiarity ratings, participants were asked to rate the distinctiveness of the movement of the individuals face and head, on a scale of 1 (not at all distinctive) to 7 (very distinctive). Distinctiveness was defined as how much a movement would stand out in a crowd, or how individual or idiosyncratic it was to the person. Participants were instructed to try and ignore how familiar the person was, or how distinctive their face shape or features may be, and give a rating based purely on the movement of the face and head. Finally, participants were asked to rate each video for the amount of movement of the face and head, on a scale of 1 (no movement) to 7 (a lot of movement). Once again, participants were instructed to try and ignore how familiar the person in the video was to them, and how distinctive their movements were, and concentrate purely on how much the face and head moved.

For all tasks, participants had to wait until the full eight-second clip had finished before giving their rating. All responses were entered on a computer keyboard. Participants had unlimited time to enter and edit their responses once the video ended. Once they were happy with their response, participants pressed the enter key to move onto the next trial.

4.1.3 Results

Due to computer problems recording the data, the final analysis included familiarity ratings and identification from 23 participants, distinctiveness ratings from 29 participants, and amount of movement ratings for 32 participants.

Participants were able to correctly name or give other unambiguous identifying information for between one and eleven of the identities (M = 6.56, SD = 2.29). The
video clips were split into two categories based on familiarity ratings and consistent
naming performance: the highest six and lowest six rated individuals (with consistent
naming performance) were retained as “famous” and “unfamiliar” faces for use in
Experiments 1 and 2 (this chapter) and Experiments 4 - 7 (Chapters 5 and 6). The
famous faces had a mean familiarity rating of 6.46, and were correctly identified by
at least 9 out of 23 participants in the naming task. The unfamiliar faces had a mean
familiarity rating of 1.43 and were correctly identified by none of the participants in
the naming task. Separate two-way ANOVAs (2 x familiarity; 3 x set) were carried
out on the distinctiveness and amount of movement ratings. No main effects or
interactions were significant in either analysis (all $F$s < 1), confirming that on
average the famous and unfamiliar clips were rated as equally distinctive and had
equal amounts of movement present.

Finally, one eight-second clip was selected for each famous and unfamiliar
identity. The average familiarity, distinctiveness and amount of movement rating for
the final clips are shown in Table 5. The clips were paired based on similar ratings of
familiarity, distinctiveness and amount of movement. These pairings remained
consistent across Experiments 1, 2 and 3 (see Chapter 5), to limit the amount of non-
identity based cues that participants could use in the matching tasks, and negate any
response bias a participant might have displayed towards one particular clip.

4.2 Experiment 1: Match-to-sample Task

Experiment 1 used a sequential match-to-sample task. Match-to-sample tasks
have been established as an effective test of recognition in a number of studies of
face and cross-modal person identification (Hill & Johnston, 2001; Kamachi et al.,
2003). The match-to-sample testing format does not require overt naming – it can be
applied to both familiar and unfamiliar faces, so that the two may be directly
compared.

Before examining the effect of task and familiarity, it was necessary to check
the validity of the PLD stimuli. One of the primary reasons that PLDs were chosen in
this study was that they minimise form cues in moving images, by displaying only a
collection of dynamic points that represent the movement of the underlying face. It is
possible that important elements of facial motion are not captured in this display.
Table 5:  
*Mean ratings for the famous and unfamiliar face videos used in Experiments 1-4 (Chapters 4 and 5).*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Russell Crowe</th>
<th>Kyle Sandilands</th>
<th>Ian Dickson</th>
<th>Ben Stiller</th>
<th>Jerry Seinfeld</th>
<th>Dave Hughes</th>
<th>Unfamiliar 1</th>
<th>Unfamiliar 2</th>
<th>Unfamiliar 3</th>
<th>Unfamiliar 4</th>
<th>Unfamiliar 5</th>
<th>Unfamiliar 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Familiarity</td>
<td>6.78</td>
<td>6.65</td>
<td>6.56</td>
<td>6.56</td>
<td>6.35</td>
<td>5.83</td>
<td>1.65</td>
<td>1.61</td>
<td>1.43</td>
<td>1.43</td>
<td>1.35</td>
<td>1.09</td>
</tr>
<tr>
<td>SD</td>
<td>0.67</td>
<td>1.07</td>
<td>1.27</td>
<td>0.99</td>
<td>1.52</td>
<td>1.40</td>
<td>1.40</td>
<td>1.56</td>
<td>0.79</td>
<td>0.84</td>
<td>0.93</td>
<td>0.29</td>
</tr>
<tr>
<td>Distinctiveness</td>
<td>3.33</td>
<td>1.78</td>
<td>5.22</td>
<td>3.00</td>
<td>4.33</td>
<td>5.44</td>
<td>2.89</td>
<td>3.56</td>
<td>3.00</td>
<td>4.11</td>
<td>3.33</td>
<td>3.00</td>
</tr>
<tr>
<td>SD</td>
<td>1.66</td>
<td>0.98</td>
<td>1.56</td>
<td>1.12</td>
<td>1.32</td>
<td>1.33</td>
<td>1.17</td>
<td>1.67</td>
<td>1.12</td>
<td>1.97</td>
<td>1.80</td>
<td>1.32</td>
</tr>
<tr>
<td>Amount</td>
<td>4.40</td>
<td>2.80</td>
<td>5.90</td>
<td>3.00</td>
<td>4.70</td>
<td>5.70</td>
<td>4.50</td>
<td>4.80</td>
<td>4.70</td>
<td>5.40</td>
<td>3.90</td>
<td>3.50</td>
</tr>
<tr>
<td>SD</td>
<td>1.43</td>
<td>1.14</td>
<td>1.10</td>
<td>1.41</td>
<td>1.43</td>
<td>1.95</td>
<td>1.58</td>
<td>1.55</td>
<td>1.70</td>
<td>1.26</td>
<td>1.37</td>
<td>1.35</td>
</tr>
</tbody>
</table>

*Note:* Familiarity ratings were averaged across all three clips of each individual. Distinctiveness and amount of movement ratings are reported for the final clip only. Famous pairings: Crowe and Seinfeld; Sandilands and Stiller; Dickson and Hughes. Unfamiliar pairings: Unfamiliar 1 and 3; Unfamiliar 2 and 4; Unfamiliar 5 and 6.
(Knight & Johnston, 1997), or that participants may have trouble generalizing from a full video image to the more sparse PLD. To test whether participants would have trouble extracting movement information from the PLDs, Experiment 1a examined whether participants could match faces when the same movement was present in the non-degraded sample image and the PLD test image. In this case, participants could perform the task by matching the movement alone, rather than extracting identity information from sample image and comparing it to the movement or structure of the PLD. In Experiment 1b the same match-to-sample task was used, but different video clips of the same person were used to create the sample and test images. Therefore, in Experiment 1b participants were asked to generalise across format and movement in this task. Experiment 1b eliminated the opportunity to match based on superficial movement patterns – in order to perform well at the task, participants had to match based on identity, using either characteristic movement patterns or the structural information present in the PLDs.

4.2.1 Methods of Experiments 1a and 1b

4.2.1.1 Participants. Fifteen undergraduate students (14 female), aged between 17 and 46 years (mean age 23 years) participated in Experiment 1a. One participant’s data was excluded from analysis due to invalid responses throughout the experiment. Eighteen undergraduate students (15 female), aged between 18 and 46 years (mean age 24 years) participated in Experiment 1b. All participants were recruited from the University of Western Sydney, and participated in this experiment in return for course credit. All reported normal or corrected-to-normal vision.

4.2.1.2 Stimuli. The famous and unfamiliar face videos were chosen and paired with distractors as described in section 4.1.3. Each video was edited so the face was approximately 8 cm wide. Point-light displays (PLDs) were created by tracking the movement of 27 facial regions using Motion (Apple), then superimposing small grey dots onto each point that mimicked the movement of the underlying region. Finally, the background image was set to black. The location of the points was based on the PLDs used by Hill, Jinno and Johnston (2003), with additional points added to the cheeks, orbits and temples to make the image more “face-like” (see Figure 7). Pupils were not included, as the tracking was not accurate enough to follow eye-movements. This resulted in one eight-second video per person.
displaying 27 dots that mirrored the movement of the individual’s face. These 8 s PLDs were the basis for all the PLD clips used in all subsequent experiments.

In order to test the effect of duration, PLDs of 1 s, 2 s and 4 s were created for each identity. For Experiments 1a and 2a, participants were shown clips featuring overlapping movements. The shorter video clips were created by cutting and exporting the middle 4 s, 2 s and 1 s from the 8 s PLD. To create static images, a single frame was exported from the middle of each 8 s PLD clip using Final Cut Pro (Apple). The middle sections of the clips were chosen to avoid the possibility that participants would learn just the beginning or end of one clip, and then use that information as a cue to identity in all trials. For Experiments 1b and 2b, participants were required to match clips that did not feature overlapping movement. Two versions of the shorter clips were created to ensure participants were not simply learning the movements in the clips, or using one particularly distinctive movement to identify a person. For the 4 s clips, new clips were created from seconds 1-4 and 5-8 of the 8 s PLD; for the 2 s clips seconds 3-4 and 7-8, and for the 1 s clips the 2\textsuperscript{nd} and 6\textsuperscript{th} seconds of the full 8 s clip. The static frames were extracted from the middle of the first and eighth seconds.

Figure 7: Placement of dots used when creating point-light-displays (left) and a point-light-display (right).
Two versions of Experiments 1b and 2b were created, using the short clips and static frames that did not overlap with one-another (all test PLDs overlapped with the eight-second clip). Three black frames were added to the end of each clip.

To create the sample (non-degraded) clips for Experiments 1a and 2a, the middle two seconds of the original video clips were exported using Final Cut Pro, and three black frames added to the end of each clip. The sample clips for Experiments 1b and 2b were extracted from another set of videos that had not been used in the creation of the PLD, but which received similar ratings of distinctiveness and amount of movement as the non-degraded videos from which the PLDs were derived.

This resulted in three sets of five test PLD images (1 s, 2 s, 4 s, 8 s, 2 s static), and two sets of full-video sample images (2 s, 2 s static) for each identity. Examples of the sample and PLD clips are included in Appendix A. In Experiments 1a and 2a, moving sample images and PLDs of the same identity overlapped by at least one second (in the 1s PLD), and at most 2 seconds (for all other lengths of PLD). For the static sample and test images, the frame used was always the same, so the full video and PLD static images showed exactly the same facial pose. In Experiments 1b and 2b, there was no overlap between the sample and test images, either moving or static. All final video and static stimuli measured 960 x 540 pixels, and were presented on a black background. All videos were presented at 25 frames per second. All experiments were run on a MacBook Pro using Superlab 4.0.3, and images were presented on a BENQ E2200 HD 22-inch monitor, with resolution set to 1920 x 1080 pixels. Participants were tested at an approximate viewing distance of 60 cm.

4.2.1.3 Design and Procedure. Each experiment was a repeated measures design, 2 (familiarity: famous/unfamiliar) x 2 (movement of non-degraded sample clip: moving/static) x 5 (duration of test clip = 1 s, 2 s, 4 s, 8 s or static PLD). Each condition contained six trials (one for each identity), resulting in a total of 120 trials per participant, presented in a random order. In Experiment 1b, a between-subjects factor of Version (2 levels) was added. However, the analyses revealed no main effects or interactions in any condition, so Version was excluded from any further analysis.

An example of a match-to-sample trial is shown in Figure 8. Every trial of the match-to-sample task began with a fixation cross presented in the centre of the
screen for 200 ms. This was followed by the sample (non-degraded) stimulus, presented for 2 s in the centre of the screen. In half the trials, the sample stimulus was a 2 s video. In the other half, the sample stimulus was a single static image, shown for 2 s. The sample image was followed by a 500 ms mask consisting of greyscale noise. Participants then viewed two test PLD images sequentially, one of which depicted the same person as the sample image, and the other of which depicted that person’s designated pair (matched on distinctiveness and amount of movement ratings). The PLDs could consist of static images presented for 2 s each, or video clips lasting 1, 2, 4 or 8 s each. The two test PLDs were always the same length. To prevent matching based purely on size or location of the image, each PLD was offset from the centre of the screen by 60 pixels in opposing directions. The two test PLDs were presented in random order, and separated by a 200 ms blank screen. Participants were asked to wait until the second PLD had finished, and then indicate via a keypress which of the PLD’s (1st or 2nd) depicted the same person as the initial video. Participants were instructed to respond as quickly and as accurately as possible. Once the keypress was recorded, the next trial began immediately.

Figure 8: The timeline of a match-to-sample trial in Experiment 1. Same/different trials (Experiment 2) followed the same timeline, except the second PLD was omitted, and the final screen displayed the text “same or different?”
Participants were tested individually in a darkened room. They completed 10 practice trials (without feedback) prior to beginning the full experiment, and received two breaks throughout the experiment. Following the main experiment, each participant was shown one video each of the 12 identities, and asked to rate the video for familiarity on a seven point scale, and to name the person (or provide other unambiguous identity information) if they were famous. Participants’ data were excluded from analysis if they did not rate all famous faces 6 or higher and unfamiliar faces 2 or lower, or if they could not name at least three of the six famous faces.

4.2.2 Results and Discussion of Experiment 1a

4.2.2.1 Accuracy. Three-way within-subjects ANOVAs (2 x familiarity; 2 x moving or static sample image; 5 x length of test stimulus; 1 s, 2 s, 4 s, 8 s, static) were performed on both accuracy and d’ scores (see section 4.3.2), but as the results were broadly consistent, the results section will focus on accuracy. Overall accuracy scores (proportion correct) for Experiment 1a are displayed in Table 6. Two-tailed t-tests were carried out to compare performance to chance levels (0.5), and the results are shown in Table 6. The ANOVA on the accuracy scores showed that accuracy for famous faces was higher than unfamiliar faces, $F(1,13) = 48.87, p < .0005, \eta_p^2 = .79$, and participants were better at matching moving sample images than static sample images, $F(1,13) = 23.77, p < .0005, \eta_p^2 = .65$. There was also a main effect of length, $F(4,52) = 6.12, p < .0005, \eta_p^2 = .32$, which will be discussed further in the movement advantage results. No interactions were significant, $ps > .05$.

4.2.2.2 Movement advantage. In order to assess the movement advantage (i.e., how much better participants’ matching performance was for moving rather than static point-light images), further analysis was conducted by subtracting each participant’s score for the static PLD in each condition from the scores for each clip length. For example, the score for static PLDs paired with moving sample clips of famous faces was subtracted from the score for 1 s, 2 s, 4 s, and 8 s PLDs with moving sample images of famous faces, and this was repeated for both levels of familiarity and movement of sample clip.
The difference scores (movement advantage statistics) for Experiment 1a are shown in Figure 9. Two-tailed t-tests were used to compare the movement advantage to 0. The overall mean movement advantage score was significantly greater than 0, M = 0.06, t(13) = 2.20, p = .047. This confirms the general hypothesis that participants perform significantly better on average in the conditions with moving PLDs than the conditions with static PLDs. Despite this overall finding, very few individual conditions had a significant movement advantage, suggesting that the movement advantage was small and somewhat unstable in this experiment. However, the movement advantage for clips with moving sample images (collapsed across other conditions) was significant, M=0.11, p = .045. There was no mean movement advantage for static sample clips, p = .784. Likewise, there was no mean movement advantage for famous faces, p = .620. Interestingly, the mean movement advantage for unfamiliar faces approached significance, M=0.10, p = .058.

The movement advantage was analysed with three-way within-subjects ANOVAs (2 x familiarity; 2 x moving or static sample image; 4 x length of test stimulus) for both accuracy and d’. As in the overall analysis, the movement advantage for accuracy and d’ was broadly consistent, so the results report accuracy only. Bonferroni corrections were applied to pairwise comparisons throughout this and all experiments.
The ANOVA conducted on movement advantage scores showed a significant main effect of length, $F(3,11) = 10.6, p < .0005, \eta_p^2 = .50$. Pairwise comparisons on length show that 1 s, 2 s and 4 s clips did not significantly differ from one another, but had a significantly greater movement advantage than 8 s clips ($ps < .02$). The movement by length interaction was also significant, $F(3,39) = 3.96, p = .015 \eta_p^2 = .23$. Pairwise comparisons show that the pattern of results for trials with moving sample clips mirrors the main effect of length, with 1 s, 2 s and 4 s test clips showing a greater movement advantage than 8 s clips, $ps < .012$, whereas trials with a static sample clip had a similar-sized movement advantage regardless of the length of the test clip. No other effects were significant, $ps > .17$.

4.2.2.3 Discussion of Experiment 1a. The results from Experiment 1a suggest that participants are fairly accurate at matching movements across formats (i.e., from full video to PLD). As expected, famous faces were matched more accurately than unfamiliar faces, and a dynamic non-degraded sample clip facilitated matching more
than a static sample clip. Analysis of the movement advantage showed that matching was more accurate overall if the test clip (PLD) was moving. The movement advantage was greatest for 1 s, 2 s and 4 s clips, particularly when they were being matched to a moving, non-degraded sample clip. This supports the contention that the motion advantage may be most apparent at short intervals. The effect of length disappeared when the non-degraded sample clip was static, which suggests that this motion advantage is based primarily on matching movements and/or movement style, rather than providing extra static information.

It is somewhat surprising that the movement advantage was equally large for famous and unfamiliar faces, despite the fact that famous faces were matched more accurately overall. These results probably reflect the high performance in some conditions with static PLDs, particularly when the non-degraded sample face was famous. This may be a result of the static frames being matched to or extracted from the sample clip.

The overall results from Experiment 1a indicate that our PLD stimuli carry enough movement information to facilitate movement matching and to give rise to a movement advantage. Experiment 1b investigated whether PLDs could also be used to match identities, when the sample and test clip did not contain the same movements.

4.2.3 Results and Discussion of Experiment 1b

The data from Experiment 1b underwent the same analyses as Experiment 1a. As in Experiment 1a, two-tailed t-tests were carried out to compare accuracy scores to chance levels (0.5), and the overall accuracy scores and t-test results are presented in Table 7.

Performance in Experiment 1b was dramatically worse than in Experiment 1a, and there were no main effects or interactions in the analysis of accuracy, familiarity: $F(1,17) = 0.001, p = .970, \eta^2_p = .00$; movement: $F(1,17) = 0.26, p = .617, \eta^2_p = .02$; length: $F(4,68) = 0.46, p = .767, \eta^2_p = .03$; all interactions $ps > .1$.

Movement advantage scores are illustrated in Figure 10. There was no significant movement advantage overall, and no movement advantage in any single condition, $ps > .15$. 

Chapter 4: Time, Task, and Difficulty
The only significant effect was a familiarity by length interaction, $F(3,51) = 4.38, p = .008, \eta_p^2 = .20$, which shows a greater movement disadvantage (i.e. participants were worse at matching moving PLDs than static PLDs) for 8 s clips of unfamiliar faces than famous faces. All other main effects and interactions were non-significant, familiarity: $F(1,17) = 0.01, p = .928, \eta_p^2 = .00$; movement: $F(1,17) = 1.19, p = .291, \eta_p^2 = .06$; length: $F(3,51) = 0.32, p = .811, \eta_p^2 = .02$; all other interactions $ps > .1$.

Effect sizes were greatly reduced when compared to Experiment 1a: the maximum partial eta squared was .20, as opposed to .79 in Experiment 1a. When participants were shown different clips of the same people, their ability to use movement as a cue to identity was severely compromised, even with faces that they rated as highly familiar. It is possible that these results reflect participants’ difficulty remembering and generalizing motion across varying lengths of time (up to 16 seconds in the 8-second conditions), or that there was not enough movement information present in the clips to match them (particularly the 1- and 2-second clips). Given the movement advantage in Experiment 1a was small, either factor could be enough to result in the lack of significant differences in this experiment.
4.2.4 Discussion of Experiment 1

The results of Experiment 1 suggest that participants can match non-degraded images to moving PLD faces in a match-to-sample task, but only when the two images show overlapping movement sequences. There are several possible explanations for these results. First, it is possible that participants cannot match identity across different formats. This explanation seems unlikely, as several successful experimental familiarisation studies have used different formats for learning and test phases (Bruce et al., 2001; Lander & Davies, 2007; Roark et al., 2006). Furthermore, Rosenblum et al. (2002) used similar non-degraded and PLD stimuli of unfamiliar faces and found that participants were able to match the faces based on identity. It is important to note, though, that the stimuli in this study were matched for distinctiveness and amount of movement, which may have limited the types of movement information participants could rely on (i.e., they had to rely on patterns of motion, rather than simply looking at how much a face moved). Previous studies have not explicitly controlled stimuli in this manner, which may explain why the movement advantage in Experiment 1 was small or non-existent.

Figure 10: The movement advantage for different clip lengths in Experiment 1b. The movement advantage was calculated as the difference in accuracy between moving PLD and static PLD test trials. Error bars represent +/- 1 standard error of the mean.
An alternative explanation for the poor results in Experiment 1b is the length of the trials – in the longer (8 s) trials, participants were required to remember movement patterns for more than ten seconds in order to perform a match. If it was simply a matter of trial length, participants should have performed significantly better in the shorter (1 s and 2 s) trials than the longer trials. Experiment 1a showed this pattern of results in the movement advantage analysis. However, there was no indication that trial duration impacted performance in Experiment 1b, suggesting that trial length by itself cannot account for the poor performance in that task. Finally, it is possible that the match-to-sample task was too demanding, since it required participants to compare three separate movement patterns in each trial. When the task only required participants to look for a specific movement pattern, participants could compare all three clips effectively. However, it may have been more difficult to extract and compare characteristic movement patterns, especially when participants were also asked to generalise across formats.

Since participants in Experiment 1 showed little evidence of matching based on identity, Experiment 2 used a different task. Participants were still required to match an unmanipulated video or static image to a PLD based on identity. However, participants in Experiment 2 completed a same/different task, which resulted in shorter trials, and less comparisons to be made between clips. Importantly, the same/different task still allowed the effect of familiarity, movement and clip duration to be examined.

### 4.3 Experiment 2: Same/different Task

Experiment 2 was designed to assess whether a same/different task, with shorter trials and fewer movement sequences to remember, would increase the movement advantage found in Experiment 1. While the demands of a same/different task are very similar to a match-to-sample task (i.e. matching sequentially presented images, rather than naming or old/new recognition), participants only have to make a decision based on one test clip, rather than comparing two clips and evaluating which is more similar to the sample. The memory demands are also reduced, as there is only one test clip, resulting in much shorter trials and fewer distracting movement patterns. Therefore, if the results of Experiment 1b arose due to task or memory demands, participants in the same/different task were expected to perform better than
in the match-to-sample task, especially in the longer conditions (4 s and 8 s). If the lack of movement advantage in Experiment 1b was due to problems generalising across formats, or performance was limited because the stimuli were matched for amount and distinctiveness of movement, Experiment 2 was expected to result in a similar pattern of results to Experiment 1. As in Experiment 1, experiments with overlapping clips (Experiment 2a) and non-overlapping clips (Experiment 2b) were run.

4.3.1 Methods of Experiments 2a and 2b.

4.3.1.1 Participants. Seventeen undergraduate students (12 female), aged between 17 and 38 years (mean age 21 years) participated in Experiment 2a. Nineteen undergraduate students (15 female), aged between 18 and 47 years (mean age 23 years) participated in Experiment 2b. All participants were recruited through the University of Western Sydney, and participated in return for course credit. All reported normal or corrected-to-normal vision.

4.3.1.2 Stimuli. All the stimuli used in Experiment 2 were identical to Experiment 1.

4.2.3.3 Design and Procedures. The design and procedure for Experiment 2 was identical to that of Experiment 1, except that participants completed a same/different task rather than match-to-sample. In Experiment 2 only one PLD was presented per trial, and participants were asked to indicate via keypress whether that PLD showed the same person as the initial full video sample image or a different person. As in Experiment 1, clips of each individual were always matched to the same person (e.g. the “different” image for Ben Stiller was always Kyle Sandilands) to limit the use of non-identity based cues. Each condition contained twelve trials (one same and one different trial for each identity), resulting in a total of 240 trials per participant, presented in a random order. Given the difficulty participants had in completing Experiment 1b (with non-overlapping sample and test images), participants were given twice as many practice trials for Experiment 2b as for all other experiments. As in Experiment 1b, a between-subjects factor of version (2 levels) was included in Experiment 2b. However, the analyses revealed no main
effects or interactions in any condition, so version was excluded from subsequent analyses.

4.3.2 Results and discussion of Experiment 2a

Using signal detection theory, a d’ score was calculated for each participant in each condition (Macmillan & Creelman, 2005). Using a sensitivity measure such as d’ is important in same/different experiments to limit the effect of bias in responding on the analysis – for example, it limits the effect of participants who have a bias towards responding “same” by factoring in the effect of false positives (i.e. responding “same” to a different trial). Hit and/or false positive scores of zero or one were replaced with values of 0.08 and 0.916 respectively (MacMillan & Kaplan, 1985). In this experiment, a d’ of 0 represents chance performance, while a d’ of 3.42 represents perfect performance in a condition. As in Experiment 1, we were also interested in the movement advantage, which was calculated as the difference between the d’ for the static and moving PLDs in each condition (i.e., for each level of familiarity, movement of sample clip and length of test clip). Separate analyses of bias (criterion c) and of hits and correct rejections were also carried out. A c of 0 indicates no bias, whereas a negative value of c indicates that participants had more false alarms (responding “same” to a different trial) than misses (responding “different” to a same trial). In order to simplify the results section, we will focus on d’, the movement advantage and bias, and note where the hits and correct rejection results differ or aid interpretation of the main analyses. Three-way within-subjects ANOVAs (2 x familiarity; 2 x moving or static sample image; 5 x length of test stimulus) were performed on the d’, bias and hit/correct rejection data. Three-way within-subjects ANOVAs (2 x familiarity; 2 x moving or static sample image; 4 x length of test stimulus) were performed on the movement advantage data.

4.3.2.1 Signal detection theory analysis. Two-tailed t-tests were carried out to compare each d’ to chance levels (0). The overall d’ scores and the results of the t-

---

6 Although it is possible to calculate bias statistics for a match-to-sample task, I chose not to do so in Experiment 1 because there is no meaningful difference between the two responses ("1st" or "2nd"). Therefore, the analysis of bias will be restricted to same/different tasks – Experiment 2, and Experiments 3 and 4 in Chapter 5.
tests are presented in Table 8. As in Experiment 1a, participants’ discrimination was quite good overall, but participants performed above chance levels in more conditions in Experiment 2a than in Experiment 1a. An ANOVA on the d’ scores showed that discrimination was better for famous faces than unfamiliar faces, \( F(1,16) = 29.65, p < .0005, \eta_p^2 = .65 \), and in general participants were better at matching moving sample images than static sample images, \( F(1,16) = 5.83, p = .030, \eta_p^2 = .26 \), regardless of whether the PLD was moving or static. These results are consistent with Experiment 1a. There was also a significant main effect of clip length, \( F(4,64) = 8.37, p < .0005, \eta_p^2 = .34 \), and a significant interaction between length and familiarity, \( F(4,64) = 2.81, p = .033, \eta_p^2 = .15 \). Famous faces were matched significantly better than unfamiliar faces in the 1 s and static test conditions, (1 s: \( p = .024 \); static: \( p < .0005 \)), and marginally better in the 8 s condition, \( p = .056 \), but there was no significant difference between famous and unfamiliar faces in the 2 s and 4 s conditions, \( ps > .1 \). Pairwise comparisons on the main effect of length will be discussed in the next section. No other interactions were significant, \( ps > .2 \).

Table 8: 
\textit{Mean d’ for Experiment 2a}

<table>
<thead>
<tr>
<th>Familiarity</th>
<th>Moving sample image</th>
<th>Static sample image</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 s</td>
<td>2 s</td>
</tr>
<tr>
<td>Famous</td>
<td>1.74**</td>
<td>1.91**</td>
</tr>
<tr>
<td>SD</td>
<td>1.07</td>
<td>0.86</td>
</tr>
<tr>
<td>Unfamiliar</td>
<td>1.26*</td>
<td>1.58**</td>
</tr>
<tr>
<td>SD</td>
<td>1.20</td>
<td>1.12</td>
</tr>
</tbody>
</table>

\textit{Note:} A score of 0 is chance, and the maximum is 3.42.  
*p < .05 ** p < .0005.
### 4.3.2.2 Movement advantage

The movement advantage values for Experiment 2a are shown in Figure 11. Participants showed a significant movement advantage, with an overall mean movement advantage score of 0.51, $t(16) = 3.892, p = .001$. When collapsed across other variables, unfamiliar faces showed a significant movement advantage, $M = 0.91, p < .0005$, whereas the movement advantage for famous faces was not significantly greater than 0, $M = 0.12, p = .575$. Trials that used moving non-degraded sample clips also showed a movement advantage, $M = 0.74, p < .0005$, whereas trials with static sample clips did not, $M = 0.28, p = .295$.

The ANOVA on movement advantage scores revealed several significant effects. As in Experiment 1a, the main effect of clip length was significant, $F(3,48) = 5.879, p = .001, \eta^2_p = .277$. Pairwise comparisons showed that 2 s clips were matched significantly better than 8 s clips, $p = .010$, and there was a trend for 2 s clips to be matched better than 1 s clips, $p = .061$. There was also a main effect of familiarity $F(1,16) = 6.420, p = .022, \eta^2_p = .286$. The movement advantage for unfamiliar faces was significantly higher than the movement advantage for famous faces, which was surprising given that famous faces were discriminated better overall. No interactions were significant, $ps > .25$.

![Figure 11](image)

**Figure 11:** The movement advantage for different clip lengths in Experiment 2a. The movement advantage was calculated as the difference in accuracy between moving PLD and static PLD test trials. Error bars represent +/- 1 standard error of the mean.
Bias, hits and correct rejections. Throughout Experiment 2a, participants displayed a bias towards responding “same”, $c = -.24$, SD = 0.55, illustrated by a tendency to make more false alarm errors than misses. Participants showed significantly more bias when viewing unfamiliar faces, $c = -.30$, than famous faces, $c = -.18$, $F(1,16) = 5.90, p = .027$, $\eta_p^2 = .27$. Interestingly, analyses of hit and correct rejection rates revealed that participants were not significantly more accurate when matching famous faces than unfamiliar faces, $F(1,16) = 1.57, p = .228$, $\eta_p^2 = .09$, but they were significantly better at correctly rejecting mismatched famous than unfamiliar faces, $F(1,16) = 18.88, p = .001$, $\eta_p^2 = .54$. Familiarity with a face appears to enhance performance by reducing the bias to report a match between faces, and increasing participants’ ability to reject incorrect matches.

Analyses on bias also showed a main effect of movement of the sample clip, $F(1,16) = 13.12, p = .002$, $\eta_p^2 = .45$, reflecting a greater bias towards “same” responses when the sample image was moving, $c = -.34$, then when the sample image was static, $c = -.14$. Moving sample clips resulted in significantly more hits than static sample clips, $F(1,16) = 24.02, p < .0005$, $\eta_p^2 = .60$, but movement of the sample clip had no significant effect on correct rejections, $F(1,16) = 3.18, p = .093$, $\eta_p^2 = .17$.

In contrast to the results for sample clip movement, movement of the PLD did not have a significant effect on bias, $F(1,16) = 1.19, p = .291$, $\eta_p^2 = .07$, although once again an interesting pattern of hits and correct rejections emerged. Participants made the same number of hits regardless of whether the PLD was moving or not, $F(1,16) = 1.36, p = .261$, $\eta_p^2 = .08$, but were significantly more accurate at correctly rejecting mismatched clips when the PLD was shown in motion, $F(1,16) = 6.46, p = .022$, $\eta_p^2 = .29$. The effect of movement for the sample clips and the PLDs is substantially different. It appears that, as with familiarity, movement of the PLD allowed participants to identify “different” clips more accurately, whereas movement of the comparison clip simply made participants more inclined to report a match.

Bias was not significantly affected by clip length, $F(4,64) = 0.90, p = .470$, $\eta_p^2 = .05$. However, length interacted with movement of the sample image, $F(4,64) = 5.54, p = .001$, $\eta_p^2 = .26$. Participants were significantly more biased to respond “same” for shorter clips when the non-degraded sample image was moving than static (1 s, 2 s and 4 s, all $ps < .025$). There was no difference in bias for moving and
static sample clips when 8 s and static PLDs were shown, $p > .1$. The pattern of hits and correct rejections supports the hypothesis that participants were more liberal in their responses when the clips showed superficial similarities: for example, when presented with moving sample images and short, moving PLDs. No other interactions in the analysis of bias were significant.

**4.3.2.4 Discussion of Experiment 2a.** Experiment 2a established that participants were able to match full video and PLD images somewhat more accurately in a same/different task than in a match-to-sample task. This may be a result of the shorter trials, or simpler task demands. However, it appears that both tasks reflect a similar underlying process, as the pattern of results was fairly similar between Experiments 1a and 2a. Both Experiments show an overall movement advantage, and we found that the movement advantage peaks around the 2 s clip, although this may be because the non-degraded sample image and PLD test clip in the 2 s condition show the same movement sequence. Unlike in Experiment 1a, we found a significantly larger movement advantage for unfamiliar faces than for famous faces. This is the first time, to our knowledge, that a larger movement advantage has been found for unfamiliar than familiar faces.

Experiment 2a also allowed a more in-depth analysis of the factors that impacted on participants’ responses. It appears that elements that would be expected to help participants perform well in this task (e.g. familiarity with a face, movement of the PLD test stimulus) do not simply increase the general discriminability of images, they specifically increase the number of correct rejections participants make. The exception to this is the movement of a non-degraded sample clip, which appears to increase bias to respond “same” (particularly for shorter, moving PLD test clips). Given that participants were not told that the movement of the sample and test clip would overlap, it is possible that they were deliberately applying very lenient criteria to the matching task in order to maximize their correct identifications. Despite this bias, participants were still more effective at discriminating images when the sample clip was moving, as shown in the d’ analysis, but there was no difference in movement advantage when matching from a moving or static non-degraded sample image. In general, these results support the idea that movement can play a role in matching faces when there is overlap between sample and test. In Experiment 2b, we
tested whether participants could use movement as a cue to identity when matching across format and across movement in a same/different identity task.

4.3.3 Results and discussion of Experiment 2b

4.3.3.1 Signal detection theory analysis and the movement advantage. The d’ values for Experiment 2b, and results from two-tailed t-tests comparing them to chance (0) are shown in Table 9. Similarly to Experiment 1b, performance in Experiment 2b was extremely poor; only a handful of conditions were significantly above chance, and no main effects or interactions were significant in the d’ analysis, familiarity: $F(1,18) = 0.02, p = .885, \eta^2_p = .001$; movement: $F(1,18) = 0.31, p = .59, \eta^2_p = .02$; length: $F(4,72) = 1.20, p = .317, \eta^2_p = .06$; all interactions $ps > 0.3$).

The movement advantage in Experiment 2b is illustrated in Figure 12. As in Experiment 1b, no individual conditions showed a significant movement advantage (in fact, unfamiliar faces with static comparison clips and 8 s PLDs showed a significant movement disadvantage, $p = .009$), and the analysis of movement advantage revealed no significant main effects or interactions for familiarity or length of clip, familiarity: $F(1,18) = 0.44, p = .517, \eta^2_p = .02$; length: $F(3,54) = 0.54, p = .654, \eta^2_p = .03$; all interactions $ps > .3$).

Table 9: Mean d’ for Experiment 2b

<table>
<thead>
<tr>
<th>Familiarity</th>
<th>Moving sample image</th>
<th>Static sample image</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 s</td>
<td>2 s</td>
</tr>
<tr>
<td>Famous</td>
<td>0.40</td>
<td>0.36</td>
</tr>
<tr>
<td>SD</td>
<td>1.44</td>
<td>1.60</td>
</tr>
<tr>
<td>Unfamiliar</td>
<td>0.30</td>
<td>0.52</td>
</tr>
<tr>
<td>SD</td>
<td>1.08</td>
<td>1.42</td>
</tr>
</tbody>
</table>

*Note: A score of 0 is chance, and the maximum is 3.42.
*p < .05 ** p < .0005.
Moving non-degraded sample clips, M = 0.05, resulted in a marginally greater movement advantage than static sample images, M = -0.56, $F(1,18) = 4.19, p = .056$, $\eta^2_p = .19$, but this was primarily to do with a negative movement advantage for static sample images. As in Experiment 1b, effect sizes were extremely small: the maximum partial eta squared value was .19, as opposed to .65 in Experiment 2a and .20 in Experiment 1b.

4.3.3.2 Bias, hits and correct rejections. In general, participants displayed a relatively low level of bias in Experiment 2b, c = .06, compared to Experiment 2a, c = -.24, which may reflect a higher degree of uncertainty due to the more difficult task.

Unlike Experiment 2a, familiarity did not have any significant effect on bias, hits or correct rejections, ps > .3, and movement of the sample clip only resulted in a marginal change in bias, $F(1,18) = 4.17, p = .056$, $\eta^2_p = .19$. When moving sample clips were shown, participants were relatively neutral, c = -.003, whereas presenting static sample images resulted in a tendency to respond “different”, c = .122.

![Figure 12](image-url): The movement advantage for different clip lengths in Experiment 2b. The movement advantage was calculated as the difference in accuracy between moving PLD and static PLD test trials. Error bars represent +/- 1 standard error of the mean.
These results, along with Experiment 2a, suggest that the mode of presentation of a sample image can have a significant effect on how participants are inclined to respond, with static sample images prompting more conservative criterions compared to moving sample images.

Length of the PLD test clip had no effect on bias, nor did overall movement of the PLD, both $ps > .1$. In general, the criterion scores in this analysis support the suggestion that participants were generally more cautious and less likely to report a match here than in the previous study. Finally, there was also a near-significant movement by length interaction, $F(1.69,72) = 3.47$, $p = .051$, $\eta_p^2 = .16$; Greenhouse-Geisser correction applied. Once again, the pattern of bias closely follows that found in Experiment 2a. Dynamic non-degraded sample images paired with 1 s and 2 s PLDs resulted in more lenient criteria (increased “same” responses) than static sample images, $ps < .025$. On the other hand, this pattern was almost reversed for static non-degraded sample images, which showed a marginally increased bias towards “same” responses when paired with static PLDs, $p = .084$.

4.3.3.3 Discussion of Experiment 2b. Overall, the d’ and movement advantage results for Experiment 2b confirm and extend the findings of Experiment 1b, and suggest that showing people different samples of movement from the same person can actually result in worse discriminability than extremely degraded static images alone. It is possible that participants were unable to generalize movement in these experiments because of the difficulty matching across different movement samples, combined with the difference in the visual information available in the sample (full-video) and test (point-light) images. However, analyses of bias and comparisons with Experiment 2a suggest that participants were processing the faces and reacting to the trials in a similar (although slightly more conservative) manner, despite the increased difficulty of the task. Given the similarity in the patterns of bias, it is unlikely participants were simply giving up and responding randomly to the task. One striking element of the analysis is that the effect of movement on bias is not consistent, but depends somewhat on which clip is moving (the sample or test clip), and in some cases whether both images are moving or still.
4.4 Combined Results of Experiments 1 and 2

In order to determine whether there are consistent overall effects of movement, familiarity and task on dynamic face matching performance, the discriminability and movement advantage scores (based on d’ scores) for Experiments 1 and 2 were subjected to a combined analysis.

4.4.1 Combined Signal Detection Theory Analysis

A mixed ANOVA was carried out on the combined d’ scored, with three within-subjects factors (2 x familiarity; 2 x moving or static sample image; 5 x length of test stimulus), and two between-subjects factors (2 x task; MTS and same/different; 2 x overlap; overlapping and non-overlapping clips). There was a significant effect of overlap, $F(1,64) = 46.03, p < .0005, \eta^2_p = .42$, indicating that overlapping movement sequences (Experiments 1a and 2a) lead to better matching performance than non-overlapping clips (Experiments 1b and 2b). The main effect of task was not significant, $F(1,64) = 0.23, p = .635, \eta^2_p = .00$.

The combined analysis revealed a significant main effect of familiarity, $F(1,64) = 26.69, p < .0005, \eta^2_p = .29$, and a significant familiarity by overlap interaction, $F(1,64) = 25.47, p < .0005, \eta^2_p = .28$. Famous faces were matched better than unfamiliar faces when the clips overlapped, $p < .0005$, but there was no significant difference when the clips showed separate movement sequences, $p = .929$. There was also a significant main effect of movement, $F(1,64) = 12.49, p = .001, \eta^2_p = .16$, and a movement by overlap interaction, $F(1,64) = 10.30, p = .002, \eta^2_p = .14$. The movement by overlap interaction mirrored the familiarity by overlap interaction: when the movements overlapped, moving sample clips were matched better than static sample clips, $p < .0005$, but there was no difference when the clips showed separate movement sequences, $p = .810$. Similarly, the main effect of length, $F(4,256) = 6.00, p < .0005, \eta^2_p = .08$, was qualified by an interaction between length and overlap, $F(4,256) = 6.86, p < .0005, \eta^2_p = .10$. Clips of 1s, 2 s and 4 s were matched better than 8 s and static clips in experiments with overlap, all ps < .05, but all lengths were matched equally well in the non-overlapping experiments, all ps = 1.

There were several other significant interactions, familiarity x movement x task x overlap: $F(1,64) = 6.65, p = .012, \eta^2_p = .09$; familiarity x length x overlap: $F(4,256)$
4.4.2 Combined Movement Advantage

Paired t-tests (2-tailed) were carried out to compare the combined movement advantage scores to chance (0). Overall, there was no significant movement advantage, $M = .01, t(67) = 0.97, p = .34$, but across the two experiments there was a significant movement advantage for unfamiliar faces that were presented as dynamic non-degraded sample images, paired with 1s, 2 s or 4 s PLDs, $ps < .020$.

A mixed ANOVA was carried out to investigate the overall movement advantage, with three within-subjects factors (2 x familiarity; 2 x moving or static sample image; 4 x length of test stimulus), and two between-subjects factors (2 x task; MTS and same/different; 2 x overlap; overlapping and non-overlapping clips). The analysis revealed significant main effects of clip length, $F(3,192) = 8.17, p < .0005$, $\eta_p^2 = .11$, movement of the sample clip, $F(1,64) = 8.45, p = .005$, $\eta_p^2 = .12$, and overlap, $F(1,64) = 8.45, p < .0005$, $\eta_p^2 = .12$. The movement advantage was significantly greater for clips of 1 s, 2 s and 4 s than for clips of 8 s, $ps < .02$, and significantly greater for trials containing moving non-degraded sample clips, $M = 0.375$, than static non-degraded sample clips, $M = -0.116$. The movement advantage was also significantly larger for clips that overlapped (Experiments 1a and 2a) than for clips that did not overlap (Experiments 1b and 2b).

The length by overlap interaction was significant, $F(3,192) = 3.24, p = .023$, $\eta_p^2 = .05$. Pairwise comparisons show that the movement advantage for short clips was only evident in the trials containing overlap between the sample and test images, $ps < .035$, whereas trials with no overlap showed no difference in movement advantage regardless of the clip length, $ps > .9$. There was a similar pattern for the familiarity by length interaction, $F(3,192) = 2.71, p = .046$, $\eta_p^2 = .04$. The pattern of results for unfamiliar faces mirrored the main effect of length (movement advantage for 1 s, 2 s and 4 s clips was significantly greater than for 8 s clips, $ps < .0005$), whereas for famous faces the movement advantage was of equal size for all clip lengths. Finally, the familiarity by overlap interaction was also significant, $F(1,64) = 4.30, p = .042$, $\eta_p^2 = .06$, confirming that unfamiliar faces showed a significantly larger movement advantage in conditions where the movements overlapped ($p =$
.026), whereas there was no effect of familiarity when the sample and test clips showed completely different movement sequences, $p = .569$.

There was no main effect of task, $F(1,64) = 0.001$, $p = .981$, and no interactions between task and any other variables, all $ps > .1$. No three, four or five-way interactions were significant, all $ps > .05$.

4.4.3 Discussion of the Combined Results of Experiments 1 and 2

The results from the combined analysis confirm some of the tentative conclusions from individual experiments. It appears that participants perform equally well and derive equal movement advantages from MTS and same/different tasks, which suggests that any impairment in performance in Experiment 1 is not due to memory load or task demands. The effects of clip length and familiarity on both $d'$ and the movement advantage are only present when participants can compare movements directly – that is, when there is some overlap between the sample and test clip. When there is no movement overlap, participants’ ability to discriminate faces based on their movement drops to (or below) static recognition levels, as was shown in the individual experimental results. However, participants show a consistently greater movement advantage when the non-degraded sample clip is also moving, regardless of clip overlap, clip length, or familiarity.

4.5 General Discussion of Experiments 1 and 2

This chapter presented four experiments investigating the role of movement in familiar and unfamiliar face recognition. The main aim was to compare the effect of movement on famous and unfamiliar face matching, and to test whether participants can match a fully illuminated (non-degraded) video or still image to a PLD. In general, the results indicated that participants were capable of matching both famous and unfamiliar faces to point-light images when there was some overlap between sample and test motion (Experiments 1a and 2a). When there was overlap between the clips, famous faces were matched more accurately than unfamiliar faces overall. However, unfamiliar, but not famous faces, showed a movement advantage – unfamiliar faces were matched better when the PLD was moving than when it was static (again, this effect was isolated to Experiments 1a and 2a). The benefits of movement were not isolated to the PLDs – participants were better at matching
famous and unfamiliar faces if the first, non-degraded image was a video, rather than a static image.

A secondary aim of these experiments was to assess the impact of several methodological factors: the duration of the stimuli, the type of matching task, and the difficulty of the matching task (i.e., whether the movement overlapped or not). Experiments 1a and 2a established that matching performance was better with shorter clips (1 – 4 s) than long clips (8 s). The type of task (sequential match-to-sample or sequential same/different) appeared to have a slight numerical effect on performance, but formal statistical analyses revealed no differences between the two tasks. As mentioned above, task difficulty had a very large impact on matching performance: when the clips did not show overlapping movement, participants were extremely poor, and no movement advantage emerged. In fact, all the effects mentioned above (familiarity, movement, movement of sample clip and clip duration) disappeared when participants were asked to match individuals from different movement sequences. The next sections explore these effects in more detail, and compare the findings to other studies in the area.

4.5.1 Familiarity Effects in Experiments 1 and 2

The main aim of these experiments was to assess matching performance and the presence of a movement advantage for famous and unfamiliar faces. In Experiments 1a and 2a, accuracy and d’ results show that famous faces were matched better overall. Matching short video clips of unfamiliar faces was also above chance in the majority of conditions where there was some overlap in movement between sample and test images. This indicates that familiarity with a face is not a prerequisite to extracting and matching movement information from degraded images. In fact, results from Experiments 2a and the combined analysis show that a larger movement advantage is present when matching movements from unfamiliar faces than when matching movements from famous faces. Even more surprising is the finding that there was no significant movement advantage for famous faces in any of the four experiments reported here. This effect arose because matching performance for static famous face PLDs in Experiments 1a and 2a appeared to reach an asymptote. Movement added no extra beneficial information, possibly because the moving faces were matched for amount and distinctiveness of movement, which provides a relatively stringent test of the movement advantage.
compared to previous studies (e.g., Knight & Johnston, 1997; Lander et al., 1999; Lander et al., 2001). On the other hand, unfamiliar faces were matched poorly from static PLDs (particularly in Experiments 1a and 2a), and therefore the extra shape and/or characteristic movement information in the moving clips aided matching performance. Since the effect of familiarity in the overall accuracy and d’ analysis differs from the effect of familiarity in the movement advantage, the two findings will be discussed separately.

4.5.1.1 Familiarity gives rise to an overall matching advantage. If there was no discernable effect of movement, why were famous faces matched more accurately (Experiment 1a) and discriminated better (Experiment 2a) than unfamiliar faces? As mentioned above, the matching advantage for famous faces appears to arise because famous faces were matched particularly well from static PLDs. It is likely that any movement advantage that could have arisen for famous faces was obscured by the above-chance recognition of static PLD images: across all four experiments, six out of eight conditions with a static PLDs of famous faces were matched better than chance (with the exception of Experiment 1b, static sample image; and Experiment 2b, moving sample image). On the other hand, static PLDs of unfamiliar faces were matched at chance levels for six out of eight conditions (with the exception of Experiments 1b and 2b, static sample images). Collapsing across all four experiments, matching performance for static famous PLDs was significantly better than for static unfamiliar PLDs ($p < .0005$). Alternatively, the effect of familiarity may have arisen due to poor performance with unfamiliar faces. Participants may have been distracted when they saw an unfamiliar face – perhaps trying to identify the person or memorise elements of their face – which could have lowered performance in the unfamiliar trials. However, if unfamiliar faces distracted participants, there is no obvious reason why moving PLDs would be matched better than static PLDs. Therefore, it seems most likely that the familiarity results for Experiments 1a and 2a reflect enhanced processing for famous faces, rather than impaired processing for unfamiliar faces.

An advantage for familiar face matching (regardless of movement) is not unprecedented in the research on movement and face recognition. In one of the few other studies to assess the impact of familiarity on dynamic face recognition, Bruce et al. (2001, Experiment 1) compared matching performance for static, multi-static or
moving images of personally familiar and unfamiliar faces. Like the current set of experiments, Bruce et al. (2001) used a same/different task with one high quality image (a photograph) and one degraded image (from CCTV footage). They found an overall advantage for matching familiar faces, but no movement advantage for either familiar or unfamiliar faces. They suggested that any possible movement advantage for familiar faces may have been eliminated by people being too good at identification of CCTV images, a situation that seems likely given that participants had unlimited access to 3 s looped moving clips and still images concurrently.

The current results lead to the question: why were famous faces in the current study matched so well from static PLDs, when unfamiliar faces were not? Unlike Bruce et al. (2001), participants in the current study did not have unlimited time to view and compare the degraded clips. This leaves three possible reasons for the familiarity advantage: basic stimulus characteristics, the use of static structural cues, and the use of characteristic poses or expressions.

The effect of familiarity may have arisen from basic stimulus characteristics – the same frame was used to create the sample and test images in Experiments 1a and 2a, and some identifying pictorial elements may have been preserved in the PLDs. Furthermore, while effort was made to control the distinctiveness and amount of movement in any given pairing of videos, ratings of distinctiveness or typicality were not collected for static frames. Variability in the distinctiveness of static frames may have artificially increased accuracy in some conditions. However, if stimulus characteristics such as matching frames, obvious shape similarities, or distinctiveness of PLDs accounted for the above-chance matching performance of static frames, it would be expected that these benefits would apply to unfamiliar faces as well as famous faces. As mentioned above, this was not the case, particularly for Experiments 1a and 2a.

It is possible that the discrepancy in static recognition was driven by participants’ familiarity with the shape or configuration of the famous faces. In general, studies of famous and personally familiar faces have found that familiar face recognition is robust to a number of (very dramatic) image transformations. For example, familiar face recognition is possible even when the image is stretched 200% horizontally or vertically (Hole et al., 2002, Experiment 2), and even when an image is both stretched and blurred (Hole et al., 2002, Experiment 4). Hole et al. suggested that familiar face recognition was robust to extreme image degradations
because people are sensitive to configural information – that is, the spacing of features in the face. The ability to use basic shape or configuration information as a cue to facial identity is not isolated to familiar faces – research using static faces has shown that participants can discriminate between experimentally familiar and unfamiliar faces that differ only in shape (Caharel, Jiang, Blanz, & Rossion, 2009; O’Toole, Vetter, & Blanz, 1999; Russell, Sinha, Biederman, & Nederhouser, 2006). However, several studies on experimentally familiar faces have found that familiarity increases sensitivity to small configural variations (spacing changes; Brooks & Kemp, 2007; O’Donnell & Bruce, 2001).

The hypothesis that familiarity with a face makes it easier to discriminate small spacing changes could explain the findings from Experiments 1a and 2a. In the current set of experiments, the PLDs preserved some shape information – for example, the width and height of the face, and the relative position of the eyes, nose and mouth on the head (i.e. some basic configural cues). It is possible that participants performed better in the static famous PLD condition than the static unfamiliar PLD condition because they were more sensitive to small changes to the configuration of the facial features in famous face PLDs.

Another explanation for the effect of familiarity is that participants were able to match the characteristic head poses or expressions on the static famous faces. The robustness of familiar face recognition has lead many authors to conclude that familiar faces are represented in a flexible, view-invariant manner – that is, our representation of a familiar face incorporates all of our encounters with the person; the consistent elements of their appearance; and possible variations of their face, such as viewpoint and expression changes (Bruce, 1994; Burton et al., 2011; Jenkins & Burton, 2011). There is some evidence that our representations of familiar faces also capture typical, or characteristic, static cues such as expressions. Famous faces are easier to recognise if they are displaying a smile than a neutral or negative expression, possibly because a smiling expression is more typical of our encounters with famous faces (Endo et al., 1992; Gallegos & Tranel, 2005; Kaufmann & Schweinberger, 2004). Our recognition of famous faces can also be influenced by an “iconic” image, or a typical image we have viewed in the past (Carbon, 2008). Based on these findings, it is possible that participants in the current study could match the famous PLDs, even in static form, because they displayed characteristic head poses or expressions (e.g. tilting/inclining the head; lop-sided smiles). This is not to
suggest that participants in the current experiments were only using pose and expression information to match the static PLDs – it is likely that the advantage for static famous faces arose from a combination of enhanced sensitivity to shape and configural information, and the presence of predictable characteristic face poses or expressions. Perhaps seeing the non-degraded sample image made it easier for participants to extract and compare the relevant shape and pose information in the PLD. As mentioned in the introduction to this chapter, the effect of characteristic head and/or face poses on static recognition of familiar faces may have been minimized in previous studies that chose neutral facial images as their static control condition (Knight & Johnston, 1997; Lander et al., 1999).

Experiments 1a and 2a are the first, to the authors’ knowledge, to find an advantage for matching static familiar PLDs. It is unusual for participants to perform above chance levels when presented with static PLDs, even those of highly familiar people – Rosenblum et al. (2007) and Bruce and Valentine (1988) both found chance-level performance when participants were asked to identify friends from static PLDs. However, Rosenblum et al. (2007) and Bruce and Valentine used a naming task, not a matching task, and their static frames were less likely to contain idiosyncratic head movements (Rosenblum et al. (2007) used non-rigid motions; Bruce and Valentine used expressions as their static frames). Furthermore, both Rosenblum et al. (2007) and Bruce and Valentine scattered points randomly on the face, which may have obscured the shape of the face and the location of features such as the eyes and mouth.

Overall, the advantage for matching famous faces in Experiments 1a and 2a can be interpreted as a reflection of our generally robust processing for static familiar faces, and a consequence of the particular stimuli and methodology that was used in this Experiment. However, the presence of a movement advantage for unfamiliar faces demonstrates that participants were doing more than just matching static face shapes or configurations; in some cases, movement can help face matching.

4.5.1.2 Unfamiliar faces show a movement advantage. The presence of a movement advantage for unfamiliar, but not famous faces is particularly noteworthy because previous studies have often failed to find a movement advantage for unfamiliar faces (Bruce & Valentine, 1988, Experiment 1; Bruce et al., 2001; Christie & Bruce, 1998), while movement advantages have often been reported for
famous and personally familiar faces (Bruce & Valentine, 1988, Experiment 2; Knight & Johnston, 1997; Lander et al., 1999). Given the significant advantage for famous faces in the accuracy and d’ analyses, it is possible that famous face matching reached an asymptote from static PLDs (as discussed above), whereas unfamiliar faces were matched poorly from static PLDs, and had room to benefit from movement-based cues. It is important to note, however, that these conclusions are limited to Experiments 1a and 2a – that is, the experiments with overlap between the sample and test clips. When there was no overlap, matching was close to chance levels regardless of familiarity.

The benefit of movement for unfamiliar face matching can be explained in three possible ways. First, participants may have used the extra static frames contained in the moving clips as a source of extra static information. This explanation seems unlikely to account for the entire movement advantage – several studies (using familiar, unfamiliar, and experimentally familiar faces) have tested the effect of multiple static images compared to moving clips, and found an additional benefit for movement that is unrelated to the extra static information (e.g., Lander & Bruce, 2000; Pike et al., 1997; Rosenblum et al., 2002). Furthermore, the 8 s clips, which provided the most static frames, showed a significantly smaller movement advantage than the 1 s, 2 s, and 4 s clips in Experiment 1a, and the 2 s clips in Experiment 2a. These findings strongly suggest that the movement advantage is not just a result of the extra static information contained in moving clips.

Second, participants may have matched faces based on structure-from-motion cues. This possibility was tested in the static sample image condition: if participants could match static sample images to moving PLDs, that would suggest that they were extracting structure-from-motion cues from the PLDs and matching these structural cues to the static sample image. However, there was no significant movement advantage for trials with static sample clips, for either famous or unfamiliar faces. The results in the static sample condition may explain why so many other studies have failed to find a movement advantage for experimentally familiar or unfamiliar faces. In many studies of unfamiliar and experimentally familiar faces, participants have only been asked to match moving images to static images (e.g., Bruce et al., 2001; Bruce & Valentine, 1988, Experiment 1; Lander & Bruce, 2003). This method only examines whether movement improves our static representation of a face, and does not test whether unfamiliar faces can be matched based on characteristic motion
patterns. The current experimental findings, along with the weight of evidence from previous studies, suggest that the role of structure-from-motion in movement-based face processing is limited. It is possible that structure-from-motion only plays a significant role in identification when participants view large, rigid rotational movements of the head (e.g., Farivar et al., 2009; Pike et al., 1997), and when there is no characteristic movement information to match.

Given the lack of movement advantage for static sample images, the movement advantage for unfamiliar faces in this study is likely to have arisen from matching characteristic motion patterns (e.g., Hill & Johnston, 2001; Lander & Davies, 2007), or the effect of social cues on face processing (Pilz et al., 2006, 2009). However, if participants were capable of extracting and comparing characteristic motion patterns from different movement sequences, or if the effect of movement arose because of extra attention to moving clips, there should have been a movement advantage even when there was no overlap between the sample and test clip (Experiments 1b and 2b). Since there was no movement advantage in Experiments 1b and 2b; and no movement advantage for static sample clips, it is likely that the results from Experiments 1a and 2a are based on perceptual matching, rather than facial identification. This conclusion does not imply that perceptual matching is the only possible basis of a movement advantage for unfamiliar faces. It is possible that the movement advantage for unfamiliar faces in previous studies arose because movement facilitated structure-from-motion processes and allowed participants to extract and compare structural information or characteristic motion patterns. Certainly, evidence from Hill and Johnston (2001), Pilz et al. (2006), and Farivar et al. (2009) suggests that a movement advantage for unfamiliar and experimentally familiar faces does not require overlapping movement sequences. However, without evidence of a movement advantage for static sample images or non-overlapping clips, it is impossible to determine whether these processes occurred during the unfamiliar face-matching task in Experiments 1a and 2a.

Although the movement advantage for unfamiliar faces is likely to be based on movement matching, not structure-from-motion or characteristic movement patterns, it is interesting to compare the findings from Experiments 1a and 2a to other studies of moving faces. Christie and Bruce (1998) failed to find a movement advantage for experimentally familiar faces, even when movement was present during both learning and test phases, and despite identical movements being present in study and
test conditions for half of the target faces. One possible reason for the discrepancy between their results and the current study is the length of time between their learning and test sessions: participants were tested around 30 minutes after the learning phase, which may have eroded any memory for individual movement characteristics, and therefore eliminated a movement advantage. Furthermore, the Christie and Bruce study familiarised and tested participants with staged movements, rather than candid speech. This type of stimulus might have reduced variability or idiosyncrasy in the movement, and limited the ability of participants to use individual movements as identity cues.

Bruce et al. (2001) also failed to find a movement advantage for unfamiliar faces. Once again, this maybe due to the specific methodology they employed – participants in Bruce et al.’s (2001) study were allowed unlimited viewing time for static clips. Consequently, static unfamiliar faces in Bruce et al.’s (2001) experiment were identified at levels well above chance, which may have eroded the movement advantage in their study. Bruce et al.’s findings lead to the conclusion that the movement advantage for unfamiliar faces may only be apparent for slightly more difficult tasks (e.g., cross-modal matching, Kamachi et al., 2003); tasks with time constraints (e.g., Thornton & Kourtzi, 2002); or degraded stimuli (e.g., Rosenblum et al., 2002). This is in line with research on famous faces, which has only found a movement advantage when static recognition is degraded (Knight & Johnston, 1997). Unfortunately, increasing the task difficulty too much can result in participants failing to match clips at all, as seen in Experiments 1b and 2b (and possibly Christie & Bruce, 1998). Further research is necessary to determine what task, stimuli, and procedures result in a movement advantage without the presence of floor or ceiling effects (see Chapters 5 and 6).

4.5.1.4 Conclusions on familiarity. Overall, Experiments 1 and 2 demonstrated that familiarity does not lead to a greater movement advantage, as Roark et al.’s (2003) hypotheses would seem to predict. When the task and stimuli are equated, unfamiliar faces show a greater movement advantage than famous faces. It is possible that familiar faces are simply able to call on extra static cues (for example, head pose or fine structural information), rather than having to resort to strategies based on structure-from-motion or characteristic movement patterns. Unfamiliar faces appear to benefit from movement, although in this study the
movement advantage was probably based on simple movement matching strategies. Our results, and the results from previous studies of unfamiliar and experimentally familiar faces, suggest that the movement advantage for unfamiliar faces is likely to be based on characteristic movement patterns, with a comparatively small contribution of structure-from-motion information. However, further work that equates static recognition of familiar and unfamiliar faces is necessary to confirm whether the movement advantage for unfamiliar faces is consistent.

4.5.2 The Other Movement Advantage: Moving Sample Clips

The majority of studies researching movement and face recognition have tested the effect of movement by varying whether one clip was moving or not. For example, Knight and Johnston (1997) presented one moving or static clip of each famous face, and Pike et al. (1997) tested whether presenting a moving face during the learning phase would help subsequent static recognition. In this experiment, however, the effect of movement was examined in two ways – first, the effect of movement in the degraded clip (as discussed above), and second, the effect of movement of the non-degraded sample clip. As discussed above, this allowed the current results to be compared with previous studies using different methodologies, and tested whether participants were using structure-from-motion cues during the matching tasks.

Unlike the effect of familiarity, the movement of the sample clip had a strong influence in the accuracy, d’, and movement advantage analyses. Throughout Experiments 1a and 2a participants were, on average, better at discriminating the identity of a PLD when they were shown a moving non-degraded sample image than a static frame. Furthermore, the movement advantage in the combined analysis was larger in trials that had moving sample images compared to those with static sample frames. In fact, there was no significant movement advantage in any experiment for trials with static sample images, whereas moving sample images resulted in significant movement advantages in Experiments 1a and 2a. This result leads to the conclusion, as discussed above, that participants were using dynamic characteristics (probably movement-matching) rather than structure-from-motion when matching unfamiliar faces. Like moving PLDs, it is possible that moving sample images were beneficial because they provided more static information than a single static frame. However, as discussed above, previous studies using multiple static images have
found a movement advantage (e.g., Lander & Bruce, 2000; Rosenblum et al., 2002), suggesting that there is some role for structure-from-motion or characteristic movement patterns over and above the extra static cues carried in moving images.

It is surprising that the movement advantage related to seeing the sample face in motion was equal for famous and unfamiliar faces. Theoretically, familiar faces should be matched relatively well regardless of whether the sample face is moving or static, because static faces can be matched to moving PLDs using structure-from-motion information, and static frames should also allow access to stored characteristic motion patterns for that individual. On the other hand, static images of unfamiliar faces would only facilitate structure-from-motion based comparisons, as a person would not have any stored representations of motion for an unfamiliar face. Therefore adding movement to the sample image (and adding another dimension of comparison) should increase the movement advantage for unfamiliar faces. The fact that there was no significant movement by familiarity interaction in the movement advantage analyses suggests that participants were using the same motion-based cues regardless of whether the face was familiar to them or not. Given the movement advantage disappeared in the more difficult, non-overlapping task, it is probable that those cues were primarily movement matching, with little, if any, contribution from structure-from-motion.

The movement advantage for the sample clips might also be explained by the social signals hypothesis (Roark et al., 2003). Presenting the initial clip in motion may have prompted participants to attend to the face more than in a static trial, since the moving clip contained more social cues such as speech information, changes in expression, and direction of eye gaze. The idea that participants paid more attention to the moving clips than the static frames is consistent with the finding that trials with moving sample clips were matched more accurately in the overall accuracy and $d'$ analysis (i.e., when performance with static PLDs was taken into account), and also resulted in a significantly greater movement advantage than static sample images. Once again, though, the fact that Experiments 1b and 2b found no significant effect of movement for the sample clip suggests that movement matching was the primary source of the movement advantage; the contribution of social cues and enhanced attention to the movement advantage was, at most, small.

Overall, it appears that movement is an important factor in both the sample and test clip. Movement in the sample clip facilitates matching for both familiar and
unfamiliar faces, which is somewhat unexpected. Once again, the results from Experiments 1 and 2 support the idea that participants performed better when matching moving than static clips because they were matching individual movement sequences (as opposed to using structure-from-motion cues). However, the results do not rule out an attention-based contribution – the social signals in moving clips may have encouraged participants to pay more attention to the face structure or the way a face was moving.

4.5.3 Bias, Hits and Correct Rejections in Experiments 1 and 2

The results of Experiment 2 suggest that the familiarity of a face and the presence of movement during a matching task have a significant effect on the way participants approach the task. Participants in Experiment 2a adopted more conservative criteria when matching famous than unfamiliar faces: overall, participants displayed an bias towards responding “same”, and this bias increased when participants viewed unfamiliar faces. The results of Experiment 2a are consistent with Bruce et al.’s (2001, Experiment 1) analyses of bias. In trials that showed unfamiliar faces, Bruce et al.’s (2001) participants tended to be relatively liberal in their responses; in trials that included a photograph of a familiar face, participants were more cautious.

Interestingly, in Experiment 2a, the hit rates were statistically equal for famous and unfamiliar faces, but there were significantly more correct rejections for famous than unfamiliar faces. This suggests that the overall discriminability benefit for familiar faces is not based on better overall performance, or even an improved ability to match two of the same faces, but instead is derived from an increased ability to correctly reject items that do not match. Participants in Experiment 2a may have adopted a liberal strategy for unfamiliar faces, erring towards responding “same” unless they were certain that they were different. It is possible that participants displayed a movement advantage for unfamiliar faces because movement of the test PLD provided better “different” cues than static images (perhaps the non-matching

---

7 Note that this result is slightly different to Bruce et al.’s (2001) overall finding that familiarity did not have an effect on bias. This discrepancy is due to the multiple trial types included in their analysis – the effect of familiarity reported here was only apparent in a trial type by familiarity interaction.
movement helped participants reject mismatched clips); whereas famous faces could be discriminated, and incorrect matches rejected, based on the information contained in static PLDs. The fact that moving PLDs elicited significantly more correct rejections than static PLDs on average (but a similar proportion of hits) supports the idea that movement of the test clip was helpful when making “different” decisions.

Once again, the effects of familiarity were not replicated in Experiment 2b: there was no significant difference in bias between famous and unfamiliar faces. This could be because participants were highly uncertain in both the famous and unfamiliar face conditions, which could have eliminated any systematic responding strategy. Alternatively, the overall task difficulty may have encouraged participants to take a more conservative approach to responding – this could have resulted in minimal bias by eliminating the overall predisposition to respond “same”.

Analyses of bias also show that movement of the sample clip can significantly affect how participants approach the matching task. In Experiment 2a and 2b, moving sample clips resulted in more bias towards “same” responses than static sample images. Participants may have been more liberal in their responding when they had to match moving clips for the same reason they were more liberal towards unfamiliar than famous faces – trouble definitively rejecting non-matching images, resulting in uncertainty and a generous response strategy. Alternatively, participants may have been basing their decisions on perceptual similarities. Static sample images may have seemed less similar than moving sample images to the moving PLDs, resulting in a greater tendency to respond “different”. An interaction between movement and clip length supports this explanation: the increased “same” bias for moving sample clips was strongest when they were paired with short (1 s – 2 s) PLDs; conversely, static sample images paired with static PLDs also invoked an increased “same” bias. In other words, short sample clips resemble short PLDs, and static images resemble static PLDs, so participants focused on these superficial cues when performing the matching task.

Notably, the effect of moving sample images and the interaction between moving sample images and clip length are the only significant effects that appear in both Experiments 2a and 2b. The fact that the bias was present even when participants could not complete the task strongly supports the argument that these bias effects arose from perceptual similarity of the stimuli, since this does not rely on participants being able to complete the task at above chance levels.
It appears that there were two separate influences on bias overall: familiarity and movement of the PLD; and movement of the sample clip. Overall, familiarity and the movement of the PLD clip (both of which were expected to help matching performance) had a similar effect on participants’ responses – moving PLDs did not significantly change participants’ response bias, but both familiarity with a face and moving PLDs resulted in more correct rejections. On the other hand, providing a short moving sample clip (which significantly improved accuracy and discrimination performance) increased bias towards a “same” response, possibly because participants were influenced by the perceptual similarity between the sample and test clips. The effect of familiarity, movement and methodological factors on bias will be explored further in Chapter 5.

4.5.4 Methodological Factors in Experiments 1 and 2

The primary aim of Experiments 1 and 2 was to investigate the role of familiarity in movement-based face recognition, but they were also designed to address some methodological questions. First, these experiments acted as a large-scale test to determine whether PLDs could be used in matching tasks with famous and unfamiliar faces. Second, these studies assessed the impact of varying task and task difficulty: while previous studies have used varying tasks and task difficulty, they have rarely compared them directly or attempted to quantify their impact on the movement advantage. Finally, these experiments investigated whether the length of the moving stimuli also contributed to the presence (or, in some previous studies, the absence) of a movement advantage. Below we discuss some of the results in relation to these methodological questions.

4.5.4.1 The effectiveness of point-light-displays as stimuli in identification tasks.

One aim of this series of experiments was to examine whether the movements preserved in point-light-displays were sufficient for movement-based face matching using famous and unfamiliar faces. This is especially important since no other studies have created PLDs using famous faces, and it is unclear whether the effects of movement that have been found using PLDs of personally familiar faces (e.g., Bruce & Valentine, 1988; Rosenblum et al., 2007) also apply to famous faces. Accuracy and d’ analyses of Experiments 1a and 2a showed that participants were generally performing well above chance levels. This demonstrates that participants
were quite good at extracting and identifying the matching movement information. This trend was particularly strong for clips between 2 s and 4 s in length. Experiments 1a and 2a, combined with previous research on gender and identity (Hill et al., 2003; Rosenblum et al., 2002, 2007), suggest that facial PLDs contain sufficient information to extract and match individual facial movements that have been encoded as non-degraded images. Based on these results, it seems unlikely that the null results in Experiments 1b and 2b were the result of participants’ inability to extract appropriate movement information from the PLD stimuli.

4.5.4.2 Task and exposure duration – an effect of memory? A second methodological aim of this study was to compare results from different face recognition tasks, and to determine the optimum length of clip for studying any movement advantage that may arise in face recognition. Experiment 1 used a match-to-sample paradigm, whereas Experiment 2 used a same/different task. In all experiments, clips of one to eight seconds long were tested. The overall pattern of results in match-to-sample and same/different tasks was quite similar – a combined analysis of all four experiments indicated that task made no overall difference to discriminability scores, and did not significantly interact with any variables. Despite the fact that overall pattern of results was similar across the tasks, general performance (i.e. number of conditions above chance) was better in the same/different paradigm, and a movement advantage was present in more conditions. This result may reflect memory constraints for complex facial movement patterns. In the match-to-sample experiments, participants had to remember and compare a 2 s non-degraded video or still image to two PLDs, which ranged from 1 to 8 s in length. This resulted in a minimum of 4 s of movement (in the one second PLD trials), up to a maximum of 18 s of complex dynamic patterns to be stored and compared over a short period of time. In contrast, movement in the same/different trials ranged from 3 s to 10 s long, and only one comparison was necessary. Given the inherent difficulty of dynamic face recognition (especially for degraded images of unfamiliar faces), it is not surprising that participants showed a more consistent movement advantage in shorter trials. These conclusions are further supported by the pattern of results across different lengths of clip for Experiments 1a and 2a. In the combined analysis, the movement advantage in the easier, overlapping tasks peaked in the 1 s, 2 s and 4 s long conditions, and dropped dramatically for 8 s long clips.
Our results support the theory that the ability to match facial movements is driven by short-term dynamic representations that decay quickly (Thornton & Kourtzi, 2002), and the 8-second long clips were simply too long to be remembered. Alternatively, it is possible that participants extract structure-from-motion information rapidly from the moving clips (i.e., within the first 2 s), and extra movement information beyond this point “distracts” them from the identification task. However, the results from static sample trials suggest that participants were not using structure-from-motion to match clips (see section 4.5.2 and 4.5.3). Furthermore, the lack of significant results in Experiments 1b and 2b suggests that shorter clips (e.g., 1 s, 2 s, and 4 s clips) are only helpful in easier tasks, which facilitate direct comparisons between movement in the sample and test clip. When participants were faced with extremely difficult conditions (i.e., no overlapping clips), performance on all clip lengths was equally poor. This further supports the argument that a structure-from-motion explanation (which requires no clip overlap) is less likely than a memory-based explanation of the effects of clip duration. No studies have tested the effect of clip length in recognition (e.g., old/new or naming-based) tasks, but the current results suggest that shorter clips (1-4 s) are more likely to give rise to a movement advantage in matching tasks, probably due to memory constraints for dynamic stimuli.

4.5.4.3 Generalising across motion patterns – The effects of task difficulty.

The results of Experiments 1a and 2a suggest that the use of PLDs and matching tasks may lead to new and interesting comparisons between familiar and unfamiliar face matching, and contribute to our understanding of how movement helps (or in some cases does not help) performance in face matching tasks. However, the results of Experiments 1b and 2b indicate that the conclusions from Experiments 1a and 2a (like the general role of movement in face recognition) are somewhat nebulous, and can be easily disrupted by increasing task difficulty. In Experiments 1a and 2a, participants could complete the task by searching for an overlapping movement sequence in the sample and tests clips – this is akin to basic picture-matching tasks in static face recognition, and only provides a limited amount of information about how people use movement as a cue to identity in situations where they see a variety of facial poses and movements. However, in Experiments 1b and 2b participants were asked to generalise from one movement sequence to another produced by the same
individual – a more realistic and ecologically sound task. The results from Experiments 1b and 2b suggest that this task is far more difficult than simple movement matching. Overall, participants struggled to perform above chance levels, they exhibited no movement advantages, and many reported finding the task near impossible and resorting to guessing. This is unusual, because many previous studies have reported a clear movement advantage for famous and unfamiliar faces (when tested separately – see Tables 1, 3 and 4 in Chapter 3), even when matching tasks used non-overlapping clips (e.g., Rosenblum et al., 2002). This leads to the question: why was there no movement advantage in Experiments 1b and 2b? There are several possible explanations.

First, it is possible that participants in Experiments 1b and 2b were attempting to rely on basic movement cues that were not identity-specific – for example, the amount or distinctiveness of movement produced by the individual in the sample clip. Previous research has indicated that more distinctive movements give rise to a larger movement advantage (Lander & Chuang, 2005), but most studies of famous faces in the past have not controlled for distinctiveness and amount of movement (e.g., Knight & Johnston, 1997; Lander & Bruce, 2000; Lander et al., 2001; Lander et al., 1999). It is difficult to assess whether the amount someone moves or the distinctiveness of that movement were used as default identity markers in previous studies that have used television-based movement samples (i.e., “that person moves their head a lot, they must be person X”). Since both amount and distinctiveness of movement were matched in this study, participants may have found themselves unable to use the most obvious and readily available movement cues, and therefore unable to perform the task. If participants in past experiments have been using these cues, the fact that they were matched here may also explain why the current experiments did not find a movement advantage for famous faces.

Second, it is possible that participants are unable to identify characteristic motion patterns in such a short time frame. However, a number of papers have demonstrated a motion advantage from short degraded clips in famous-face identification tasks (e.g. Knight & Johnston, 1997), and unfamiliar face matching tasks (e.g., Thornton & Kourtzi, 2002), so it seems more probable that the task in the current experiments was simply more difficult than the tasks in previous studies. This leads to the third possibility – perhaps the task was difficult because participants had to generalise across movement and format. In other words, it is possible that
participants had trouble extracting and matching individual movement patterns whilst also trying to generalise motion from a non-degraded sample image to a PLD test image. Bruce et al. (2001) and Burton, Wilson et al. (1999) showed that identification was difficult when generalising from a CCTV clip to a high quality photograph, so it seems reasonable to assume that generalising from a high-quality video or still to a PLD would be even more difficult. The results from Experiment 1a and 2a support this idea: even when the video and PLD clips were identical (as in the 2 second condition in Experiments 1a and 2a), participants were unable to match video clips with perfect accuracy, which suggests that the format-switch does have a detrimental effect on accuracy in general. However, previous research in the field has not tested the effect of generalising to different formats – it is unclear whether participants would have performed better if asked to generalise to low quality or degraded videos (e.g., Burton, Wilson et al., 1999; Lander & Davies, 2007), or whether PLDs are particularly difficult. Consequently, Chapter 5 investigated the effect of different formats (PLDs and shape-normalised avatars), and whether changing or maintaining formats across sample and test images has an impact on matching performance.

Despite these drawbacks and generally poor performance when generalising across motion, it appears that participants approached the more difficult task in a very similar manner to the easier, movement-matching task. Bias results from Experiments 2a and 2b were very similar, with the exception of the main effect of familiarity. Both experiments showed more liberal response patterns for moving clips, and a tendency to respond “same” more often for perceptually similar clips (e.g., trials with two short moving clips or two static images).

Although participants showed similar response patterns to the easy and hard tasks, the results of Experiments 1 and 2 suggest that task difficulty is an important factor to consider when investigating a movement advantage in face recognition. Although participants are adept at matching overlapping movement sequences, generalizing from one movement sequence to another is a difficult task, even for highly familiar faces.

4.5.5 Conclusions from Experiments 1 and 2

Overall, this study demonstrates that movement can improve matching performance for degraded images of faces. Consistent with previous literature (e.g.,
Bruce et al., 2001), familiar faces – in this case famous faces – were matched significantly better than unfamiliar faces. However, the movement advantage in both experiments was larger for unfamiliar faces than for famous faces. While previous research suggests that familiar face recognition may also benefit from movement, none of the four experiments reported here support this hypothesis, possibly due to above-chance static recognition performance for famous faces. The movement advantage is small, and can easily be disrupted by varying the motion sequences between clips. The results suggest that participants were relying primarily on motion matching (a basic version of the supplemental information hypothesis, Roark et al., 2003), rather than extracting characteristic patterns of motion that could be identified in different movement sequences, or using structure-from-motion processes to develop three-dimensional representations of the head and face. This theory is also supported by the fact that matching performance was significantly better when both the sample and test image showed the face in motion. The results from Experiments 1a and 2a suggest that motion patterns can be extracted and compared from short (2-4 s) clips more effectively than from longer (8 s) clips, which may support the theory that the basis of a motion advantage is short-term dynamic representations. This research also suggests that familiarity with a face and movement of sample and test clips both contribute to participant bias. It is possible that these effects are distinct and separable, however further research will be needed to confirm that these bias effects are constant and exist independently of each other. The effects of familiarity and movement on face matching will be explored further in the next chapter, which examines the effect of different stimulus manipulations on the movement advantage.
Chapter 5

The Role of Movement in Famous and Unfamiliar Face Matching

The Effect of Stimulus Type
CHAPTER 5: THE ROLE OF MOVEMENT IN FAMOUS AND UNFAMILIAR FACE MATCHING: THE EFFECT OF STIMULUS TYPE

The experiments in Chapter 4 confirmed that movement can improve face matching under some circumstances. Most notably, Experiments 1 and 2 established that in match-to-sample and same/different matching tasks with overlapping movement sequences, participants showed a movement advantage for unfamiliar, but not famous faces. However, the overall pattern of results showed the opposite pattern – when static matching performance was taken into account, famous faces were matched more accurately and discriminated better than unfamiliar faces. One possible explanation for these results is that participants were able to match famous faces to their corresponding PLD because the participants were highly familiar with the shape cues contained in the PLD stimuli. Participants may have used these shape cues, rather than the movement information, to match the famous faces.

This chapter presents two experiments that build on the findings of Chapter 4. Like Experiments 1 and 2, Experiments 3 and 4 were designed to test the effect of familiarity in movement-based face matching, and to address some methodological questions. Specifically, Experiments 3 and 4 compared the effect of movement on famous and unfamiliar face matching when the faces were presented as PLDs (as in Experiments 1 and 2) and shape-normalised avatars, which are created by projecting individual motion onto a uniform face and head structure.

There were multiple reasons to compare matching performance using different stimuli. The results of Experiments 1 and 2 raised the question: were participants unable to match faces with different movement sequences because they could not extract and generalise characteristic movement patterns; because they could not generalise across different presentation formats in general (i.e., from non-degraded images to degraded images); or because they could not generalise to PLDs specifically? Experiments 3 and 4 addressed these questions in two ways. First, by comparing matching performance from non-degraded images to PLDs and avatars (Experiment 3), to test whether the problems generalising across movement sequences extended to degraded stimuli other than PLDs (particularly stimuli that do not contain individual shape or configural information); and second, by comparing
matching performance from PLD to PLD or avatar to avatar, to determine whether matching performance was better when participants did not have to generalise across formats at all (Experiment 4).

Another reason to compare PLDs and shape-normalised avatars is methodological, and extends on the work presented in Chapter 4. As discussed in Chapter 3, not all studies investigating the role of movement in face recognition have found a movement advantage, and the role of familiarity in movement-based face recognition and matching is unclear. It is possible that studies have found a consistent movement advantage for famous and personally familiar, but not experimentally familiar or unfamiliar faces because we have good internal representations of famous and personally familiar faces – we have seen them from different angles, with different expressions, and we have had enough exposure to become familiar with the structure of the face (Burton et al., 2011) and the idiosyncratic way the person moves their face and head (O’Toole et al., 2002). Alternatively, it is possible that the different methodologies used in famous, personally familiar, unfamiliar, and experimentally familiar face studies have resulted in more consistent movement advantages for familiar than unfamiliar faces. For example, many famous face studies have used samples of natural movement (e.g., Knight & Johnston, 1997; Lander & Bruce, 2000, Lander et al., 2001; see Table 1 in Chapter 3 for more detail), whereas many studies of unfamiliar or experimentally familiar faces have used scripted movements (e.g., Bruce & Valentine, 1988; Pike et al., 1997; Thornton & Kourtzi, 2002; cf. Hill & Johnston, 2001; see Tables 2 and 3 in Chapter 3 for more detail). In an effort to understand the effects of methodology on the movement advantage, Chapter 4 examined the effect of varying the task, task difficulty, and clip duration on famous and unfamiliar face matching. The experiments presented in this chapter examined whether the conflicting results from past studies arose because of the type of stimulus used.

Many studies have used degraded face images to study motion and face recognition (e.g., Hill & Johnston, 2001; Knight & Johnston, 1997), but the exact type of stimulus (i.e., the way the face is degraded) has varied between the studies. Very little research has been carried out to investigate whether the stimulus type affects the movement advantage. Knight and Johnston (1997) and Lander et al. (1999, Experiment 1) found that a movement advantage for famous faces was present when the face was negated (i.e. reversing the brightness contrast in an image,
see Figure 4 in Chapter 2), but not when the face was non-degraded. The authors put this down to the fact that static recognition in the non-degraded condition was at ceiling levels. Lander et al. (1999, Experiments 2 - 4) expanded on this finding, and confirmed that the movement advantage was also present for thresholded images of famous faces. Unfortunately, due to methodological differences between the experiments it is difficult to tell whether negation and thresholding have a comparable effect. Lander et al. (2001) also investigated the effect of multiple stimulus manipulations – pixelation and blurring – and whether movement was more helpful when the images were more severely degraded. They found that the movement advantage for famous faces was a similar size, regardless of the level of pixelation (they used 10 and 20 pixels-per-face), but the movement advantage was larger at higher levels of blur (i.e., for more blurry faces). Lander et al.’s (2001) results confirm that movement can be particularly helpful when famous faces are presented under non-optimal conditions, and provide the first evidence that different types of stimulus manipulation may elicit different effects of movement. Once again, though, methodological differences make it difficult to directly compare the effects of the different stimulus types.

At this point, no studies have compared the effect of stimulus type on the movement advantage for unfamiliar faces. Consequently, Experiments 3 and 4 are the first studies designed to directly compare the movement advantage for famous and unfamiliar faces with two different types of stimulus manipulation.

The final and most important reason for choosing to compare PLDs and shape-normalised avatars is theoretical. Previous studies that have directly compared the movement advantage for familiar and unfamiliar faces have used personally or experimentally familiar faces (no comparisons have been made with moving famous faces), and have focused on whether these types of familiar faces show a larger movement advantage than unfamiliar faces – in other words, is there a quantitative difference between the movement advantage for familiar and unfamiliar faces. These studies have found mixed results – Bruce et al., (2001) found no movement advantage for familiar or unfamiliar faces; several studies found no interaction between face familiarity and the movement advantage (Lander & Chuang, 2005; Lander & Davies, 2007; Lander et al., 2006); and two studies found that increasing exposure to a face increases the effect of movement (Bonner et al., 2003; Roark et al., 2006). The experiments presented in Chapter 4 found different results again: in
matching tasks with overlapping movement sequences, unfamiliar faces showed a movement advantage, whereas famous faces did not.

The two experiments presented in this chapter compare the size of the movement advantage for famous and unfamiliar faces (as in Chapter 4), but they also examine whether we use the same types of movement-based information when matching familiar and unfamiliar faces – in other words, is there a qualitative difference between the movement advantage for familiar and unfamiliar faces.

The PLDs used in Experiments 3 and 4 were identical to those in Experiments 1 and 2: they were created by tracking multiple points on the face, and then editing the video so only the movement of the points is visible. These PLDs preserve the majority of facial movement information (for example, the movement of the mouth, eyebrows, cheeks and head), and cues to the shape of the head and face. Therefore, any movement advantage from PLDs could be a result of participants matching structure-from-motion, characteristic motion patterns, or both.

Shape-normalised avatars are created by tracking motion from multiple face areas and projecting it onto one standard face form – for example, an “average” face (e.g., Hill & Johnston, 2001). Like PLDs, this method preserves an individual’s movement information, but unlike PLDs, the stimuli contain uniform structure, texture, and facial features. Since the face shape and configuration is normalised, there is only minimal structure-from-motion information available in the avatar stimuli. Consequently, any movement advantage with avatars is likely to arise because participants can match characteristic motion patterns. Comparing performance with PLDs and avatars should reveal whether participants match faces based on characteristic motion patterns only (equal performance for both stimuli); structure-from-motion only (movement advantage for PLDs, but not avatars); or a combination of both (unequal movement advantage for PLDs and avatars).

A review of previous research suggests that characteristic motion patterns contribute to face recognition for familiar faces. Disrupting the speed or rhythm of motion reduces the movement advantage for famous faces (Lander & Bruce, 2000, 2004; Lander et al., 1999), and personally familiar faces (Cook, Johnston, & Heyes, 2012; Lander et al., 2006). Interestingly, slowed, speeded or staggered movements still convey a small movement advantage for famous faces (Lander & Bruce, 2000; Lander et al., 1999), which suggests that there may also be a role for structure-from-motion cues in famous face recognition. However, no studies to this point have used
shape-normalised stimuli to investigate famous face matching, so it is unclear whether some of the movement advantage for slowed, speeded or staggered clips may have resulted from residual characteristic motion pattern information.

Previous studies have also found evidence for the use of characteristic motion patterns in experimentally familiar face recognition and unfamiliar face matching. Lander and Davies (2007) found that experimentally familiar faces only showed a movement advantage when movement was present in both the learning and test phases. This finding mirrors the results presented in Chapter 4. Both studies suggest that participants were learning or matching the patterns of facial movement (although in Experiments 1 and 2, participants were likely matching individual movement sequences, not characteristic movement patterns), as opposed to extracting structure-from-motion information from the moving face and comparing it to the structure of the face in the static image. Moving unfamiliar faces can also be matched to audio tokens (Kamachi et al., 2003; Lander et al., 2007), which suggests that there is some identifying information common to both visual and auditory modalities – most probably timing or rhythm cues.

Some of the most compelling evidence for the use of characteristic motion patterns in unfamiliar face recognition comes from other studies using shape-normalised avatars. Hill and Johnston (2001) projected the movement from twelve different actors onto one average-shaped head. Participants were able to sort the stimuli into groups based on identity and complete an odd-one-out task at above chance-levels, despite the fact that there was no discriminating shape or texture information, and the participants were completely unfamiliar with the actors whose movements they were viewing. Similar results have been obtained when testing infants. Despite the uniform shape and texture, 4-8 month-old infants could discriminate between different actors, demonstrated by a preference for a novel actor following habituation (Spencer et al., 2006). Hill and Johnston’s and Spencer et al.’s results suggest that characteristic face and head movement alone is sufficient to perform matching tasks, even when structural cues are limited. However, it is important to note that neither study included a static control condition. Hill and Johnson presented backwards and inverted motion in an odd-one-out task, and found that even inverted animations were discriminated at above-chance levels (backwards and forwards motion were discriminated equally well). This suggests that participants may have been able to use some cues that could be discriminated even
when the head was upside-down – for example, static poses, rhythm or speed information, or movement patterns of individual features (the inverted and backwards stimuli included both rigid and non-rigid movements).

The fact that people can use characteristic motion patterns to discriminate unfamiliar faces does not discount the use of structure-from-motion cues as well. Several studies have found that rigid head movement alone (which does not provide any characteristic motion pattern information) can support experimentally familiar face recognition (Pike et al., 1997; Schiff et al., 1986) and unfamiliar face matching (Farivar et al., 2009). However, like famous faces, there has not been a direct comparison of the movement advantage for unfamiliar faces using stimuli that either preserve or normalise face and head shape.

The fact that both familiar and unfamiliar faces appear to use a combination of movement-based cues to aid recognition and matching strongly suggests that participants should show a movement advantage for both avatars and PLDs. However, whether participants will be able to match PLDs and avatars equally well is unclear.

Although both avatars and PLDs have been used widely in face processing research, only one study has directly compared performance with avatars and PLDs using natural movements, and that study used sex discrimination performance as the dependent variable. Hill et al. (2003) used motion capture to record the face and head movements of 12 actors, and then presented the resulting recordings as PLDs, shape-normalised PLDs, or shape-normalised avatars. Participants were equally good at discriminating the sex of the face whether it was presented as a PLD, a shape-normalised PLD, or shape-normalised avatar, but there was a trend for PLDs to be discriminated better than the shape-normalised avatars, which suggests that the structural information in PLDs may be useful over and above characteristic movement information.

The results of Hill et al.’s (2003) study suggest that PLDs could lead to significantly better matching performance than avatars. However, sex discrimination is quite different from matching based on perceived identity, which may rely on different types of movement (Hill & Johnston, 2001). Consequently, an identification task may show a different pattern of performance compared to the sex discrimination task. For example, asking participants to match faces based on identity may lead participants to rely heavily on structure-from-motion cues that are more easily
accessible in PLD stimuli. It is also possible that participants could find it easier to recognise or match idiosyncratic movements in the more abstract PLD form than in shape-normalised avatars, since the movements in PLDs are not being masked by incongruent form information as they are for avatar stimuli. One major limitation of shape-normalised avatars is the fact that averaged form cues still provide some static information, even if it is misleading and/or unhelpful to the task (i.e., the skin texture or shape of the avatars, although constant across the different identities, may mislead or distract participants). Knappmeyer et al. (2003) used morphed face stimuli to demonstrate that even subtle form cues such as texture information can reduce the effect of motion in an identification task. Knappmeyer et al.’s results support the theory that form and motion are encoded together – it may not be possible to dissociate them completely, even when the task requires it. This may be particularly important when matching an avatar to a non-degraded image, or when trying to identify a familiar face based on the avatar’s motion.

On the other hand, it is possible that avatars preserve more useful movement cues than PLDs, because they present information from the whole face. PLDs have been criticised for only showing a sparse array of movement information (Knight and Johnston, 1997), which might only be sufficient to support individual recognition in highly familiar faces (e.g., Bruce & Valentine, 1988; Rosenblum et al., 2007). Although other experiments have successfully used PLDs in unfamiliar face matching tasks (Rosenblum et al., 2002), the fact that participants in Experiments 1 and 2 could not match non-degraded images to PLDs without some overlap suggests that the PLDs used in this experiment may not contain enough information to support matching (as stated above, this is one of the main reasons behind directly comparing the PLDs and avatars). Therefore, it is possible that avatars will be matched more accurately than PLDs – participants may use the extra movement information carried by avatars in order to extract better characteristic motion patterns from a face.

5.1 Experiment 3: Matching Between Formats – Video to PLD and Video to Avatar

This experiment investigated participants’ ability to match a non-degraded video or a still image to a degraded image. Participants were presented with a still image or moving clip of a famous or unfamiliar person, followed by a still image or
moving clip of a PLD or an avatar. Unlike Experiments 1 and 2, there were no overlapping movement sequences in any condition of Experiment 3. Therefore, it was unclear whether participants would be able to match either famous or unfamiliar faces at above-chance levels. If participants could match PLDs or avatars at above-chance levels, there were three main predictions, similar to those in Experiments 1 and 2.

First, based on previous research and the findings from Experiments 1 and 2, it was expected that famous faces would be matched more accurately than unfamiliar faces overall (Bruce et al., 2001; Burton et al., 1999).

Second, it was predicted that a movement advantage would be present for at least some faces. The effect of familiarity on the movement advantage is somewhat inconsistent in previous research, and there was no clear prediction about the possible interaction between familiarity and movement. Overall, though, it was expected that participants would perform worst when presented with two static images, and best when presented with two moving images. Based on Experiments 1 and 2, providing a moving, non-degraded sample image was expected to improve performance dramatically.

Finally, it was predicted that participants would perform somewhat better when viewing PLDs than shape-normalised avatars. However, previous evidence suggests that even shape-normalised stimuli can be matched well above chance levels, so a movement advantage was expected for both stimulus types.

5.1.1 Methods of Experiment 3

5.1.1.2 Participants. Thirty-eight undergraduate students (27 female) from the University of Western Sydney, aged between 18 and 59 years (mean age 23.5 years) participated in this experiment in return for course credit. All reported normal or corrected-to-normal vision. Four participants’ data were excluded from analysis due to their failure to follow the instructions. One participant’s data was excluded due to failure to recognise any of the famous faces.

5.1.1.3 Stimuli and Materials. A set of video images of six highly familiar (famous) and six unfamiliar adult males was obtained from the online content of two talk shows. The clips were the same as those used in Experiments 1 and 2. They were chosen based on undergraduate ratings of familiarity, and were matched based on
ratings of distinctiveness and amount of movement (see section 4.1, Chapter 4). All
clicks showed the person speaking, facing towards the camera in an interview
situation. The clips were selected to show each face from approximately the same
viewpoint and distance from the camera, and to exclude extreme facial movements.

Each 8 s clip was cut into four sequential 2 s videos (with no overlapping
frames), which were used as the full video sample clip, and also as the basis for the
PLDs and avatars in these experiments. The PLDs were created in the same way as
in Chapter 4 (see Figure 7). However, all the clips were 2 s long, as opposed to 1 – 8
s in Experiments 1 and 2. Avatars were created using a custom face-tracking
program (Saragih, Lucey & Cohn, 2010; Saragih, Lucey & Cohn, 2011). The
program tracks 66 points on the face, (including eyes and pupils) and creates an
“avatar” which mimics the movements of the original video sequence, but displays a
uniform shape and texture for all actors. In these experiments, a white “mask” was
chosen for the avatar shape (Figure 13). Examples of the PLD and avatar stimuli
used in Experiments 3 and 4 are included in Appendix A.

The mask was chosen to be a neutral image that did not resemble any of the
faces, and was presented on a black background. In total, 144 videos were created:
four 2 s videos per stimulus type (non-degraded videos, PLDs and avatars), for 12
identities. There was also a corresponding static image, created by taking a single
frame from the middle of each 2 s PLD, avatar and original video clip.

![Positions of the points tracked in the shape-normalised avatar stimuli.](image)

*Figure 13:* Positions of the points tracked in the shape-normalised avatar stimuli.
Participants were only shown the completed image on the right during the experiments.
All video and static images measured 960 x 540 pixels, and were presented on a black background. All videos were presented at 25 frames per second. The experiment was run on a MacBook Pro using Superlab 4.0.3, and images were presented on a BENQ E2200 HD 22-inch monitor, with resolution set to 1920 x 1080 pixels. Participants were tested at an approximate viewing distance of 60 cm.

5.1.1.4 Design and Procedure. Participants completed a same/different identity-matching task. The experiment was a fully repeated measures design, 2 (familiarity: famous/unfamiliar) x 2 (stimulus type: PLD/avatar) x 4 (movement of clips: moving/moving (M/M); moving/static (M/S); static/moving (S/M); static/static (S/S)). To prevent simple matching of identical movement sequences and to encourage a focus on identity matching, the movement sequences and/or static pictures shown in the two images were extracted from different sections of the original video clip. Each condition contained twenty-four trials (two same and two different trials for each identity), resulting in 384 trials per participant. Trials were presented in two blocks, and different clip pairings were presented in each block, to prevent learning of particular movement sequences. Within each block, the order of presentation of trials was randomized. Block presentation order and stimulus pairings were counterbalanced between participants.

Every trial of the main experiment began with a fixation cross, presented in the centre of the screen for 200 ms. The fixation cross was extinguished and replaced by two sequentially presented facial images, each 2 s long, and separated by a 500 ms grayscale noise mask (identical to that used in Experiments 1 and 2, see Figure 8 in Chapter 3). To prevent matching based purely on the location of the face onscreen, the initial face image was randomly offset from the centre of the screen by 60 pixels to the left or right, and the subsequent face image was offset 60 pixels in the opposite direction. Participants were asked to indicate via key press whether the two facial images showed the same person, or two different people. Participants were instructed to respond as quickly and as accurately as possible after the second video had finished. Once the key press was recorded, the next trial began immediately. In Experiment 3, the first stimulus shown in each trial was always a non-degraded video or still image. The second stimulus was either a PLD or an avatar.

Participants were tested individually in a darkened room. They completed 12 practice trials (without feedback) prior to beginning the main experiment, and
received several breaks throughout testing. Following the main experiment, each participant completed the same familiarity check used in Experiments 1 and 2. Once again, participants’ data were excluded from analysis if they did not rate all famous faces 6 or higher and unfamiliar faces 2 or lower, or if they could not name, or provide other unambiguous identity information for, at least three of the six famous faces.

5.1.2 Results of Experiment 3

5.1.2.1 Signal Detection Theory analysis: As in Experiment 2 (Chapter 4), a criterion-independent measure of sensitivity (d’) was calculated for each participant in each condition (MacMillan & Creelman, 2005). Hit and/or false positive scores of zero or one were replaced with values of 0.042 and 0.948 respectively (MacMillan & Kaplan, 1985). In this study, a d’ of 0 represents chance performance, while a d’ of 3.46 represents perfect performance in a condition. One sample t-tests were carried out to compare the resulting d’ scores to chance performance levels. Participants performed significantly above chance in 13 out of 16 conditions, all ps < .05 (see Table 10). All three conditions that were not significantly above chance involved matching unfamiliar faces to avatars – the M/S, S/M and S/S conditions, ps >.08.

A three-way within-subjects ANOVA (2 x familiarity; 2 x stimulus type, PLD or avatar; 4 x movement of clips; M/M, M/S, S/M, S/S) was carried out on the d’ scores. The main effects of familiarity and stimulus type were significant. Famous faces were matched significantly better than unfamiliar faces, F(1,32) = 6.33, p = .017, ηp² = .16, and there was a trend for PLDs to be matched better than avatars, F(1,32) = 3.99, p = .054, ηp² = .11. The familiarity by stimulus type interaction was significant, F(1,32) = 4.73, p = .037, ηp²=.13. Pairwise comparisons (Bonferroni corrected) revealed that famous faces were recognised equally well from PLDs and avatars, p = .845, but unfamiliar faces were recognised significantly better from PLDs than avatars, p = .003. No other interactions were significant, ps > .6.

---

8 Analyses were also carried out to examine the effects of block, but no main effects or interactions were significant (ps > .2), therefore block was excluded from any further analysis.
Table 10:
*Mean d’ for Experiment 3*

<table>
<thead>
<tr>
<th>familiarity</th>
<th>Avatars</th>
<th>Point-light-displays</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M/M</td>
<td>M/S</td>
</tr>
<tr>
<td>Famous</td>
<td>0.85**</td>
<td>0.62*</td>
</tr>
<tr>
<td>SD</td>
<td>1.13</td>
<td>0.92</td>
</tr>
<tr>
<td>Unfamiliar</td>
<td>0.47*</td>
<td>0.29</td>
</tr>
<tr>
<td>SD</td>
<td>1.16</td>
<td>1.23</td>
</tr>
</tbody>
</table>

*Note: Trials had two videos (M/M); a non-degraded video paired with a static PLD/avatar (M/S); a non-degraded static image paired with a moving PLD or avatar (S/M), or two static images (S/S). A score of 0 is chance, and the maximum is 3.46. *p < .05 ** p < .0005.*

Surprisingly, there was no main effect of movement of clips, $F(3,93) = .992$, $p = .40$, $\eta^2_p = .031$, and no interactions, $ps > .15$. However, in the interest of investigating the familiarity effect, Bonferroni-corrected pairwise comparisons for famous and unfamiliar faces were carried out in each movement condition and for each stimulus type. In general, famous and unfamiliar faces were matched equally well in all conditions, $ps > .14$, except for trials with S/S avatar images, where famous faces were matched significantly better than unfamiliar faces, $p = .006$.

**5.2.1.2 Movement advantage:** In order to investigate whether moving stimuli were matched better than static stimuli, a secondary analysis of the movement advantage was conducted. To calculate the movement advantage, the d’ value for the S/S condition was subtracted from each other movement condition. The resulting movement advantage scores are shown in Figure 14. Movement advantage scores for each condition were subjected to one sample t-tests comparing them to 0 (no movement advantage), and a three-way within-subjects ANOVA (2 x familiarity; 2 x stimulus type; 3 x movement of clips). No conditions were significantly greater than 0, $ps > .17$, and there were no main effects or interactions, all $ps > .05$. 
5.2.1.3 Bias, hits and correct rejections. As in Experiment 2 (Chapter 4), bias was measured using criterion c (MacMillan & Creelman, 2005). Three-way within-subjects ANOVAs (2 x familiarity; 2 x stimulus type, PLD or avatar; 4 x movement of clips, M/M, M/S, S/M, S/S) were carried out on the bias, hit and correct rejection (CR) scores. Overall, participants displayed a neutral response bias, c = 0.07, SD = 0.79. There was no significant effect of stimulus type, $F(1,32) = 3.27, p = .080, \eta_p^2 = .09$. Participants remained relatively neutral when viewing famous faces, c = 0.03, but were significantly more biased towards a “different” response when viewing unfamiliar faces, c = 0.10, $F(1,32) = 4.82, p = .035, \eta_p^2 = .13$. Famous faces also resulted in significantly more hits than unfamiliar faces, $F(1,32) = 9.74, p = .004, \eta_p^2 = .23$, but familiarity had no effect on CRs, $F(1,32) = 0.01, p = .921, \eta_p^2 = .000$. This pattern is the opposite of that found in Experiment 2b, but it is not immediately apparent why participants response patterns differed so much between the experiments. The bias results also revealed a significant interaction between familiarity and stimulus type, $F(1,32) = 16.4, p < .0005, \eta_p^2 = .34$. Participants
showed equal response bias for PLDs, regardless of familiarity, $p = .552$, but significantly more bias towards a “different” response for unfamiliar than famous faces presented as avatars, $p < .0005$.

The movement condition also had a significant effect on bias, $F(3,96) = 35.78$, $p < .0005$, $\eta^2_p = .53$. Pairwise comparisons showed that M/M trials resulted in a bias towards a “same” response, $c = -0.41$, whereas M/S and S/M trials resulted in a “different” bias, M/S: $c = .47$, S/M: $c = .27$. S/S trials resulted in a relatively neutral response bias, $c = -.07$. All pairwise comparisons on bias between movement conditions were significant after Bonferroni-correction, $ps < .05$. There was also a significant effect of movement in the hit and CR analyses. Not surprisingly, the M/M condition had significantly more hits than any other movement condition, all $ps < .0005$, and the M/S and S/M conditions had significantly more CRs than either the M/M or S/S conditions, all $ps < .05$. These findings are very similar to those in Experiment 2 – when the clips were perceptually dissimilar (i.e., one moving and one static), participants were more likely to respond “different” than if the clips were more perceptually similar (i.e., both moving or static).

There were also several significant interactions in the bias analysis. The effect of movement condition interacted significantly with stimulus type, $F(3,96) = 10.25$, $p < .0005$, $\eta^2_p = .24$, and familiarity, $F(3,96) = 10.23$, $p < .0005$, $\eta^2_p = .24$. Overall, the pattern of bias was similar to the main effect of movement for PLDs and avatars, except that for avatars, bias for S/M and S/S was only approaching significance, $p = .061$; and for PLDs the M/S and S/M conditions were not significantly different, $p = 1$. Similarly, famous and unfamiliar faces followed the same bias pattern as the main effect of movement except that pairwise comparisons between famous faces in the S/M and S/S conditions; and unfamiliar faces in the M/S and S/M conditions, failed to reach significance, $ps > .1$.

### 5.1.3 Discussion of Experiment 3

Experiment 3 showed that participants are capable of matching non-degraded images of famous faces to PLDs and avatars, but performance was equal for moving and static stimuli. Similarly, participants were capable of matching non-degraded images and videos of unfamiliar faces to PLDs (and avatars, when both clips were moving), but performance was equal for moving and static stimuli. In line with the results from Experiments 1 and 2, famous faces were matched better overall.
Perceptual similarity of the stimuli also appeared to have a similar effect on bias in Experiments 2 and 3 – participants were more inclined to respond “same” when the stimuli were both moving (or both static) than when one was moving and one was static. In contrast to Experiment 2b, unfamiliar faces resulted in a greater bias towards a “different” response, and significantly fewer hits, than famous faces.

There are several explanations for the lack of movement advantage. It is possible that participants were reliant on static cues throughout the experiment. Participants could have been using primarily shape-based cues to match the famous and unfamiliar PLDs, since the PLDs contained some structural information even in the absence of motion. When presented with avatars, which had no distinguishing shape cues, participants may have relied on characteristic static face and head poses to match videos of famous faces with their avatar counterparts. This may also explain why participants had a higher hit rate for famous faces – it is possible that participants were more familiar with characteristic head poses of famous faces, and more confident making a positive match. The use of static cues may also account for the finding that PLDs were matched somewhat more accurately than avatars. It is possible that having conflicting form and texture cues in the non-degraded image and the shape-normalised avatar confused participants and resulted in poor performance (Knappmeyer et al., 2003). In other words, the switch from non-degraded image to uniform avatar might have overshadowed the movement cues, by attracting attention to irrelevant or useless information such as the structure, features, or texture of the face.

Alternately, it is possible that participants are unable to extract enough useful characteristic movement information from PLDs and avatars to support a movement advantage. This seems unlikely, since movement advantages have been demonstrated with similar avatar and PLD stimuli before (Hill & Johnston, 2001; Rosenblum et al., 2007). However, previous studies have rarely asked participants to generalise from non-degraded videos to PLDs or avatars. Burton et al. (1999) and Bruce et al. (2001) demonstrated that generalising from a degraded image to a high-quality photograph is difficult, particularly for unfamiliar faces. Having a non-degraded image may be helpful for comparison of movement or form cues, but it is possible that the extra information is distracting, or that participants simply cannot generalise from non-degraded videos to PLDs and avatars in an identification task.
The presence of a non-degraded image may also explain the overall matching advantage for famous faces – it is possible that participants were paying more attention to the faces that they knew, which could have led to better performance in famous face trials. These possibilities were tested in Experiment 4 by replacing the non-degraded image with a second degraded image.

### 5.2 Experiment 4: Matching Within Formats – PLD to PLD and Avatar to Avatar

Experiment 3 failed to find a movement advantage for either famous or unfamiliar faces, presented as PLDs or avatars. This replicates the lack of movement advantage in Experiments 1b and 2b (although in general performance levels in Experiment 3 were much higher than Experiments 1b and 2b). Experiment 4 was designed to determine whether this lack of effect was due to problems with the stimuli, or whether participants could not generalize from non-degraded videos to degraded stimuli.

Experiment 4 investigated whether participants could match two degraded images; that is, two PLD or avatars. If the nature of the movement information in the PLD and avatar stimuli (i.e., the fact that the moving clips were matched for amount and distinctiveness of movement) caused the lack of movement advantage in Experiment 3, it was predicted that similar results should arise in Experiment 4. However, if participants had trouble generalizing motion from non-degraded video sequences to PLDs and avatars, it would be expected that the results for both types of stimulus would improve significantly.

As each trial only contained degraded stimuli (either PLDs or avatars), the effect of familiarity was expected to diminish considerably, because participants were no longer able to recognize a face and draw on memories of that person to match characteristic head poses or typical movement sequences. However, previous studies have found that participants can match both unfamiliar (Hill & Johnston, 2001) and personally familiar faces (Cook et al., 2012) presented as shape-normalised avatars, so both familiar and unfamiliar faces were still expected to give rise to a movement advantage. No previous studies have created avatars using famous faces, though, or compared performance with familiar or unfamiliar face
avatars, so it was unclear whether famous faces should result in a larger movement advantage than unfamiliar faces.

Finally, if the superior matching performance for famous faces in Experiment 3 was due to different levels of interest or attention for famous and unfamiliar face trials – for example, participants paying more attention to trials in which a famous face appeared – presenting two degraded images should eliminate the matching advantage for famous faces altogether.

5.2.1 Methods of Experiment 4

5.2.1.1 Participants. Seventeen undergraduate students (12 female) from the University of Western Sydney, aged between 18 and 45 years (mean age 22 years) participated in this experiment in return for course credit. All reported normal or corrected-to-normal vision. One participant’s data was excluded from analysis due to their failure to follow the instructions.

5.2.1.2 Stimuli, Design and Procedure. The experiment was identical to Experiment 1, except that each trial showed either two PLDs or two avatars, and image manipulation was blocked (participants viewed all PLDs in one block and all avatars in another). Block presentation order and stimulus pairings were counterbalanced between participants.

5.2.2 Results of Experiment 4

5.2.2.1 Signal Detection Theory analysis. Once again, a d’ score was calculated for each participant in each condition, and t-tests were carried out to compare each condition to chance. The d’ results for Experiment 4 are shown in Table 11. As in Experiment 3, performance was above chance in the majority of conditions ($p < .05$ in 12 out of 16 conditions). Chance-level performance was found for the S/S condition for PLDs (famous and unfamiliar) and unfamiliar avatars, and for the S/M condition for unfamiliar avatars. This pattern of results is markedly different from Experiment 3, where only unfamiliar faces with avatar test images were matched at chance levels. The results from Experiment 4 suggest that removing the non-degraded image encouraged participants to rely more heavily on movement-based cues, rather than matching known famous face images to similar-shaped avatars and PLDs.
Table 11: Mean $d'$ for Experiment 4.

<table>
<thead>
<tr>
<th></th>
<th>Avatars</th>
<th>Point-light-displays</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M/M</td>
<td>M/S</td>
</tr>
<tr>
<td>Familiar</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Famous</td>
<td>1.24**</td>
<td>1.08**</td>
</tr>
<tr>
<td>SD</td>
<td>0.72</td>
<td>0.75</td>
</tr>
<tr>
<td>Unfamiliar</td>
<td>1.01**</td>
<td>0.78*</td>
</tr>
<tr>
<td>SD</td>
<td>0.90</td>
<td>0.98</td>
</tr>
</tbody>
</table>

Note: Trials had two videos (M/M); a non-degraded video paired with a static PLD/avatar (M/S); a non-degraded static image paired with a moving PLD or avatar (S/M), or two static images (S/S). A score of 0 is chance, and the maximum is 3.46. *

A three-way within-subjects ANOVA (2 x familiarity; 2 x stimulus type, PLD or avatar; 4 x movement of clips, M/M, M/S, S/M, S/S) was carried out on the $d'$ scores. Replicating the results from Experiment 3, there were significant effects for familiarity and stimulus type. Famous faces were matched better than unfamiliar faces, $F(1,15) = 17.46, p = .001, \eta_p^2 = .54,$ and PLDs were matched better than avatars, $F(1,15) = 8.56, p = .010, \eta_p^2 = .36.$ However, unlike in Experiment 3, there was also a significant main effect of movement of clips, $F(1.85, 27.72) = 9.94, p = .001, \eta_p^2 = .40$ (Greenhouse-Geisser correction applied for sphericity violations), and a significant interaction between stimulus type and the movement condition, $F(2.48, 37.17) = 4.04, p = .019, \eta_p^2=.212.$ Overall, M/M and M/S clips were matched better than S/S clips ($ps < .01$), but no other pairwise comparison were significant ($ps > .05$). The interaction reflected the fact that avatars were matched equally well in all movement conditions, all $ps > .3,$ whereas PLDs in the M/M, M/S and S/M conditions were matched significantly better than in the S/S condition, all $ps < .02.$ PLDs in the M/M, M/S, and S/M conditions were matched equally well, all $ps > .1.$

There was a marginally significant interaction between familiarity and movement condition, $F(2.07, 31.81) = 2.74, p = .077, \eta_p^2 = .15.$ As in Experiment 3, Bonferroni-corrected pairwise comparisons were carried out to assess the difference.
between famous and unfamiliar faces in all conditions. Once again, famous faces presented as avatars were matched significantly better than unfamiliar faces in the S/S condition, \( p = .001 \), but the effect of familiarity in all other conditions was not significant, \( ps > .05 \).

### 5.2.2.2 Movement advantage

To investigate the movement advantage, and in particular the interaction between manipulation and movement from the d’ analysis, the d’ value for the S/S movement condition was subtracted from the values of each other condition for avatars and PLDs and subjected to secondary analyses. The movement advantage score for each condition is shown in Figure 15. Unlike Experiment 3, the movement advantage was significant in the majority of conditions, with the exception of all three famous avatar conditions (\( ps > .4 \)) and the unfamiliar avatars in the S/M movement condition (\( p = .194 \)).

A three-way within-subjects ANOVA (2 x familiarity; 2 x stimulus type; 3 x movement of clips) showed all main effects were significant. There was a larger movement advantage for unfamiliar faces, \( M = 1.15 \), than for famous faces, \( M = 0.56 \), \( F(1,15) = 5.64, p = .031, \eta_p^2 = .27 \), although both famous and unfamiliar faces showed an overall movement advantage score greater than 0, \( ps < .05 \). The advantage for unfamiliar faces is surprising, as famous faces were matched more accurately overall in the signal detection theory analysis. PLDs, \( M = 1.31 \), showed a greater movement advantage than avatars, \( M = 0.41 \), \( F(1,15) = 7.80, p = .014, \eta_p^2 = .34 \), across all three movement conditions, \( ps < .05 \). Overall, the avatars did not show any significant movement advantage, \( p = .147 \), but PLDs did, \( p = .001 \), which suggests that the better matching performance for PLDs in the d’ analysis was not simply a consequence of better static cues in the PLDs. Finally, the effect of movement of clip was also significant in this analysis, \( F(2,30) = 3.94, p = .030, \eta_p^2 = .21 \). Pairwise comparisons (Bonferroni-corrected) reveal that the M/M condition, \( M = 1.07 \), resulted in a larger movement advantage than the S/M condition, \( M = 0.65, p = .033 \), but not the M/S condition, \( M = 0.85, p = .220 \); all three conditions (M/M, M/S, S/M) led to a significant movement advantage, \( ps < .025 \). No interactions were significant, \( ps > .1 \).
5.2.2.3 Bias, hits and correct rejections. The overall bias in Experiment 4 was close to neutral, $c = 0.03$, $SD = 0.44$. There was no significant effect of stimulus type, $F(1,15) = 0.12$, $p = .737$, $\eta^2_p = .01$. Unlike Experiment 3, there was no main effect of familiarity on bias, $F(1,15) = 3.99$, $p = .064$, $\eta^2_p = .21$. The analysis of hits also found no main effect of familiarity, $F(1,15) = 2.68$, $p = .123$, $\eta^2_p = .15$. However, familiarity did have a significant effect on CRs. Mismatched famous faces were correctly rejected more often than mismatched unfamiliar faces on average, $F(1,15) = 18.38$, $p = .001$, $\eta^2_p = .55$.

The bias analysis revealed a significant main effect of movement condition, $F(3,45) = 24.52$, $p < .0005$, $\eta^2_p = .62$. In M/M trials, participants showed a “same” bias, $c = -0.27$; in M/S and S/M trials there was minimal bias, M/S: $c = .032$; S/M: $c = -.018$; and in S/S trials participants showed a “different” bias, $c = 0.38$. All pairwise comparisons were significant, $ps < .005$, except for the M/S to S/M comparison, $p = 1$. The hit statistics mirrored this effect almost perfectly: the M/M condition, $c = 0.73$, had significantly more hits than M/S, $c = 0.62$, and S/M, $c = 0.61$, which in turn had significantly more hits than the S/S condition, $c = 0.41$. 

---

**Figure 15**: The movement advantage for different movement conditions in Experiment 4. The movement advantage was calculated as the difference in accuracy between moving and static PLD or avatar trials. Error bars represent +/- 1 standard error of the mean.
Again, all comparisons were significant, $ps < .005$, except the M/S to S/M comparison, $p = 1$.

There was a significant interaction between movement and stimulus type in the bias analysis, $F(3,45) = 12.51, p < .0005, \eta^2_p = .46$. Like in Experiment 3, the pattern of bias for PLDs mirrored the main effect of movement (all $ps < .05$ except M/S and S/M, $p = 1$). However, the bias for avatars only varied slightly across the movement conditions: M/M showed significantly more bias towards a “same” response than M/S and S/S conditions, $ps < .05$, but no other comparisons reached significance. Finally, the interaction between familiarity and movement condition was also significant in the bias analysis, $F(3,45) = 5.77, p = .002, \eta^2_p = .28$, and in the hits analysis, $F(3,45) = 3.44, p = .025, \eta^2_p = .19$, but not in the CR analysis, $F(3,45) = 2.18, p = .111, \eta^2_p = .12$. The bias interaction reflected the fact that participants viewing famous faces had a significantly larger bias towards a “different” response than unfamiliar faces in the M/S condition ($p < .0005$). Pairwise comparisons on the hits interaction revealed a different pattern: famous faces had significantly more hits than unfamiliar faces in the S/S condition, $p = .005$, but not in any other condition, $ps > .1$. Famous faces were correctly rejected more often than unfamiliar faces in the M/S, S/M, and S/S conditions, all $ps < .05$.

5.2.2.4 Image dissimilarity analysis. The overall $d'$ results for Experiment 4 showed that participants were very good at discriminating between famous faces presented as avatars when both images were static. This is puzzling, since the majority of identifying static cues (texture, structure, features) had been rendered uninformative in the avatars condition. Furthermore, the effect was isolated to famous face avatars: the same effect did not arise for famous faces presented as PLDs, or unfamiliar faces presented as avatars. It is possible the effect arose because the static famous face avatars were more similar than the static unfamiliar face avatars, or because the famous face stimuli in the “different” trials were more dissimilar than the equivalent unfamiliar face stimuli. To test this effect, an analysis of image dissimilarity was carried out on all the static avatar pairs used in Experiment 4. Image dissimilarity was calculated as the root-mean-square difference in greyscale values across each pair of images, after the greyscale values of the pixels were normalised (Henson, Mouchlianitis, Matthews, & Kouider, 2008; Vuilleumier, Henson, Driver, & Dolan, 2002). In this analysis, identical images
would have a score of 0 and completely dissimilar images would have a score of 1. Since the images were offset, and participants may have considered this when judging the similarity of the image pairs, the image analysis was run at every possible pixel offset, and the minimum score was retained (this represents the closest possible match between the two images). A 2 (familiarity) x 2 (same or different trial) ANOVA was conducted on the resulting minimum scores. Both main effects were significant, and the interaction approached significance, familiarity: $F(1,92) = 18.70, p < .0005, \eta^2_p = .17$; same/different: $F(1,92) = 41.82, p < .0005, \eta^2_p = .31$; familiarity x same/different: $F(1,92) = 2.98, p = .087, \eta^2_p = .03$. Unsurprisingly, images of different people, $M = .028$, were, on average, less similar than images of the same person, $M = .017$. Unexpectedly, though, famous faces images, $M = .026$, were less similar than unfamiliar face images on average, $M = .019$ – this discounts the idea that static face matching of famous faces was better because the famous face static pairs were easier match to when they were the same. Pairwise comparisons on the interaction between familiarity and same/different revealed that famous and unfamiliar faces in “same” trials were equally similar, $p = .07$, but famous faces in the “different” trials were more dissimilar than unfamiliar faces, $p = .002$.

The image similarity statistics suggest that the matching advantage for static famous face avatars could have arisen in the “different” trials – the mismatched famous faces were more dissimilar than the mismatched unfamiliar faces. However, that pattern of results is not borne out by the hit and CR analyses. If image dissimilarity explained the results for famous faces presented as static avatars, it would be expected that famous faces should have had a similar hit rate, and a significantly higher CR rate, compared to unfamiliar faces. Pairwise comparisons (with Bonferroni-correction) show that famous face avatars resulted in significantly more hits ($p = .022$) and CRs ($p = .013$) than unfamiliar face avatars in the S/S condition. This effect was not a result of differences in bias – famous and unfamiliar avatars in the S/S condition had similar levels of bias, $p = .741$. Perceptual similarity between two static images may explain why famous faces were correctly rejected more than unfamiliar faces, but it is still unclear why famous faces were correctly matched more often than unfamiliar faces in the static to static avatar condition.
5.2.3 Discussion of Experiment 4

Overall, Experiment 4 showed a strong, significant movement advantage for famous and unfamiliar faces presented as PLDs. It is interesting that, once again, unfamiliar faces had a larger movement advantage than famous faces, despite the fact that famous faces were matched more accurately overall. The results from Experiments 1a, 2a, and 4 suggest that participants are not better at matching movement patterns that they are familiar with – rather, movement is more helpful for faces they have not seen before.

It is possible that the beneficial effect of familiarity in Experiment 3 could have arisen because participants paid more attention to trials with famous faces than unfamiliar faces. However, in Experiment 4 participants were not aware that some of the faces presented were famous until after the experiment. As they had no way of knowing who the famous people were, the results from Experiment 4 eliminate the possibility that participants were consciously looking for particular facial characteristics or poses, or that they were paying more attention to the trials with famous faces in them.

5.3 General Discussion of Experiments 3 and 4

In this chapter, two experiments were presented that compared matching performance for famous and unfamiliar faces, presented as point-light-displays and shape-normalised avatars. Overall, famous faces were matched better than unfamiliar faces, and PLDs were matched better than avatars. When matching from a non-degraded image to a degraded face, movement did not improve matching performance (Experiment 3). However, moving stimuli were recognised significantly better than static stimuli when two degraded face images were presented in each trial (Experiment 4). The movement advantage in Experiment 4 was larger for unfamiliar faces than famous faces, and larger for PLDs than for avatars.

5.3.1 Familiarity and Movement Effects in Experiments 3 and 4

We found an overall benefit of familiarity for matching in both experiments. Famous faces were matched better than unfamiliar faces, even when participants were not aware that the trials contained famous faces (Experiment 4). Despite the
overall familiarity advantage, there was only a small benefit of viewing famous faces in motion, which disappeared when participants were asked to match non-degraded videos to degraded stimuli. Unfamiliar faces also failed to show any motion advantage when generalizing from non-degraded images to degraded stimuli. However, there was a large movement advantage for unfamiliar faces in Experiment 4, when participants were matching the same stimulus type in each trial. Interestingly, the movement advantage for unfamiliar faces in Experiment 4 was larger than for famous faces. This is the third time in this series of experiments that a larger movement advantage for unfamiliar faces than for famous faces. However, in Experiments 1 and 2 (Chapter 4), the advantage for unfamiliar faces was confined to PLDs that overlapped with the non-degraded comparison video, suggesting that participants were relying primarily on movement matching. In the current chapter, the images did not contain any overlap. Participants in Experiment 4 could generalize from one movement sequence to another moving or static image, as is shown by the significant movement advantage in the majority of conditions.

The difference between the size of the movement advantage for famous and unfamiliar face matching in Experiment 4 is surprising, as previous studies on famous faces have often found a significant movement advantage (Knight & Johnston, 1997; Lander & Bruce, 2004; Lander et al., 2001), whereas several studies using experimentally familiar or unfamiliar faces have failed to find any effect of movement (Bruce & Valentine, 1988; Bruce et al., 2001; Christie & Bruce, 1998). However, the experimentally familiar face studies have generally used old/new recognition tests, which require participants to recall faces after a long delay (e.g., 30 mins in Christie and Bruce’s study). It is possible that using a same/different task, which allows participants to compare motion more directly, is more conducive to finding a motion advantage in unfamiliar faces (Thornton & Kourtzi, 2002). This idea is also supported by the findings reported in Chapter 4.

The fact that there was a larger movement advantage for unfamiliar faces in Experiment 4 is also surprising because our participants should have been able to use the motion of famous faces in several ways to help identification – via characteristic motions patterns, and via structure-from-motion processes. However, unfamiliar faces should only have access to structure-from-motion, since the accumulation of supplemental information is likely to require multiple exposures to the face (Roark et al., 2003). Although it is possible that characteristic motion information is extracted
from brief periods of exposure, it seems unlikely that a movement advantage from such a rapid process would be more robust than the motion patterns built up over multiple exposures to famous faces.

There are several possible explanations for the famous face matching advantages and unfamiliar face movement advantage. One possibility is that there were systematic differences between the famous and unfamiliar video stimuli. However, the videos were matched based on ratings of the amount and distinctiveness of movement prior to the experiment, in order to minimize movement differences within the stimuli. Another possibility is that the famous face stimuli contained more static cues than the unfamiliar stimuli. However, matching performance in both famous and unfamiliar PLD-to-PLD and unfamiliar avatar-to-avatar trials in Experiment 4 was at chance levels. Furthermore, the analysis of image dissimilarity revealed that static famous and unfamiliar avatars were equally similar in “same” trials; perceptual similarity cannot account for the higher overall hit rate for static famous face avatars (although image dissimilarity may explain why participants were better at correctly rejecting static famous avatars than static unfamiliar avatars in “different” trials).

An alternative explanation for the famous face matching advantage in this study, as in Experiments 1a and 2a, is that participants might use movement information to aid identification of both famous and unfamiliar faces, but they could be more sensitive to the static cues present in famous faces than unfamiliar faces. For example, participants may be better at extracting shape cues from static images of faces they know, or they may simply be better at recognising characteristic face and/or head poses. This is particularly true for Experiments 1a, 2a, and 3, where participants could recognise the face shown in the initial image, and therefore may have been responding to remembered images of the person as well as the stimuli that were presented.

Participants in Experiment 4 could not match faces based on memories of characteristic poses, but it is possible that participants were using the extra details from the avatar stimuli (such as the eyes, or more detailed representations of the lips) as an extra static cue that contained characteristic face information. Research using static faces has shown that presenting a famous face in an “iconic” pose aids recognition (Carbon, 2008), and it is possible that familiarity with a face means we become more sensitive to typical expressions (and perhaps poses) that occur
regularly in our experience of a person (Burton et al., 2011; Kaufmann & Schweinberger, 2004). If participants were better at extracting and matching static face cues from famous faces, it would increase overall matching performance for famous faces (as was the case in both experiments), but, as discussed in Chapter 4, increasing static matching can lead to an asymptotic effect, and reduce or eliminate the effects of motion for famous faces. Previous studies have often used neutral static images of famous faces (e.g., Knight & Johnston, 1997), which may have reduced the amount of distinctive face and head poses in their static conditions. Since these distinctive poses were still likely to be present in the dynamic conditions, it is possible that previous research has artificially inflated the movement advantage for famous faces by not including all identifying static cues, as discussed in Chapter 4.

In general, the findings across Chapters 4 and 5 suggest that the benefit of familiarity for moving faces is not as large and as clear-cut as has previously been assumed. Many studies have established that people are better at recognising familiar than unfamiliar faces when they are presented as static pictures (Johnston & Edmonds, 2009). Furthermore, several studies on movement-based face recognition, including Experiment 4, have shown a movement advantage for familiar faces (e.g., Knight & Johnston, 1997). However, our results suggest that the benefits of familiarity and of movement for matching are not additive – rather, movement may help improve recognition only up to a certain level. Since static processing of familiar faces is already very good, familiar faces may only derive a small benefit from movement. On the other hand, static processing of unfamiliar faces is known to be less robust – for example, people are less sensitive to spatial changes in unfamiliar than familiar faces (Brooks & Kemp, 2007). As people are less proficient at using static cues in unfamiliar face matching, they may rely more on movement to provide a cue to identity, and hence derive more benefit from movement than familiar faces. Since this result was evident in Experiment 4, where people did not know which faces were famous and which were unfamiliar, it is clear that the benefit is not based on participants paying more attention to the famous faces, or deliberately focusing more on the static cues present in famous faces.

5.3.2 Stimulus Type Effects in Experiments 3 and 4

The comparison of shape-normalised avatars and non-normalised PLDs in Experiments 3 and 4 also resulted in some surprising findings. Participants were
poorer at matching avatars than PLDs overall, regardless of whether they were matching from a non-degraded image or another avatar. Furthermore, avatars resulted in no overall movement advantage in either experiment (although unfamiliar faces presented as avatars showed a significant movement advantage in all three movement conditions in Experiment 4). On the other hand, participants showed a significant movement advantage when matching from one PLD to another. If participants were relying purely on characteristic movement information to match the faces, PLDs and avatars should show an equal movement advantage, since they both portrayed the same face and head movements. If anything, the avatars portrayed more movement than the PLDs (66 points tracked compared to 27), which would be expected to lead to a slight advantage for the avatars if characteristic movement patterns were the only basis of the movement advantage. Likewise, if the movement advantage arose simply because the moving images had more frames, and hence more information about head poses and expressions (or low-level information such as screen position), the movement advantage for PLDs and avatars should have been similar in size. However, the difference between PLDs and avatars in both experiments suggests that participants were also using information other than characteristic face and head movements.

There are at least two possible explanations for the different patterns of matching performance for PLDs and avatars. It is possible that participants were distracted by the presence of a face when they were viewing the avatars. The fact that the avatars had a shape, texture and features (even if these details were uniform across the stimulus set) may have caused participants to attend to superficial form cues, rather than relying solely on the movement information. This may explain the matching advantage for famous faces. If participants were basing their decisions on superficial cues such as the position of the head or whether the person was smiling, they may have been more sensitive to characteristic head and face poses in the avatar stimuli. Since participants were highly familiar with all of the famous faces, they may have been able to match the face and head poses to representations from memory, explaining the unusually good matching performance for famous avatars in the static-to-static condition of Experiment 3. However, if participants had no characteristic poses stored in memory (as for unfamiliar faces), focusing on the form cues would not have been a good strategy, which could explain why participants
performed poorly when matching non-degraded unfamiliar images to avatars in Experiment 3.

An increased focus on form information may also explain the differing results for moving PLDs compared to avatars. Knappmeyer et al.’s (2003) results suggest that people cannot completely disassociate form and motion cues – even subtle static information, like texture cues, reduced the influence of movement in their experiments. If a person is familiar with the actor whose motion is being projected onto the averaged head, it may be difficult for them to ignore the strong, yet uninformative, form cues that are present in the avatar and to focus on the task-relevant motion cues. In other words, if participants were focusing primarily on matching form-based cues, it is possible that they were somewhat distracted when those form cues (e.g. the shape of the face, or the spacing between the eyes) did not match, as was the case when matching from non-degraded images to shape-normalised avatars in Experiment 3. Furthermore, it is possible that participants focused on form cues in the avatar stimuli at the expense of characteristic movement information, leading to the lack of overall movement advantage in Experiment 4.

Another possible reason for the effect of stimulus type in the current study is the presence of structural cues in the PLDs, but not in the avatars. Participants may have used this basic structural information to match non-degraded images to PLDs in Experiment 3. In Experiment 4, where there was no non-degraded image to provide reliable face-shape information, participants may have used the moving PLDs to extract structure-from-motion information, which then gave rise to a significant movement advantage.

It is important to note that trials containing two dynamic unfamiliar avatars had a significant movement advantage in Experiment 4, which suggests that characteristic movement patterns can still give rise to a movement advantage in the absence of identifying structural cues. However, given that the movement advantage for PLDs was strong and consistent throughout Experiment 4, it seems that it arose primarily from structure-from-motion. Furthermore, this effect does not just arise due to better static processing of PLD images: in Experiment 4, trials with two static

---

9 It is unclear why participants were unable to use this information in Experiments 1b or 2b – perhaps the presence of avatars, which did not contain any structural cues, highlighted the fact that the PLDs contained identifying structural information.
PLDs were matched worse than trials with two static avatars. This is not the first experiment to find a movement advantage based on structural information: Pike et al. (1997) showed that participants were more accurate at recognising experimentally familiar faces that were learnt in rigid rotational motion, and Farivar et al. (2009) showed that participants could match unfamiliar faces based on structure-from-motion cues alone. Our research extends this finding, demonstrating that structural information can be extracted from conversational movement, without a familiarisation/training phase, and even when the face image is severely degraded.

Once again, though, it is unclear why the same results did not arise in Experiments 1, 2, and 3. It is possible that participants cannot extract and compare structure-from-motion information across different formats (i.e., from non-degraded images to PLDs or avatars). Alternatively, participants might have been distracted by the social cues in the non-degraded videos: non-degraded videos may have prompted people to try to identify the person, speech-read, or interpret the conversational and emotional expressions present in the initial clip.

Finally, and more probably, it is possible that a combination of factors influenced the use of structure-from-motion and characteristic motion pattern information across Experiments 1-4. When the format of the images changed, participants may have been unable to extract or compare structure-from-motion information or characteristic movement patterns, perhaps because they were distracted by the presence of social cues, or simply had difficulty generalising from one image format to another. Instead, they relied on movement matching, as in Experiments 1a and 2a; or superficial form cues, as for the famous faces (particularly in the avatar condition in Experiment 3). When the format of the images did not change (Experiment 4), participants may have attempted to match the faces based on superficial form cues (i.e., static pose cues in famous face stimuli). If these static cues were not present or informative, as in the case of unfamiliar faces, participants may have been using structure-from-motion cues preferentially, and resorted to matching characteristic motion patterns last. It is unclear why structure-from-motion would be the most prevalent movement cue – perhaps because it is useful when matching to or from moving and static images; perhaps because structure-from-motion information (i.e., information about the shape of the face and head) is more invariant than characteristic motion patterns, which could change significantly depending on context (see, for example, Lander et al., 2007, where participants could
not match cross-modal stimuli when the manner of speech was changed). In sum, the results from Experiments 1-4 suggest that participants use of static cues and movement information is flexible, possibly hierarchical, and may vary depending on the availability of static information – for example, participants may concentrate on matching static cues, and only use movement information when this is ineffective. The use of movement information may be influenced by the reliability of the cues: identical movement sequences are more reliable indicators of identity than structure-from-motion, which is in turn more reliable than characteristic movement patterns in an identity-based task. In short, a number of factors, including the presence of static information and the reliability of movement information, may determine which cues participants attend to when matching from a non-degraded image to a PLD or avatar, or when matching two degraded faces. This idea will be discussed further in Chapter 8.

Both the presence of a face in the avatars and the presence of shape information in the PLDs could account for the effects of image manipulation in the current study, but it is unclear how these explanations relate to previous research that compared PLDs and solid-body animations. It is possible that Hill et al. (2003) found no significant difference between PLDs and shape-normalised avatars because their participants were focusing primarily on the movement cues, and did not pay as much attention to the extra structural cues offered by the PLDs. Alternatively, it is possible that participants in Hill et al.’s study were not distracted by the form information carried by the avatars because they were not performing a matching task, and as such were not aware of the minor structural differences between the original images and the final avatars. Regardless, the fact that PLDs were discriminated better overall than avatars in our study suggests that participants may pay attention to different cues in an identity-matching task than a gender classification task.

Overall, the effect of stimulus type in this study suggests that PLDs are more conducive to a movement advantage than avatars. At this point it is unclear whether this is due to underlying structural cues in PLDs giving rise to a structure-from-motion-based motion advantage, as suggested by Hill et al. (2003); or whether having an artificial “face” with the wrong structure or texture interferes with movement-based identification, as suggested by Knappmeyer et al.’s (2003) research on facial morphs.
5.3.3 Matching Between Formats vs. Matching Within Formats

An overall examination of both experiments revealed that participants were much better at matching faces in Experiment 4 than Experiment 3. As with the avatar stimuli, it is possible that presence of a face image in Experiment 3 may have encouraged participants to pay more attention to static shape and structure information, which would usually be the most informative cues to identity. A focus on form cues would explain why matching performance for static PLDs was above chance in Experiment 3, but not Experiment 4. It may also explain why participants showed generally better performance for avatars in Experiment 4 compared to Experiment 3: in Experiment 4, there was no mismatch between form cues for the avatars (i.e., the faces were always the same shape), whereas in Experiment 3 the shape-normalised avatars contained different structures, shapes, textures, and features compared to the matching non-degraded images. However, participants were also better at matching PLDs in Experiment 4 when compared to Experiment 3, even though there was no mismatch in underlying face structure. Therefore, it is possible that the difference between the two experiments is due to overarching problems generalizing from one type of image to another, rather than specifically due to the focus on form information, or the mismatch of form cues in avatars. This problem may have arisen because the PLD and avatars stimuli contain different features and a different texture, compared to the non-degraded images. A non-stimulus-specific problem generalizing from one moving image to another could also account the findings in Experiments 1 and 2. It seems that participants are quite capable of matching identities from different moving sequences when the stimuli are in the same format (as in Experiment 4), but this task is far more difficult when participants are asked to simultaneously generalize across stimulus format and movement sequence (as in Experiments 1b and 2b).

5.3.4 Bias, Hits and Correct Rejections in Experiments 3 and 4

Unlike in Experiment 2, there was no consistent pattern of familiarity and bias across Experiments 3 and 4. In Experiment 3, famous faces resulted in a greater “different” bias, whereas in Experiment 4, there was no effect of familiarity on bias. This indicates that the source of the familiarity bias in Experiments 2a and 3 may be the fact that people could identify the initial image. Initially, it appears that the effect of familiarity on bias was inconsistent across Experiments 2 and 3. However, on
closer examination, both Experiment 2b and the PLD condition of Experiment 3 showed no effect of familiarity on bias. The effect of familiarity on bias in Experiment 3 was primarily driven by avatars – participants were significantly more biased towards a “different” response for unfamiliar than famous avatars. It is unclear why unfamiliar faces paired with avatars resulted in a “different” bias – it is possible that participants were simply less confident responding “same” to unfamiliar faces when structural cues were not available.

The only consistent bias pattern across Experiments 2, 3, and 4 was that participants were significantly biased towards a “same” response when the two images to be matched were perceptually similar – that is, when participants were shown two moving images. When one image was moving and the other static, participants were generally more likely to report that the stimuli were “different”, regardless of familiarity or stimulus type. Unlike in Experiment 2, the effect of perceptual similarity did not appear for S/S trials in Experiments 3 and 4. Once again, it is unclear why participants in Experiments 3 and 4 had different response strategies compared to those in Experiment 2.

5.3.5 Conclusions from Experiments 3 and 4

The experiments presented in this chapter confirm and extend on those presented in Chapter 4. In general, it appears that participants have trouble generalising from a non-degraded image to a degraded image, regardless of whether the degraded image in a PLD or shape-normalised avatar. However, participants in Experiment 3 performed better overall than those in Experiments 1 and 2, and participants in Experiment 4 performed better again. Allowing participants to view a non-degraded image in a matching task appear to hinder, rather than help performance.

There was a general matching advantage for famous (compared to unfamiliar) faces across both experiments in this chapter, but this advantage arose primarily due to good performance with static images. It is important to note, though, that the advantage was not solely due to low-level image similarity (at least in Experiment 4) – participants may be able to match famous faces based on characteristic static poses or expressions. When the effect of static recognition is removed, participants show a larger movement advantage for unfamiliar than famous faces when matching from PLD-to-PLD or avatar-to-avatar (Experiment 4).
The results suggest that participants can use different movement-based mechanisms in a flexible manner. Experiment 4 provided evidence that participants can match characteristic movement patterns when structural cues are redundant (as suggested by the supplemental information hypothesis, Roark et al., 2003). However, participants were significantly more accurate when matching PLDs than avatars. These findings suggest that the movement advantage observed in unfamiliar and famous faces was primarily derived from structure-from-motion, whereby movement is used to build a more robust three-dimensional representation of the face and head (the representation enhancement hypothesis, Roark et al., 2003). The next chapter extends on the finding that people can use characteristic movement patterns to complete matching tasks when the images are degraded and shape-normalised.
Chapter 6

The Role of Movement in Famous and Unfamiliar Face Matching

Which Movement Matters?
CHAPTER 6: THE ROLE OF MOVEMENT IN FAMOUS AND UNFAMILIAR FACE MATCHING: WHICH MOVEMENT MATTERS?

Chapters 4 and 5 addressed the questions: Is there a movement advantage for famous and unfamiliar faces; is there a difference in the size of the movement advantage for famous and unfamiliar faces; and what mechanisms (structure-from-motion or characteristic movement patterns) are responsible for the movement advantage in famous and unfamiliar face matching? There were several important findings. In Chapter 4, there was a movement advantage for unfamiliar faces, but it was only present when it was possible to match individual motion sequences. In Chapter 5, a movement advantage was found for both famous and unfamiliar faces when participants were asked to match from one PLD to another (Experiment 4). The results from Chapter 5 suggest that the movement advantage arose primarily because participants were comparing shape-based cues when matching the famous and unfamiliar faces. However, Chapter 5 also found evidence that participants could match faces based on characteristic movement patterns: for unfamiliar faces, participants could match two moving shape-normalised avatars significantly better than two static shape-normalised avatars, even though the moving clips showed different movement sequences. Famous faces presented as avatars showed no movement advantage, probably because the static avatars were matched unusually well.

The experiments in this chapter extend these findings from Chapter 5, by examining what type of movement contributes to characteristic movement patterns. Specifically, Experiments 5, 6, and 7 sought to examine whether participants can match characteristic movement patterns when only rigid or non-rigid movement is present (see section 6.1), and, if non-rigid movements do carry characteristic movement information, whether identity information is concentrated in different areas of the face (see section 6.2). As in Chapters 4 and 5, the experiments in this chapter also examined whether the use of different types of movement or areas of the face is influenced by familiarity with a person. However, Experiments 5, 6, and 7 used a new paradigm – a sorting task – to investigate the use of characteristic movement patterns in movement-based face matching.
6.1 The Type of Movement: Rigid and Non-rigid Movements in Face Recognition

As mentioned in Chapter 3, face and head movements can be divided into two broad categories: rigid movements and non-rigid movements. Rigid movements involve translations of the head – for example, nodding or shaking actions, or the change in viewpoint when someone walks around you (or you around them). On the other hand, non-rigid movements occur when the face temporarily changes or deforms in some way (e.g., when someone speaks, smiles or changes their eye gaze), but the overall shape or view of the head does not change. There is a large amount of evidence that both rigid and non-rigid movements can give rise to a movement advantage (to facilitate comparisons, Tables 1-4 in Chapter 3 list the types of movement used in various studies). However, it is possible that rigid and non-rigid movement contribute to face recognition or matching in different ways.

Characteristic movement patterns could be made up of both rigid head movements and non-rigid face movements (e.g., speech and expression). Likewise, social cues can be present in both rigid and non-rigid movements (e.g., head nodding and smiling both carry social information). By contrast, structure-from-motion cues are more likely to be facilitated by rigid movements of the head, because these allow more “views” of the face that can be extrapolated into 3D representations. In other words, rigid movement may give rise to a movement advantage via structure-from-motion, characteristic movement patterns or social signals, whereas non-rigid movement can only support recognition via characteristic motion patterns or social signals.

Studies using non-rigid stimuli (e.g., expressions, speech) have found a movement advantage for personally familiar (Lander et al., 2006; Rosenblum et al., 2007), experimentally familiar (Bruce et al., 2001, Experiment 3; Knappmeyer et al., 2003; Lander & Davies, 2007, Experiment 1; Pilz et al., 2006; Pilz et al., 2009), and unfamiliar faces (Rosenblum et al., 2002; Thornton & Kourtzi, 2002). Although the exact details of the studies vary, the overall findings suggest that non-rigid movement allows people to extract and identify/match characteristic movement patterns (e.g., Knappmeyer et al., 2003; Rosenblum et al., 2007; Lander & Davies, 2007). However, some studies have not used moving test images, meaning that
participants could not have matched movement patterns. This suggests that some of the benefits of non-rigid movement may occur because people pay more attention to moving than static faces (e.g., Pilz et al., 2006; Pilz et al., 2009; Thornton & Kourtzi, 2002).

Studies using rigid head movements (e.g., rotational movements) have found a movement advantage for learning or recognising experimentally familiar faces (Pike et al., 1997; Schiff et al., 1986), and for matching unfamiliar faces (Farivar et al., 2009). These findings support the idea that rigid movement helps face recognition via structure-from-motion information. However, none of these studies featured characteristic movement patterns (they used still faces rotating at a fixed rate), so, based on these results, it is unclear whether rigid movements can also contribute to face recognition via characteristic movement patterns.

To determine what type of movement is most helpful for face recognition, several studies have attempted to directly compare performance with rigid and non-rigid movements. Unfortunately, many of the results conflict with each other. Bruce and Valentine (1988) and Lander and Chuang (2005) asked participants to identify personally familiar faces using staged rigid or non-rigid movement (e.g., nodding/shaking rigid movements; smiling/frowning/speech non-rigid movements). Bruce and Valentine found that participants could name their friends more accurately from PLD stimuli displaying either kind of movement than from a single static frame. Furthermore, there was no significant difference in accuracy between rigid head movements and non-rigid face movements. On the other hand, Lander and Chuang found that participants were significantly better at identifying degraded faces (blurred, contrast- and brightness-adjusted) when they saw non-rigid movements (either speaking or smiling) than static frames, but there was no movement advantage for viewing faces moving rigidly (looking up and down). These studies suggest that we can recognise familiar faces based on non-rigid movements alone, but the role of rigid movement is less clear. It is possible that the choice of stimuli can explain the different results: participants viewing PLDs (Bruce & Valentine, 1988) may have used the rigid head movements to extract structural cues, whereas participants in Lander and Chuang’s (2005) study may have been able to obtain enough structural information from the static frames alone, or were unable to extract additional structure-from-motion information from the up-and-down head movements.
Christie and Bruce (1998) and Lander and Bruce (2003) also compared rigid and non-rigid movements, but they used experimentally familiar faces. Christie and Bruce tested the effect of movement for learning and test images displaying rigid (head nodding and shaking) and non-rigid (expressive) movements. They found that participants performed better overall when they were presented with rigid than non-rigid movements. Participants were also better at generalising across viewpoints when they were presented with rigid movement, compared to non-rigid movement. However, these results are qualified by the fact that Christie and Bruce did not find an overall movement advantage – in fact, their participants performed better when they were familiarised with static faces (although there was a marginal beneficial effect for testing with moving images). In contrast to Christie and Bruce, Lander and Bruce (2003) found a significant movement advantage for face learning. Faces learnt in movement (either rigid or non-rigid) were subsequently recognised more accurately than those learnt from a single static frame, but the benefit for rigid movement was erased when multiple static images were used as a control condition. It is important to note, though, that Lander and Bruce (2003) only tested recognition of static faces – participants could not use characteristic movement patterns to identify experimentally familiar faces, which suggests that social cues were driving the movement advantage for non-rigid movement (see Chapter 3, section 3.3.2.3 for further discussion).

It appears that any benefit of rigid movement for experimentally familiar faces arises primarily because of extra structure-from-motion or viewpoint information in rigid clips (e.g., Lander & Bruce, 2003; Pike et al., 1997) – this is also consistent with the reduced viewpoint dependence in Christie and Bruce (1998). These findings do not rule out the possibility that rigid movements can also contribute to characteristic movement patterns, but at this point, no studies using experimentally familiar faces have tested this hypothesis. On the other hand, the benefit on non-rigid movement for face learning probably arises due to a combination of characteristic movement patterns (e.g., Lander & Davies, 2007) and enhanced attention to social cues (e.g., Lander & Bruce, 2003).

Some of the most interesting studies that have compared rigid and non-rigid movement have used unfamiliar faces as stimuli. Hill and Johnston (2001) projected isolated rigid, non-rigid, or combined rigid and non-rigid movements onto shape-normalised avatars (similar to those used in Experiments 3 and 4). Unlike the
personally and experimentally familiar face studies, the movement sequences were natural movements, filmed during conversational speech (not staged or restricted movements). Hill and Johnston found that participants could sort the avatars into identity groups at above chance levels when the animations showed rigid or combined movement, but sorting was only marginally better than chance when the animations showed only non-rigid movement. Most importantly, Hill and Johnston’s results confirmed that rigid movement can act as a cue to identity, independently of structure-from-motion (participants would not have been able to extract structural information from the shape-normalised heads). Watson et al. (2005) also tested participants using shape-normalised avatars of unfamiliar faces. In line with Hill and Johnston, matching was best overall for combined (rigid and non-rigid) movement, and worst for non-rigid movement only (Experiment 3). However, Watson et al. found that non-rigid movement matching was relatively viewpoint-invariant, whereas performance dropped when participants were asked to match rigid movements from different viewing angles (but cf., Christie & Bruce, 1998).

Overall, then, the findings suggest that non-rigid movement is less important than rigid movement for unfamiliar face matching (Hill & Johnston, 2001; Watson et al., 2005), but as a face becomes more familiar, non-rigid movement becomes as important as, or even more important than rigid movement for face recognition (Bruce & Valentine, 1988; Lander & Chuang, 2005). Watson et al.’s findings may account for this shift in cue use. It is possible that the importance of non-rigid movement increases with familiarity, because non-rigid movement can act as a viewpoint-invariant cue that is useful when attempting to identify a familiar face from novel perspectives.

Alternatively, it is possible that characteristic rigid movements are important when matching familiar as well as unfamiliar faces (Bruce & Valentine, 1988; Hill & Johnston, 2001), but we require some level of familiarity to generalise across viewpoints (e.g., Christie & Bruce, 1998). As mentioned above, it is difficult to assess the role of characteristic rigid movements in familiar face recognition. Currently, no studies of familiar or experimentally familiar faces have used shape-normalised faces; this makes it impossible to determine the relative contribution of structure-from-motion and characteristic movement patterns. Furthermore, all of the familiar and experimentally familiar studies recounted above used staged movements, as opposed to the natural movement sequences used in Hill and
Johnston (2001) and Watson et al. (2005). Staged movements may have limited the presence of characteristic rigid movements (such as a person consistently raising their head at the end of a sentence), which may have reduced participants’ use of characteristic movement cues.

Finally, it is difficult to compare the use of rigid and non-rigid movements in familiar and unfamiliar face recognition because no studies to date have manipulated familiarity and the type of movement concurrently. Furthermore, no studies have examined the use of rigid and non-rigid movements using famous faces, probably because it is difficult to get spontaneous examples of pure rigid or non-rigid movement, or motion capture data from famous faces. However, the advent of new video tracking technology means that it is possible to extract the rigid and non-rigid movements from any video. Therefore, Experiments 5, 6, and 7 compared matching of famous and unfamiliar faces using rigid only, non-rigid only, or combined rigid and non-rigid movements. Like Hill and Johnston (2001) and Watson et al. (2005), the stimuli were shape-normalised, and depicted natural speech and expressive movements. Consequently, the results for unfamiliar faces were expected to mirror Hill and Johnston’s and Watson et al.’s: an advantage for combined and rigid movement compared to non-rigid movements. There was one caveat to this prediction: Neither Hill and Johnston or Watson et al. used a static control condition (they compared performance with upright and inverted faces instead), and it is possible that rigid head movements provide more static identity cues than non-rigid movements (e.g., Lander & Bruce, 2003). If this is the case, rigid movements were expected to lead to an overall matching advantage, but non-rigid movements were expected to lead to a larger movement advantage.

It was unclear whether famous faces would show the same pattern of results as unfamiliar faces. Lander and Chuang’s (2005) and Bruce and Valentine’s (1988) results indicate that non-rigid movement is at least as important as rigid movement for personally familiar faces, and it was expected that famous faces would show a similar pattern. Therefore, famous faces were expected to show a movement advantage for non-rigid movements, and possibly also for rigid movements. Based on previous chapters in this thesis, it was expected that famous faces would be matched better than unfamiliar faces overall, but unfamiliar faces would show a larger movement advantage than famous faces.
6.2 The Location of Movement Information in Faces: Eyes and Mouths

As reviewed in section 6.1, there is a large amount of evidence that non-rigid movement can give rise to a movement advantage. The results from several studies suggest that this occurs because people can extract and match characteristic movement patterns from non-rigid movement. For example, Lander et al. (2006) found that personally familiar faces were recognised more accurately from natural than morphed smiles – this suggests that the participants were familiar with the temporal characteristics of their friends’ expressions. Furthermore, participants can sort and match shape-normalised animations of unfamiliar faces based on non-rigid movements alone (Hill & Johnston, 2001; Watson et al., 2005). However, it is unclear whether characteristic movement patterns occur across (and are extracted from) the whole face, or whether movement-based face matching is based primarily on the movements of features such as the mouth or eyes. No research has directly compared the role of different facial areas in movement-based face recognition, but evidence from expression recognition and static face recognition studies suggests that both the eye and mouth regions may be important during face processing tasks.

In one of the earliest studies of facial PLDs, Bassili (1979) investigated the role of upper and lower face movements in expression recognition. Participants who viewed the lower half of the face were more accurate at identifying happiness and disgust, whereas the upper half of the face was better for anger and fear. In a similar experiment, Cunningham, Kleiner, Wallraven, and Bulthoff (2005) examined which areas of the face were sufficient for the recognition of nine conversational and emotional facial expressions – for example, agree/disagree, happy, sad, surprise – by selectively isolating rigid head movement, eye, eyebrow and mouth movements. They found that some expressions were conveyed primarily by rigid movement (e.g., agree/disagree), some relied on a combination of areas (e.g., disgust, confusion and clueless), and others still relied on specific facial areas. Happiness and surprise were conveyed almost totally by the mouth region, whereas confusion was recognised best when the eyebrows were present, and thinking was recognised best from the eye region (see also Nusseck et al., 2008). Overall, these results do not suggest that the movement of one facial area dominates expression recognition. However, they do suggest that individual areas of the face can carry a large amount of movement.
information, and it is possible that those same areas (i.e., the eyes and mouth) can support movement-based face recognition.

Experiments 5, 6, and 7 used a different procedure to examine the use of different regions in movement-based face recognition. Rather than isolating movement from one area (Cunningham et al., 2005), or restricting the movement of other areas of the face (Nusseck et al., 2008), the PLDs in Experiments 6 and 7 had extra detail added to the eyes or mouth. No studies of face recognition or matching have used this procedure before, and it was unclear whether participants would be able to use this extra information during a matching-based task. The following sections review evidence for the importance of the eyes and mouth during the perception and recognition of faces – both static and moving – to assess whether either region should confer an advantage during movement-based face matching.

### 6.2.1 The Eye Region

The eyes are an extremely important area in face processing tasks, from face detection to complex social judgements (Emery, 2000). Research on static face recognition suggests that the eye area is particularly useful during identification tasks. When identifying a face, people fixate on the eye region more often than the mouth (Emery, 2000; although familiarity with a face can modify scanning patterns, Barton, Radcliffe, Cherkasova, Edelman, & Intriligator, 2006). This fixation pattern may reflect the fact that the eyes carry useful cues to identity, particularly for familiar faces. Using the “bubbles” task, in which selected areas of the face are revealed by a sequence of masks, Vinette, Gosselin, and Schyns (2004) established that eye information is diagnostic of identity very quickly (47-94 ms) after an experimentally face is presented. Furthermore, people are more sensitive to spatial changes to the eye region than mouth region of experimentally familiar faces, but not unfamiliar faces (Brooks & Kemp, 2007; O’Donnell & Bruce, 2001). On top of their importance in identification tasks, eyes also contribute to a myriad of social judgements, particularly gaze direction. Gaze direction is an important social cue: not only does it provide information about the location a person is looking, but it is involved in complex social cognitive processes such as visual perspective-taking, joint attention, and theory of mind (Emery, 2000). Gaze direction can also influence face recognition and emotion processing – people are generally slower to recognise a face when the gaze is averted (Itier & Batty, 2009), and the direction of eye gaze can
influence the perceived intensity of an emotional expression (Willis, Palermo & Burke, 2011).

In general, then, the eyes and the way they move are extremely important for identification and social processing, and they attract a large amount of attention when we view a face. What is unclear, though, is whether eye movements can also provide identity information. As mentioned above, no research on movement-based face recognition has selectively manipulated movement in the eye region, and as such it is unclear whether people use characteristic patterns of gaze, blinking, or other eye movements (e.g., someone might frequently narrow or close their eyes) as a cue to identity. However, if characteristic movements from the eye region do contribute to face matching, it is possible that participants will be better at matching famous than unfamiliar faces when eye-movement information is included in the stimuli (Experiment 6). This effect may arise because participants are more sensitive to the eye region in familiar than unfamiliar faces (e.g., Brooks & Kemp, 2007, O’Donnell & Bruce, 2001), which could lead to familiarity with characteristic eye movement patterns.

6.2.2 The Mouth Region

Like the eyes, the mouth carries a large amount of social information. As mentioned above, many emotional and conversational expressions are conveyed via the movement of the mouth (e.g., Nusseck et al., 2008), and isolated visual speech information can improve auditory speech-in-noise perception by up to 10 dB (Rosenblum et al., 1996). While the mouth does not attract as much attention as the eyes, people still look at the mouth frequently when making identity decisions (Barton et al., 2006). However, most importantly for this study, the way a mouth moves can carry a large amount of identity information for both unfamiliar and familiar faces. Rosenblum et al. (2002) created PLDs using 30 fluorescent dots, at least 15 of which were placed around the lips, teeth, tongue, and jaw. Their stimuli therefore presented detailed articulatory information, with very little movement information from the remainder of the face, and no rigid movement at all. Despite these restrictions, participants could match unfamiliar faces at above-chance levels in a match-to-sample task. Participants were significantly better when matching original movement sequences than jumbled or staggered clips. Rosenblum et al. (2007) used similar stimuli, with movement information concentrated around the mouth and jaw
area, for familiar faces. Participants were remarkably accurate when naming friends from their mouth movements – five out of seven speakers were named above chance levels in a two-alternative forced choice task, and four out of seven were named above chance in a seven-alternative version of the task. As mentioned above, Lander et al. (2006) also provided evidence that natural mouth movements – in this case, natural smiles – improve identification performance for personally familiar faces, compared to morphed or speeded movements.

The fact that timing disruptions have such a detrimental effect on recognition tasks involving mouth movements (Lander et al., 2006; Rosenblum et al., 2002) suggests that mouth movements may be important for recognition due to the precise temporal cues they carry. However, speech and expression movements are somewhat constrained by production requirements – for example, changing the duration of an emotional expression can change its perceived intensity (Kamachi et al., 2001), and speeding visual speech information impairs perception even in trained lipreaders (Spehar, Tye-Murray, & Sommers, 2004). Therefore, it is unclear whether people should show an increased movement advantage for mouth movements as well as eye movements, which may have other social cues attached, but have fewer constraints.

Like eye movements, familiarity with a face should increase familiarity with mouth movements – therefore, it was expected that adding detailed articulatory information to the PLDs (Experiment 7) would improve performance for famous faces more than for unfamiliar faces. However, both famous and unfamiliar faces were expected to benefit from the addition of detailed mouth information (Rosenblum et al., 2002; Rosenblum et al., 2007).

### 6.3 The Sorting Task

Experiments 5, 6, and 7 used a different task to Experiments 1-4: participants were asked to sort video and still clips into groups based on identity. A sorting task was used by Hill and Johnston (2001) to compare rigid and non-rigid movements in unfamiliar face matching. The sorting task has several advantages over the match-to-sample and same/different tasks used in Chapters 4 and 5. First, participants could watch the clips as many times as they wished and complete the task at their own pace, leading to higher confidence and better overall performance; second, the sorting task offered extra dependent variables for analysis (see section 6.4, General...
Methods); and third, it was a shorter, more engaging task than the other matching tasks, again conducive to better performance. As in Experiment 4, participants were asked to match faces based purely on shape-normalised stimuli (they did not have a non-degraded video or still for comparison). Half of the underlying faces were unfamiliar, half were famous, but participants were not informed in advance that any of the faces were famous. Consequently, any effects of familiarity were unlikely to arise from increased attention to the famous face videos.

To confirm that participants could sort faces based on identity, a pilot sorting study was run that compared shape-normalised PLDs, shape-normalised PLDs with eyes (referred to as PLDe) and shape-normalised heads based on those used by Hill and Johnston (2001). Subsequently, Experiment 5 presented participants with shape-normalised PLDs that were similar to those used in Chapters 4 and 5. The PLDs showed the movement of the eyebrows and lips, but did not show any detailed eye movements (in fact, no eyes were shown) or the movement of the teeth and tongue. This experiment acted as a baseline for subsequent experiments, while comparing rigid and non-rigid movement for famous and unfamiliar faces. Experiment 6 examined the benefit of including eye movements: the stimuli were the same PLDs as Experiment 5, but included animated eyes that showed gaze direction and extent that the eyes were open or closed. Experiment 7 examined the benefit of including detailed mouth movements: once again, the stimuli were the same PLDs as Experiment 5, but included the teeth and tongue from the original (greyscale) video. As mentioned above, it was unclear whether one facial area would be more useful than the other, but it was expected that both eyes and mouths would improve matching performance compared to the baseline shape-normalised PLDs.

### 6.4 Sorting Task General Method

#### 6.4.1 Stimuli and Materials

The video clips used to create the stimuli in this chapter were identical in Experiments 5, 6, and 7. As in Experiment 1-4 (Chapters 4 and 5), Experiments 5, 6, and 7 used videos of highly familiar (famous) and unfamiliar faces obtained from the online content of two talk shows. The clips showed each person speaking to the camera in an interview situation. All of the people chosen were male, to prevent participants completing the task based on gender, rather than identity. Each video
Chapter 6: Which Movement Matters?

was eight seconds long, and showed the face from approximately the same angle and distance from the camera. Unlike Experiments 1-4, only four famous and four unfamiliar faces were used in Experiments 5, 6, and 7 (i.e., two famous and two unfamiliar faces that were used in Experiments 1-4 were excluded). The four famous and unfamiliar faces were selected based on the distinctiveness and amount of movement ratings (see Chapter 4, Pilot study 1; for further details). Since the sorting task involved participants viewing and comparing all four famous or unfamiliar people, the clips for each level of familiarity were chosen to be as close in distinctiveness and amount of movement ratings as was practicable – this meant that Experiments 5, 6, and 7 used different clips (although sourced from the same television programs) to those used in Experiments 1-4. The mean familiarity, distinctiveness and amount of movement ratings for each clip in Experiments 5, 6, and 7 are shown in Table 12.

The distinctiveness and amount of movement ratings for each clip were subjected to separate one-way ANOVAs with familiarity as the between-subjects variable. Results show that mean distinctiveness and amount of rating movements did not differ for famous and unfamiliar faces: distinctiveness, $F(1,6) = .001, p = .976$; amount of movement, $F(1,6) = .62, p = .463$.

Shape-normalised PLD stimuli were created using the same custom written tracking program that was used to create the avatars in Chapter 5 (Saragih et al., 2010, 2011). The program tracks 66 points on the face, and projects the movement of those points onto a synthetic face with uniform shape and texture. The program can animate the synthetic faces with rigid head movements, non-rigid face and eye movements, or both movements together. Examples of the shape-normalised PLDs used in Experiments 5, 6, and 7 are included in Appendix A. To create the videos for the sorting tasks, each 8 s video was cut into four segments of 2 s each. A frame from the centre of each 2 s clip served as the static control condition. All shape-normalised PLDs were presented at 25 frames per second. The videos and the static control frames both measured 960 x 540 pixels. The experiments were run using a custom-written program on Lenovo T500 laptop computers, running Windows XP, with screen resolution set to 1280 x 800 pixels.
Table 12: 
Mean ratings for the famous and unfamiliar face videos used in Experiments 5, 6, and 7 (Chapter 6).

<table>
<thead>
<tr>
<th>Measure</th>
<th>Russell Crowe</th>
<th>Kyle Sandilands</th>
<th>Ben Stiller</th>
<th>Jerry Seinfeld</th>
<th>Unfam 3</th>
<th>Unfam 4</th>
<th>Unfam 5</th>
<th>Unfam 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Familiarity</td>
<td>6.78</td>
<td>6.65</td>
<td>6.56</td>
<td>6.35</td>
<td>1.43</td>
<td>1.43</td>
<td>1.35</td>
<td>1.09</td>
</tr>
<tr>
<td>SD</td>
<td>0.67</td>
<td>1.07</td>
<td>0.99</td>
<td>1.52</td>
<td>0.79</td>
<td>0.84</td>
<td>0.93</td>
<td>0.29</td>
</tr>
<tr>
<td>Distinctiveness</td>
<td>3.10</td>
<td>3.30</td>
<td>4.20</td>
<td>4.30</td>
<td>4.06</td>
<td>4.48</td>
<td>3.33</td>
<td>3.60</td>
</tr>
<tr>
<td>SD</td>
<td>1.37</td>
<td>1.49</td>
<td>1.40</td>
<td>1.42</td>
<td>1.74</td>
<td>1.86</td>
<td>1.80</td>
<td>1.50</td>
</tr>
<tr>
<td>Amount</td>
<td>4.64</td>
<td>3.09</td>
<td>5.54</td>
<td>4.54</td>
<td>6.18</td>
<td>6.36</td>
<td>3.90</td>
<td>4.54</td>
</tr>
<tr>
<td>SD</td>
<td>0.81</td>
<td>1.22</td>
<td>1.13</td>
<td>1.04</td>
<td>1.22</td>
<td>0.95</td>
<td>1.37</td>
<td>1.81</td>
</tr>
</tbody>
</table>

*Note:* Familiarity ratings were averaged across all three clips of each individual, as in the pilot study (see Chapter 4). Distinctiveness and amount of movement ratings are reported for the clips used in this chapter only. To facilitate comparisons, the unfamiliar faces are numbered the same as in Table 5.

6.4.2 Design and Procedure

Experiments 5, 6 and 7 used a mixed design, with within-subjects factors of familiarity (famous and unfamiliar) and presentation (moving and static), and the between-subjects variable of movement type (combined; rigid; non-rigid). Participants were randomly assigned to each movement type.

Each participant completed four sorting tasks (famous/moving; famous/static; unfamiliar/moving; unfamiliar/static). In each task, they were presented with 16 boxes on-screen (four boxes each x four people, see Figure 16a), and were asked to sort them into four equally sized groups based on identity. When they clicked on a box, a 2 s video or static image of a shape-normalised PLD appeared in the centre of the screen (see Figure 16b). The image disappeared after 2 s. Participants could then drag the box into a numbered column on the other side of the screen (see Figure 16c). Boxes in the same column indicated the same identity, but there was no restriction as to which identity had to be sorted into which column. Participants were advised that they could watch the clips as many times as they wanted, but they must watch each clip at least once before sorting it into a group. Participants were told that
they could change their mind and re-sort the boxes into different groups as many times as necessary until they were happy that each group contained four clips of the same person.

Each task had a 10 min time limit, and participants could monitor the time on their screen. When they were finished, participants clicked a button to indicate this and stop the timer. The experimenter then started them on the next task. The order of the four sorting tasks was counterbalanced across participants, with the constraint that participants never completed two moving or two static tasks sequentially.

Participants were not informed that some of the faces were famous until after the sorting tasks. Following the sorting tasks, each participant was shown a 2 s non-degraded video of the eight people in the sorting task. They were asked to rate the video for familiarity on a scale of 1-7 (1 = completely unfamiliar; 7 = extremely familiar), and to name the person (or provide other unambiguous identity information) if they were famous.

![Figure 16: An illustration of the sorting task. a) A partially completed sorting task. Participants sorted 16 numbered boxes into four columns according to identity. b) When participants double-clicked on a box, a PLD would appear in the centre of the screen for 2 s. c) Participants dragged the box to a column to complete the task.]
Participants’ data were excluded from analysis if they did not rate at least three out of the four famous faces 5 or higher or if they could not name or unambiguously identify at least three of the four famous faces and if they did not rate all of the unfamiliar faces 3 or lower.

6.4.3 Analysis
In each experiment, participants’ performance was analysed using four measures: accuracy, average number of times participants viewed each clip (views), time to complete the task (time), and a combined measure of accuracy divided by the number of views (accuracy per view). Accuracy was calculated based on how many times each image was placed in a group with another image of the same person. For example, if a group contained 3 videos of Person A and 1 video of Person B, the score for the group would be 6 – each of the three Person A clips was matched with another Person A clip twice. This resulted in a maximum score of 48 (4 people x 4 images x 3 other possible matches) and a minimum score of 0 (Hill & Johnston, 2001). Chance was determined by randomly allocating sixteen images to groups 100,000 times. In Experiments 5, 6, and 7, chance performance was 9.6. The accuracy score was then divided by the views score to create the accuracy per view measure. Accuracy per view measure is a new measure – previous studies (e.g., Hill & Johnston, 2001) have only measured accuracy, without taking into account the amount of exposure participants have to each clip. The views measure and the accuracy per view measure were created to control for the possibility that participants were more accurate in some conditions simply because they watched the clips more often.

In each experiment, accuracy scores were compared to chance using 2-tailed t-tests. Each measure was then subjected to a separate three-way mixed-design ANOVA (2 x familiarity; 2 x presentation – moving or static; 3 x movement type – combined, rigid, or non-rigid). Bonferroni corrections were applied to pairwise comparisons throughout all experiments.

6.5 Sorting Pilot Study
The pilot study was conducted to determine whether participants could complete a sorting task using the PLD stimuli, and to choose an appropriate time
limit for subsequent experiments. It also provided initial data on the sorting task, using the computer-tracked stimuli. The methods were exactly as stated in the general methods, except as noted below.

6.5.1 Methods of Sorting Pilot Study

6.5.1.1 Participants. Thirty undergraduate students (17 female, $M$ age: 21.4 years, range: 18 to 40) from the University of Western Sydney took part in the study in return for course credit. Participants reported normal or corrected-to-normal vision. Three participants’ data was excluded from analysis due to failure to pass the familiarity check.

6.5.1.2 Stimuli and materials. Three different stimulus sets were created for the pilot experiment. All stimulus sets used identical movement sequences and the same animation program, and all stimuli showed combined rigid and non-rigid movement. One stimulus set consisted of avatar faces, with the same head shape, features and texture as those used by Hill and Johnston (2001) and Watson et al. (2005); the second set consisted of a shape-normalised PLD with the same distribution of points as used in Chapters 4 and 5\(^{10}\); and the final set consisted of the same PLDs with animated eyes (PLDe). An example of each stimulus, along with the PLDs with mouths (PLDm) used in Experiment 7, is shown in Figure 17.

6.5.1.3 Design and procedure. In the pilot study, there were two within-subjects factors (familiarity: famous or unfamiliar; presentation: moving or static), and one between-subjects factor (stimulus, with 10 participants per group). Participants were given 15 mins to complete each of the four sorting tasks.

\(^{10}\) It is important to note that these were not the exact same stimuli as used in Chapters 4 and 5. The current stimuli had a uniform shape and were generated using an automated tracking program, rather than being hand-tracked in a video-editing program. The majority of stimuli were also created from different movement sequences for the current experiments, as is detailed in section 6.4.
Figure 17: An example of the stimuli used in the pilot sorting task and Experiments 5, 6, and 7: a) a shape-normalised avatar (based on Hill & Johnston, 2001); b) a shape-normalised point-light display, as used in Experiment 5; c) a shape-normalised point-light display with eyes, as used in Experiment 6; d) a shape-normalised point-light display with teeth and tongue, as used in Experiment 7.

6.5.2 Results and Discussion of the Pilot Sorting Task

6.5.2.1 Accuracy. T-tests on the accuracy figures showed that avatars were sorted at above-chance levels in all conditions, \( p < .025 \), while PLDs and PLDe were sorted at above chance levels in all conditions except unfamiliar static PLDs/PLDe, \( p > .09 \), all other \( ps < .05 \).

The ANOVA on accuracy revealed a main effect of familiarity, \( F(1,24) = 4.59, p = .042, \eta^2_p = .15 \). As predicted, famous faces, \( M = 16.63 \), were sorted more accurately than unfamiliar faces, \( M = 14.30 \), regardless of whether they were moving or static. No other main effects or interactions, including those involving movement, were significant, all \( ps > .3 \). Planned pairwise comparisons revealed no movement advantage for famous or unfamiliar faces, and no movement advantage for any individual stimulus type, all \( ps > .1 \).

The results from the pilot study confirm that people can sort these degraded faces quite well, regardless of the type of stimulus used. Thus, in this case avatars, PLD and PLDe provided similar information.
6.5.2.2 Views. On average, participants watched each clip 4.26 times while completing the sorting task. The ANOVA on average number of views per clip revealed a main effect of presentation, $F(1,24) = 11.30, p = .002, \eta^2_p = .33$. Moving faces, $M = 3.90$, were viewed fewer times on average than static faces, $M = 4.62$. No other main effects or interactions were significant, all $p$s > .1.

6.5.2.3 Time. The average time to complete each sorting task was 5.66 mins. The ANOVA revealed no significant main effects or interactions for time to complete each sorting task, all $p$’s > .2. Based on the average time to complete the task, participants in subsequent experiments were restricted to a maximum of 10 mins per sorting task – no participants in the pilot study exceeded 10 mins in any task.

6.5.2.4 Accuracy per view. Like the accuracy analysis, the ANOVA on accuracy per view revealed a main effect of familiarity, $F(1,24) = 5.23, p = .031, \eta^2_p = .18$. Once again, famous faces, $M = 5.11$, were sorted more accurately per view than unfamiliar faces, $M = 3.90$. No other main effects or interactions were significant, all $p$s > .1. However, planned pairwise comparisons revealed that there was a movement advantage for unfamiliar faces, $p = .028$, but not for famous faces, $p = .882$, and there was no movement advantage for any stimulus type, $p$s > .01.

6.5.3 Conclusions from the Pilot Sorting Task

There were two main findings from the pilot sorting task: first, as predicted, there was a significant advantage for sorting famous faces, in both the accuracy and accuracy-per-view analyses and regardless of whether they were moving or static. This finding confirms that the results from Experiment 4 (and, to a lesser extent, Experiments 1-3) were not due to abnormal static frames, since new clips and static frames were chosen for this experiment. Famous faces may simply be easier to match from static frames – once again, it is unclear whether this is because participants are better at matching characteristic static poses for famous faces, or whether famous individuals are more stereotyped in their movements and poses. This question will be addressed further in Chapter 7, which studies personally familiar faces.

The second important result was the significant effect of presentation format (moving or static) on the number of times participants viewed the clips, combined
with the lack of movement advantage in the accuracy analysis, in other words, participants watch the moving clips fewer times than the static clips, but still achieve the same level of accuracy. These findings indicate that movement may not simply improve matching performance, but may make it faster or easier to match faces when they are degraded.

The pilot study results also extend the work presented in Chapter 5. The lack of effect of stimulus type indicates that the advantage for PLDs over avatars in Experiment 4 was probably due to the shape-based cues in the PLDs, rather than the presence of a facial form or eyes in the avatar stimuli. Nonetheless, PLDs were sorted well in the present experiment, which suggests they carry sufficient identity cues to support matching. PLDs were chosen as the stimuli for the subsequent baseline experiment (Experiment 5) because creating an avatar without eyes (e.g., by removing the pupils, setting them to a constant point, or occluding them with “sunglasses”) resulted in an odd-looking stimulus that may have distracted participants from the task.

6.6 Experiment 5: Rigid and Non-Rigid Movement in Shape-normalised Point-light-displays

As mentioned in section 6.3, Experiment 5 acted as a baseline for Experiments 6 and 7. Using PLDs with no eyes or mouths, Experiment 5 assessed the contribution of rigid and non-rigid movements to famous and unfamiliar face sorting. Once again, the methods were identical to the general methods except as indicated below.

6.6.1 Methods of Experiment 5

6.6.1.1 Participants. Sixty-three undergraduate students (49 female, $M_{age}: 21.8$ years, range: 18 to 48) from the University of Western Sydney took part in the study in return for course credit. Participants reported normal or corrected-to-normal vision. Three participants’ data was not included due to failure to complete the experiment. A further nine participants’ data was excluded from analysis due to failure to pass the familiarity check.

6.6.1.2 Stimuli and Materials. The stimuli in this experiment were shape-normalised PLDs, created using the custom animation program described in the
general methods. Each PLD consisted of 25 white dots, positioned in consistent locations on the face. (see Figure 17b). The images were normalised to a single face and head shape, so participants could not use the location or spacing of the markers as a cue to identity. Separate groups of participants saw rigid head movements \( (n = 17) \), non-rigid face movements \( (n = 17) \) or combined rigid and non-rigid movements \( (n = 17) \).

6.6.2 Results and Discussion of Experiment 5

6.6.2.1 Accuracy. Accuracy for each condition of Experiment 5 is shown in Table 13. Participants who viewed clips with combined movement were above chance in all conditions, \( p \)'s < .05. Participants in the non-rigid movement condition sorted famous faces (both moving and static) and unfamiliar moving faces at chance levels, \( p \)'s > .05. Static unfamiliar faces in the non-rigid condition were sorted better than chance, \( p < .0005 \). Participants who viewed clips with rigid head movement sorted famous faces better than chance, moving clips: \( p = .015 \); static clips: \( p = .002 \), but unfamiliar faces were sorted at chance levels, \( p \)'s > .05.

The ANOVA on accuracy revealed no main effects, familiarity and presentation: \( F(1,48) < 1 \); movement type: \( F(2,48) = 2.80, p = .071, \eta_p^2 = .10 \). However, there was a highly significant interaction between familiarity and movement type, \( F(2,48) = 5.43, p = .007, \eta_p^2 = .18 \). For famous faces, rigid movement was sorted more accurately than non-rigid movement, \( p = .024 \), whereas for unfamiliar faces, combined movement was sorted more accurately than rigid movement, \( p = .019 \). The familiarity by movement type interaction was qualified by a three-way interaction between familiarity, presentation and movement type, \( F(2,48) = 5.03, p = .010, \eta_p^2 = .17 \). Pairwise comparisons showed an advantage for rigid over non-rigid clips for famous faces, but the effect was isolated to static clips – that is, static frames extracted from the rigid or non-rigid videos – moving: \( p = .605 \); static: \( p = .020 \). This indicates that, for famous faces, the head pose information contained in the rigid static frames was more useful for sorting than the expressions information contained in the non-rigid static frames, but this benefit disappeared when movement was introduced to the clips (probably because accuracy in the non-rigid condition increased). For unfamiliar faces, combined movement was sorted more accurately than both rigid, \( p = .003 \), and non-rigid movement, \( p = .022 \), but only when the clips were shown moving, static \( ps > .1 \). This pattern of results is quite
different to the famous faces – for unfamiliar faces, participants could not use head
pose or expression information to sort the static frames, but when movement was
introduced to the clips the combination of rigid and non-rigid movement cues helped
sorting significantly.

Pairwise comparisons were also carried out to examine whether moving clips
were sorted better than static clips in any of the conditions. None of the individual
movement types showed a movement advantage, \( p > .1 \). There was no overall
movement advantage for either famous or unfamiliar faces, \( p > .4 \), and famous
faces were sorted equally well from moving and static clips regardless of movement
type, \( p > .1 \).

Table 13:
\textit{Mean Accuracy and Views for Experiment 5}

<table>
<thead>
<tr>
<th>Measure</th>
<th>Combined</th>
<th>Non-Rigid</th>
<th>Rigid</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Moving</td>
<td>Static</td>
<td>Moving</td>
</tr>
<tr>
<td>Accuracy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Famous</td>
<td>12.94*</td>
<td>14.24*</td>
<td>11.41</td>
</tr>
<tr>
<td></td>
<td>(5.01)</td>
<td>(5.70)</td>
<td>(3.72)</td>
</tr>
<tr>
<td>Unfamiliar</td>
<td>15.41**</td>
<td>12.12*</td>
<td>11.18</td>
</tr>
<tr>
<td></td>
<td>(4.84)</td>
<td>(3.97)</td>
<td>(4.48)</td>
</tr>
<tr>
<td>Views</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Famous</td>
<td>2.73</td>
<td>3.11</td>
<td>3.51</td>
</tr>
<tr>
<td></td>
<td>(1.35)</td>
<td>(1.23)</td>
<td>(1.41)</td>
</tr>
<tr>
<td>Unfamiliar</td>
<td>2.80</td>
<td>2.96</td>
<td>3.32</td>
</tr>
<tr>
<td></td>
<td>(1.41)</td>
<td>(1.60)</td>
<td>(1.85)</td>
</tr>
</tbody>
</table>

\textit{Note:} Chance level accuracy was 9.6. Standard deviations are shown in brackets.
*\( p < .05 \) ** \( p < .0005 \)
Unfamiliar faces showed a movement advantage for combined movement, \( p = .019 \). There was a movement disadvantage for unfamiliar faces displaying non-rigid movement, \( p = .006 \).

The accuracy results indicate that participants were able to sort shape-normalised PLDs relatively well, but there is very little evidence that they were using movement cues to do so. There was no overall movement advantage, regardless of whether the faces were famous or what type of movement was displayed, and only a single condition – unfamiliar faces with combined movement – was sorted better from moving than from static clips.

### 6.6.2.2 Views

The average number of views per clip is also shown in Table 13. The ANOVA on average number of views per clip revealed no significant main effects or interactions (\( p \)’s > .1). There was no evidence that participants were compensating for the lack of movement information by watching the clips more frequently, as they did in the pilot study.

### 6.6.2.3 Time

The ANOVA on completion time revealed no significant main effects or interactions (\( ps > .1 \)).

### 6.6.2.4 Accuracy per view

The accuracy per view results are displayed in Figure 18. The accuracy per view ANOVA mirrored many of the accuracy results. There were no significant main effects for familiarity or presentation, both \( F(1,48) < 1 \). However, there was a main effect of movement, \( F(2,48) = 3.71, p = .032, \eta_p^2 = .13 \), and a significant interaction between movement and familiarity, \( F(2,48) = 7.34, p = .002, \eta_p^2 = .23 \). Pairwise comparisons showed that combined movement was sorted significantly better than rigid movement overall, \( p = .044 \). When the results were broken down by familiarity, combined movement was sorted better than rigid movement for unfamiliar faces, \( p = .006 \); as in the accuracy analysis, the effect was present for moving clips, \( p = .002 \), but not static clips, \( p = .272 \). Combined movement was sorted better than non-rigid movement for famous faces, \( p = .030 \); unlike accuracy, this effect was not present when moving and static clips were
analysed separately, $ps > .1$. No other pairwise comparisons were significant, $ps > .05$.

It is interesting to note that the relationship between familiarity and movement type was distinctly different for accuracy and accuracy-per-view, particularly for famous faces (accuracy: advantage for rigid over non-rigid movement, static frames only; accuracy-per-view: advantage for combined over non-rigid movement). It is possible that participants in the rigid movement condition simply watched the clips more often, although not enough to result in a significant effect in the views analysis.

Planned pairwise comparisons were carried out to look for a movement advantage in each condition. There was no movement advantage for any individual movement type, $p's > .08$. Neither famous nor unfamiliar faces showed an overall movement advantage, $p's > .7$, and famous faces showed no movement advantage for any movement type, $p's > .1$. Unfamiliar faces displaying combined movement were sorted significantly better from moving than static clips, $p = .034$.

Figure 18: Accuracy-per-view results for different movement types in Experiment 5. Error bars represent +/- 1 standard error of the mean.
The accuracy per view results support the conclusions from the accuracy analysis: participants could sort shape-normalised PLDs well, but there was no overall benefit of movement. Despite the lack of a movement advantage, the type of movement (which also carried through to the type of static clip, since the static clips were extracted from the moving sequences) had a significant impact on how well participants could sort the faces. Famous faces were sorted consistently poorly from non-rigid clips. This result extends on previous research using the sorting task (Hill & Johnston, 2001), which found an advantage for rigid over non-rigid movement using unfamiliar faces. On the other hand, unfamiliar faces showed a consistent advantage for combined movement clips. It is difficult to say why these results do not replicate Hill and Johnston’s (2001) findings, which also used shape-normalised, unfamiliar faces, but found an advantage for rigid movement when sorting unfamiliar faces. One possible explanation is that Hill and Johnston included eyes in their avatars, which could have distracted participants from attending to identity in the non-rigid condition, and even in the combined movement condition. This was investigated further in Experiment 6.

6.6.3 Conclusions from Experiment 5

Overall, participants in Experiment 5 showed a very different pattern of responding from those in the pilot study. There was no effect of familiarity, and no effect of movement on the average number of views per clip. This is not entirely surprising – the PLD stimuli were very basic, and provided no individual structural cues (e.g., the shape of the face). Furthermore, even in the pilot study there was no movement advantage for PLDs individually. However, unfamiliar faces did show a significant movement advantage for combined rigid and non-rigid movements, suggesting that both rigid and non-rigid movement cues carry identity information.

The effect of movement type varied significantly based on familiarity, but it is unclear how much of that effect arose from static cues. The advantage for rigid static information for famous faces in the accuracy analysis indicates that head pose information, rather than expression information, might be an important cue in famous face matching, which may begin to explain the high static accuracy levels for famous faces in Experiments 1a, 2a, and 4.

Based on the results from Experiment 5, it appears that the movement advantage for both famous and unfamiliar faces may require information from more
facial areas than were presented in this study. As discussed above, there are many reasons why the eye and mouth areas might offer a movement advantage. Consequently, Experiment 6 tested the effect of adding eye movement information, and Experiment 7 tested the effect of adding the movements of the teeth and tongue to the basic PLDs. As in the Pilot study and Experiment 5, the methods for each experiment are identical to the general methods except as noted.

6.7 Experiment 6: Rigid and Non-Rigid Movement in Shape-normalised Point-light-displays with Eyes

Experiment 6 was designed to investigate whether sorting performance is affected by showing participants eye movements.

6.7.1 Methods of Experiment 6

6.7.1.1 Participants. Seventy-one undergraduate students (60 female, \( M \) age: 21.9 years, range: 17 to 52) from the University of Western Sydney took part in the study in return for course credit. Participants reported normal or corrected-to-normal vision. Ten participants’ data was not included due to failure to understand the instructions or complete the experiment within the time limits. A further ten participants’ data was excluded from analysis due to failure to pass the familiarity check.

6.7.1.2 Stimuli and Materials. The stimuli in this experiment were identical to those in Experiment 5, except that they contained eyes. The pupil location and opening and closing of the eyes was tracked and projected onto synthetic, shape-normalised eyes in the original PLDs (see Figure 17c). Separate groups of participants saw rigid head movements \((n = 17)\), non-rigid face movements \((n = 17)\) or combined rigid and non-rigid movements \((n = 17)\).

6.7.2 Results and Discussion of Experiment 6

6.7.2.1 Accuracy. Accuracy for Experiment 6 is shown in Table 14. Participants who viewed clips with combined or rigid movement were above chance in all conditions (all \( ps < .05 \)). Performance in the rigid condition is a distinct improvement on Experiment 5, where participants were at chance level for both
unfamiliar face tasks. Participants in the non-rigid movement condition sorted famous faces (both moving and static) at chance levels, \( p_s > .05 \), although there was a trend for moving famous faces to be sorted better than chance, \( p = .061 \). Unfamiliar faces showing non-rigid movement were sorted better than chance, \( p < .0005 \).

The ANOVA on accuracy showed a main effect of familiarity, \( F(1,48) = 5.08, p = .029, \eta^2_p = .10 \). On average, unfamiliar faces were sorted more accurately than famous faces. There was also a main effect of movement type, \( F(2,48) = 3.81, p = .29, \eta^2_p = .14 \), and a significant familiarity by movement type interaction, \( F(2,48) = 5.15, p = .009, \eta^2_p = .18 \). Overall, combined movement was sorted significantly more accurately than rigid movement, \( p = .024 \). For unfamiliar faces, non-rigid movement was sorted more accurately than rigid movement, \( p = .004 \), and there was a trend for combined movement to be sorted more accurately than rigid movement, \( p = .052 \). Combined and non-rigid movement were sorted equally well, \( p = 1 \).

<table>
<thead>
<tr>
<th>Measure</th>
<th>Combined Moving</th>
<th>Static</th>
<th>Non-Rigid Moving</th>
<th>Static</th>
<th>Rigid Moving</th>
<th>Static</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Famous</td>
<td>16.59**</td>
<td>16.59*</td>
<td>13.29</td>
<td>10.59</td>
<td>15.29*</td>
<td>13.06*</td>
</tr>
<tr>
<td></td>
<td>(6.32)</td>
<td>(8.88)</td>
<td>(7.55)</td>
<td>(2.72)</td>
<td>(8.15)</td>
<td>(5.79)</td>
</tr>
<tr>
<td>Unfamiliar</td>
<td>16.82**</td>
<td>18.00**</td>
<td>18.71**</td>
<td>19.18**</td>
<td>12.00*</td>
<td>14.82*</td>
</tr>
<tr>
<td></td>
<td>(6.25)</td>
<td>(5.92)</td>
<td>(6.78)</td>
<td>(7.32)</td>
<td>(4.12)</td>
<td>(5.92)</td>
</tr>
<tr>
<td>Views</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Famous</td>
<td>3.50</td>
<td>4.39</td>
<td>4.34</td>
<td>4.81</td>
<td>3.32</td>
<td>3.52</td>
</tr>
<tr>
<td></td>
<td>(1.19)</td>
<td>(1.38)</td>
<td>(1.64)</td>
<td>(1.47)</td>
<td>(1.45)</td>
<td>(1.59)</td>
</tr>
<tr>
<td>Unfamiliar</td>
<td>2.96</td>
<td>4.07</td>
<td>3.73</td>
<td>4.86</td>
<td>3.13</td>
<td>3.26</td>
</tr>
<tr>
<td></td>
<td>(0.93)</td>
<td>(1.65)</td>
<td>(1.16)</td>
<td>(2.18)</td>
<td>(1.32)</td>
<td>(0.94)</td>
</tr>
</tbody>
</table>

*Note: Chance level accuracy was 9.6. Standard deviations are shown in brackets. *\( p < .05 \) ** \( p < .0005 \)
For famous faces, a different pattern emerged: combined movement was sorted more accurately than non-rigid movement, $p = .043$, but equally as well as rigid movement, $p = .579$.

There was no significant main effect of presentation (moving versus static), $F(1,48) < 1$, but the interaction between familiarity and presentation just reached significance, $F(1,48) = 4.06, p = .050, \eta_p^2 = .08$, reflecting the fact that, numerically, famous faces were sorted better from moving than static clips, whereas unfamiliar faces showed the opposite pattern (neither comparison reached significance, $ps > .1$).

Pairwise comparisons revealed no movement advantage in any individual condition, and no other interactions were significant, all $p$’s $> .1$.

**6.7.2.2 Views.** Average views for Experiment 6 are also shown in Table 14. The ANOVA on average number of views per clip revealed a significant main effects of presentation, $F(1,48) = 22.08, p < .0005, \eta_p^2 = .315$, and movement type, $F(2,48) = 5.33, p = .008, \eta_p^2 = .182$. Static clips were viewed significantly more times on average than moving clips, and non-rigid clips were viewed significantly more times on average than rigid clips, $p = .007$. There was also a significant interaction between presentation and movement, $F(2,48) = 3.22, p = .049, \eta_p^2 = .12$. Combined and non-rigid movement clips were viewed significantly more times in the static condition than the moving condition, combined: $p < .0005$; non-rigid: $p = .002$, but there was no effect of presentation for rigid movement clips, $p = .489$. The main effect of familiarity was not significant, $F(1,48) = 2.83, p = .099, \eta_p^2 = .06$, confirming that participants did not perform better at the unfamiliar clips simply because they watched them more times before sorting them. No other interactions were significant, $p$’s $> .1$.

**6.7.2.3 Time.** The ANOVA revealed no significant main effects or interactions for time to complete each sorting task, all $p$’s $> .1$.

**6.7.2.4 Accuracy per view.** Accuracy per view results for Experiment 6 are shown in Figure 19. All three main effects were significant in the accuracy per view ANOVA. Participants were better at sorting unfamiliar than famous faces, $F(1,48) = 5.08, p = .029, \eta_p^2 = .10$, and better at sorting moving than static clips, $F(1,48) =$
Chapter 6: Which Movement Matters?

4.24, \(p = .045, \eta_p^2 = .08\). Pairwise comparisons on the main effect of movement \((F(2, 48) = 3.27, p = .047, \eta_p^2 = .12)\) showed that combined movement was sorted better than non-rigid movement, \(p = .044\). There was also a significant interaction between familiarity and movement type, \(F(2, 48) = 3.26, p = .047, \eta_p^2 = .12\).

Further analysis of the interaction between familiarity and movement type showed that famous faces were sorted significantly better from rigid than non-rigid movement, \(p = .013\), and marginally better from combined than non-rigid movement, \(p = .054\). The advantage for rigid over non-rigid cues was once again driven by static matching performance: for static frames, but not moving clips, famous faces were matched better from rigid than non-rigid images, static: \(p = .032\); moving: \(p = .129\).

The pattern of results for famous faces mirrors those found in Experiment 5: for famous faces, combined and/or rigid movements are generally sorted better than non-rigid movement, probably because of the static head pose information available in the rigid and combined conditions. These results conflict with those of Lander and Chuang (2005), who found a movement advantage for personally familiar faces shown in non-rigid, but not rigid movement. It is unclear why famous faces in this Experiments 5 and 6 show a consistent disadvantage for non-rigid movement – it is possible that the benefits of non-rigid movement are most apparent when detailed mouth movements are included in the stimuli (e.g., Lander & Chuang, 2005; Lander et al., 2006; Rosenblum et al., 2007). This possibility was tested in Experiment 7.

Unlike famous faces, unfamiliar faces were matched equally well from all movement types in the accuracy-per-view analysis (all \(ps > .1\)). There does not appear to be a consistent pattern of results for unfamiliar faces and movement type: in Experiment 5, there was an advantage for combined movement, but Experiment 6 failed to replicate this advantage when accuracy was adjusted for the number of times people watched the clips. Once again, the current results conflict with previous research that found an advantage for rigid movements in unfamiliar face matching (Hill and Johnston, 2001; Watson et al., 2005).

Similarly to the analysis of accuracy, pairwise comparisons between moving and static clips revealed no significant movement advantage in any individual condition, all \(ps > .1\), or for any individual movement type, \(ps > .05\). However, there was an overall movement advantage for famous faces, \(p = .040\), but not unfamiliar faces, \(p = .238\). No other interactions were significant, all \(p’s > .1\).
6.7.3 Neutral Static Frames: A Control Task for the Non-Rigid Condition

An examination of the results of Experiment 6 showed that sorting accuracy for unfamiliar faces in the non-rigid condition was remarkably high, for both moving and static clips. The analysis of views and the results from the accuracy-per-view analysis indicated that this was probably because participants were watching the non-rigid clips more frequently than the combined movement or rigid movement clips. However, it is also possible that participants were watching the clips more often and performing better overall because they were using abnormal cues to sort the clips – several participants commented that the non-rigid condition was very difficult, and they were using cues such as “how wide the eyes looked” or “whether the mouth was open” to sort the clips.

To determine whether the advantage for non-rigid movement for unfamiliar faces arose because the PLDe stimuli had unusually good static information, a second set of static PLDe stimuli were generated. Rather than selecting the static frame from the centre of the moving clip, these static stimuli were created from neutral frames – that it, frames showing a relatively neutral expression, looking ahead, with the mouth closed and eyes open. This was expected to minimise unusual or idiosyncratic static cues and reduce performance in the static condition, especially

Figure 19: Accuracy-per-view results for different movement types in Experiment 6. Error bars represent +/- 1 standard error of the mean.
for unfamiliar faces. Since the moving clips were identical to the original non-rigid condition, performance in the moving conditions was not expected to change.

6.7.3.1 Method.

6.7.3.1.1 Participants. Twenty-two new participants took part in the neutral condition. Participants were undergraduate students (16 female, $M$ age: 23.9 years, range: 17 to 30) from the University of Western Sydney, who took part in the study in return for course credit. All participants reported normal or corrected-to-normal vision. Four participants’ data was excluded from analysis due to failure to pass the familiarity check. To ensure equal numbers for the between-subjects analysis, one extra participant was tested in the normal non-rigid condition. This left 18 participants in the normal non-rigid condition, and 18 participants in the neutral non-rigid condition.

6.7.3.1.2 Stimuli, design and analysis. This experiment used a mixed design, with the within subjects variables of familiarity (famous and unfamiliar) and presentation (moving and static), and the between-subjects variable experiment (normal experiment and neutral experiment).

To create the neutral control experiment, four static neutral frames were generated for each famous and unfamiliar face, and used in place of the static frames from the non-rigid condition. As mentioned above, these neutral frames were extracted from the non-rigid clips, and showed each person with a neutral emotion expression, eyes open and facing forward, and mouth closed. The moving stimuli were identical to those used in the main experiment.

Apart from the change to the static frames, all other details of the stimuli and procedure were identical to Experiment 6. To determine whether this manipulation had a significant effect on performance, the neutral condition was analysed together with the non-rigid condition from Experiment 6, using mixed methods ANOVA with two within-subjects factors (2 x familiarity; 2 x presentation), and one between-subjects factor (2 x experiment: neutral/normal frames).

6.7.3.2 Results and discussion of the neutral frame control experiment.

6.7.3.2.1 Accuracy. Accuracy results from the neutral frame and normal experiments are shown in Table 15. For clarity, the results from Experiment 6 (non-rigid movement only) are repeated in Table 15, to facilitate comparison to the control
experiment. Participants performed above chance in all conditions in the neutral frame experiment, $p < .01$, and in both unfamiliar face conditions in the normal experiment, unfamiliar $p < .0005$; famous static: $p = .107$; famous moving: $p = .081$, as noted in Experiment 6. The ANOVA revealed a main effect of familiarity, $F(1,34) = 5.42, p = .026, \eta^2_p = .14$, and an interaction between familiarity and experiment, $F(1,34) = 19.38, p < .0005, \eta^2_p = .36$. In the neutral frames experiment, there was no effect of familiarity, $p = .151$, whereas unfamiliar faces were sorted significantly better than famous faces in the normal version of the experiment ($p < .0005$, as noted in Experiment 6). No other main effects or interactions were significant (all $p > .1$), but pairwise comparisons were carried out to compare performance in each condition across experiments. In the static conditions, participants in the neutral frame experiment were significantly less accurate than participants in the normal experiment for both famous and unfamiliar faces, $p = .002$, but in the moving conditions, performance did not change significantly, $p > .05$. In other words, asking participants to sort neutral static frames decreased accuracy in the static condition, but not the moving condition (as was expected), compared to the main experiment. This suggests that the static frames in the normal experiment did carry some static identity cues, but they were reduced in the neutral frame experiment.

The comparison between neutral and randomly chosen static frames confirms that static cues such as expression can be important when matching faces, even when there is very little identifying static information available. Previous studies that have used neutral frames in their static condition (e.g., Knight & Johnston, 1997; see Chapter 4 for more examples) have minimised useful static cues, which may have artificially inflated the movement advantage.

6.7.3.2.2 Views. The average number of views per clip for the neutral frame and normal experiments is shown in Table 15. The ANOVA on views revealed main effects of presentation, $F(1,34) = 8.33, p = .007, \eta^2_p = .20$, and experiment, $F(1,34) = 28.55, p < .0005, \eta^2_p = .46$, and an interaction between presentation and experiment, $F(1,34) = 4.81, p = .035, \eta^2_p = .12$. Overall, participants watched static clips more often than moving clips, and clips in the normal experiment more than clips in the neutral version of the experiment. The interaction between presentation and experiment indicated that participants in the neutral frame experiment watched the moving and static clips the same number of times on average, $p = .628$, but participants in the normal experiment watched the static clips more times than the
moving clips, \( p = .001 \), as reported in Experiment 6. No other main effects or interactions were significant, all \( ps > .1 \).

It is unclear why participants in the neutral experiment did not watch the static clips more often, as would be expected if they found them more difficult. It is possible that participants in the neutral experiment realised they could not get any more useful information from the static frames, whereas participants in the normal experiment re-watched the static clips, looking for specific idiosyncratic cues to identity. Once again, this indicates that the advantage for unfamiliar PLDe in the normal experiment probably arose because participants watched the static clips more often, possibly because they were trying to extract unusual static cues.

Table 15:
*Mean Accuracy and Views for the neutral frame and normal experiment. Both showed non-rigid movement.*

<table>
<thead>
<tr>
<th></th>
<th>Neutral Frame</th>
<th>Normal^a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Moving</td>
<td>Static</td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Famous</td>
<td>15.89*</td>
<td>15.11**</td>
</tr>
<tr>
<td></td>
<td>(6.84)</td>
<td>(4.91)</td>
</tr>
<tr>
<td>Unfamiliar</td>
<td>14.56**</td>
<td>12.33*</td>
</tr>
<tr>
<td></td>
<td>(4.16)</td>
<td>(3.90)</td>
</tr>
<tr>
<td><strong>Views</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Famous</td>
<td>2.67</td>
<td>2.72</td>
</tr>
<tr>
<td></td>
<td>(1.13)</td>
<td>(0.80)</td>
</tr>
<tr>
<td>Unfamiliar</td>
<td>2.55</td>
<td>2.72</td>
</tr>
<tr>
<td></td>
<td>(1.04)</td>
<td>(0.92)</td>
</tr>
</tbody>
</table>

*Note:* Chance level accuracy was 9.6. Standard deviations are shown in brackets.

^aTaken from the non-rigid condition in Experiment 6, with one participant added.

\*\( p < .05 \) ** \( p < .0005 \)
6.7.3.2.3 **Accuracy per view.** The ANOVA on accuracy-per-view showed that all three main effects were significant. Unfamiliar faces were sorted slightly better than famous faces, $F(1,34) = 4.22$, $p = .048$, $\eta^2_p = .11$; moving faces were sorted slightly better than static faces, $F(1,34) = 8.42$, $p = .006$, $\eta^2_p = .20$; and the neutral frame experiment led to better sorting than the normal experiment, $F(1,34) = 14.97$, $p < .0005$, $\eta^2_p = .31$. There was also a significant familiarity by experiment interaction, $F(1,34) = 8.79$, $p = .005$, $\eta^2_p = .20$, reflecting the fact that unfamiliar faces were sorted better than famous faces in the normal experiment, $p = .001$, as noted in Experiment 6, but famous and unfamiliar faces were sorted equally well in the neutral frame experiment, $p = .524$.

6.7.3.3 **Conclusions from the neutral frame control experiment.** The neutral frame experiment eliminated the accuracy and accuracy-per-view advantage for unfamiliar face sorting that was found for non-rigid PLDe in Experiment 6. The analysis of accuracy-per-view in the main experiment and the results from the neutral frame experiment confirm that participants’ unusually high accuracy when sorting unfamiliar face PLDe showing non-rigid movement was most likely an anomaly – a result of watching the clips more often and focusing on idiosyncratic static information, rather than participants being particularly good at sorting unfamiliar faces from non-rigid movement. Overall, then, it appears that non-rigid movement is equally helpful for famous and unfamiliar face sorting, and the advantage for unfamiliar faces and non-rigid movement in Experiment 6 probably reflects idiosyncrasies in the static stimuli or participants’ sorting strategies in the unfamiliar PLDe condition.

6.7.4 **Conclusions from Experiment 6**

The results from Experiment 6 suggest that participants can use movement cues to accurately sort PLDe. Unlike Experiment 5, Experiment 6 showed a significant overall movement advantage in the accuracy-per-view analysis, but not the accuracy analysis. Furthermore, static clips were watched more times on average than moving clips, replicating the effect found in the pilot sorting task. The presence of a movement advantage confirms that the PLDe contained more useful movement information than the PLDs – in other words, eye movements do contain some cues to identity. These findings also confirm that the movement advantage does not just
reflect the fact that moving faces are sorted more accurately than static faces, but that movement can make it easier or faster to sort faces. Since most studies using moving faces have not taken reaction times into account (with the exception of the priming and visual search tasks, Pilz et al., 2006; Thornton & Kourtzi, 2002), this is an important finding; future research should examine the effect of movement on measures other than accuracy.

In contrast to the other experiments in this thesis, there was an overall advantage for unfamiliar face sorting compared to famous face sorting. This effect arose because unfamiliar faces were sorted more accurately than famous faces in the non-rigid condition. However, the effect remained in the accuracy-per-view analysis, which suggests that the advantage for unfamiliar faces was not simply due to participants watching the clips more often. The neutral frame control experiment confirmed that participants performed significantly worse in the static condition when neutral static frames were chosen, and there was no significant difference between static matching performance for famous and unfamiliar faces using neutral frames. Taken together, these results indicate that unfamiliar PLDe in the non-rigid condition had some sort of static information that participants were focusing on when completing the task (e.g. characteristic expressions), and the advantage for unfamiliar faces in the non-rigid condition does not reflect a true advantage for non-rigid movement matching.

Despite this anomaly, the results from Experiment 6 suggest that the type of movement is an important factor in famous and unfamiliar face sorting. In line with Experiment 5, combined movement was sorted best overall, and famous faces showed a disadvantage for non-rigid movement. The results for unfamiliar faces were inconsistent, probably due to anomalous results in the non-rigid condition.

6.8 Experiment 7: Rigid and Non-Rigid Movement in Shape-normalised Point-light-displays with Teeth and Tongues

Experiment 7 was designed to investigate whether showing participants real articulatory information, in the form of the teeth and tongue, improves sorting performance or has any impact on the effect of familiarity or movement type.
6.8.1 Methods of Experiment 7

6.8.1.1 Participants. Sixty undergraduate students (47 female, \(M\) age: 20.4 years, range: 18 to 32) from the University of Western Sydney took part in the study in return for course credit. Participants reported normal or corrected-to-normal vision. Nine participants’ data was excluded from analysis due to failure to pass the familiarity check.

6.8.1.2 Stimuli and Materials. The stimuli in this experiment were identical to those in Experiment 5, except that they contained extra information in the mouth region. The PLDs in Experiment 5 and 6 included six markers that tracked the movement of the outer lips, but the tracking program was unable to track the teeth and tongue. In Experiment 7, the inner mouth region from the original video was converted to greyscale and inserted into the PLD (see Figure 17d). Whenever the individual in the video opened their mouth, the area inside the lips was included in the resulting clip (referred to as PLDm). Separate groups of participants saw rigid head movements (\(n = 17\)), non-rigid face movements (\(n = 17\)) or combined rigid and non-rigid movements (\(n = 17\)).

6.8.2 Results and Discussion of Experiment 7

6.8.2.1 Accuracy. Accuracy for each condition is shown in Table 16. Participants who viewed clips with combined or rigid movement were above chance in all conditions (all \(p < .05\)). Participants in the non-rigid movement condition sorted famous faces (both moving and static) above chance levels, moving: \(p = .003\); static: \(p = .018\). Unfamiliar faces in the non-rigid static condition were sorted better than chance, \(p = .014\), but unfamiliar non-rigid movement was sorted at chance levels, \(p = .235\).

The ANOVA on accuracy revealed no significant main effects. However, there was a trend for famous faces to be sorted more accurately than unfamiliar faces, \(F(1,48) = 3.81, p = .057, \eta_p^2 = .07\), and for moving clips to be sorted more accurately than static clips, \(F(1,48) = 3.35, p = .073, \eta_p^2 = .06\). No other main effects or interactions approached significance, all \(p > .08\). Unlike PLDs and PLDs with eyes, participants were equally accurate when sorting PLDm from combined, rigid and non-rigid movement, for both famous and unfamiliar faces, \(p > .1\).
Planned pairwise comparisons showed that moving clips were sorted more accurately than static clips in the combined movement condition, \( p = .011 \), and for famous faces overall, \( p = .035 \). Pairwise comparisons within each condition showed a movement advantage for famous faces in combined movement, \( p = .047 \), and a near significant movement advantage for unfamiliar faces in combined movement, \( p = .064 \). No other comparisons were significant, \( ps > .1 \).

### 6.8.2.2 Views

The average number of views per clip for each condition is also shown in Table 16. The ANOVA on average number of views per clip revealed no significant main effects or interactions (\( ps > .05 \)).

Table 16: 
*Mean Accuracy and Views for Experiment 7*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Combined</th>
<th>Non-Rigid</th>
<th>Rigid</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Moving</td>
<td>Static</td>
<td>Moving</td>
</tr>
<tr>
<td>Accuracy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Famous</td>
<td>17.65**</td>
<td>13.18*</td>
<td>15.88*</td>
</tr>
<tr>
<td></td>
<td>(7.08)</td>
<td>(3.18)</td>
<td>(7.30)</td>
</tr>
<tr>
<td>Unfamiliar</td>
<td>15.65**</td>
<td>12.82*</td>
<td>10.94</td>
</tr>
<tr>
<td></td>
<td>(4.26)</td>
<td>(4.13)</td>
<td>(4.48)</td>
</tr>
<tr>
<td>Views</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Famous</td>
<td>2.56</td>
<td>3.05</td>
<td>3.19</td>
</tr>
<tr>
<td></td>
<td>(0.86)</td>
<td>(1.08)</td>
<td>(1.07)</td>
</tr>
<tr>
<td>Unfamiliar</td>
<td>2.89</td>
<td>2.91</td>
<td>3.12</td>
</tr>
<tr>
<td></td>
<td>(1.08)</td>
<td>(1.11)</td>
<td>(1.17)</td>
</tr>
</tbody>
</table>

*Note:* Chance level accuracy was 9.6. Standard deviations are shown in brackets.  
\*\( p < .05 \) \*\( p < .0005 \)
6.8.2.3 **Time.** The ANOVA revealed no significant main effects or interactions for time to complete each sorting task, all $p$’s > .1.

6.8.2.4 **Accuracy per view.** Accuracy per view results for Experiment 7 are shown in Figure 20. The ANOVA on accuracy per view showed main effects of presentation and movement type, presentation: $F(1,48) = 4.90$, $p = .032$, $\eta_p^2 = .09$; movement type: $F(2,48) = 5.30$, $p = .008$, $\eta_p^2 = .18$.

Moving clips were sorted significantly better than static clips, and combined movement was sorted significantly better than rigid or non-rigid movement, rigid: $p = .033$; non-rigid: $p = .014$. This is broadly consistent with the results of Experiment 6, which found an accuracy-per-view advantage for combined over non-rigid movement for PLDe. There was no main effect of familiarity, $F(1,48) = 1.15$, $p = .289$, $\eta_p^2 = .02$, and no interactions approached significance, all $p$s > .1.

The movement type results diverged from Experiment 6 when the accuracy-per-view results were analysed separately for famous and unfamiliar faces. Although the interactions failed to reach significance, planned pairwise comparisons were carried out to investigate the effects of familiarity and the presence of a movement advantage for PLDm. Unlike Experiment 6, unfamiliar faces were sorted better in the rigid than non-rigid movement condition, $p = .036$. This is the only result in all three experiments that replicates the findings of Hill and Johnston’s (2001) study, which found that unfamiliar, moving faces were sorted more accurately from rigid than non-rigid movement. However, like in Experiments 5 and 6, the advantage for the rigid condition was driven by better sorting of rigid static clips (i.e. those containing head pose information). Pairwise comparisons showed that static clips in the rigid condition were sorted significantly better than static clips in the non-rigid condition, $p = .015$, but there were no significant effects of movement type for the moving clips, $p$s > .1. It is also important to note that the advantage for rigid over non-rigid clips for unfamiliar faces was only apparent in this study when accuracy was adjusted for average views (not measured by Hill & Johnston, 2001). This may indicate that participants were compensating for poor performance or uncertainty in the non-rigid condition, particularly the static non-rigid condition, by watching the clips more frequently (although not enough to reach significance in the views analysis). This will be discussed further in the general discussion.
Famous faces were recognised equally well from all types of movement, with the exception of a marginal trend for combined movement to be sorted better than non-rigid movement, \( p = .069 \) – the same trend found for famous faces in Experiment 6.

As in Experiment 6, there was an overall movement advantage for famous faces, \( p = .024 \), but not unfamiliar faces, \( p = .297 \). This suggests that participants were able to use characteristic movement patterns to distinguish famous faces, but they were relying primarily on static information to discriminate unfamiliar faces.

Unlike the analysis of movement type, the movement advantage for famous faces was consistent across the accuracy and accuracy per view analyses, which confirms that participants were not more accurate simply because they watched the clips more often in the famous faces condition. In addition to the movement advantage for famous faces, there was an overall movement advantage for faces showing combined movement, \( p = .006 \), driven by a movement advantage for famous faces in the combined movement condition, \( p = .009 \). There was no overall movement advantage for rigid or non-rigid movement, \( p’s > .1 \), and no other pairwise comparisons approached significance, \( p’s > .09 \).
6.8.3 Conclusions from Experiment 7

In general, the results from Experiment 7 support the idea that detailed mouth movements, like the eyes movements tested in Experiment 6, carry useful identity information. This is particularly true for famous faces, which showed a significant movement advantage in both experiments.

The effects of presentation (the movement advantage) in this experiment were broadly consistent with the results of Experiment 6. Moving clips were sorted better than static clips in the accuracy per view analysis, and there was an overall movement advantage for famous, but not unfamiliar faces. The effect of movement type was also similar across Experiment 6 and 7: For famous faces, PLDs with eyes and PLDs with mouths were both sorted somewhat better from combined than non-rigid movement. For unfamiliar faces, on the other hand, there was no advantage for combined movement. PLDs with mouths were sorted better from rigid than non-rigid movement, but this was probably a consequence of static pose information rather than the use of different movement cues.

The most interesting finding in Experiment 7 was that, once again, there was no indication of an advantage for non-rigid movement for familiar faces, either as a movement advantage or as an overall benefit for sorting non-rigid movement (cf. Lander & Chuang, 2005). This implies that the advantage for non-rigid movement in their study did not arise because they showed detailed mouth movements. Experiment 7 did find evidence for a rigid movement advantage for unfamiliar faces (when compared to non-rigid movement, but not combined movement, cf. Hill & Johnston, 2001), but as discussed above, this probably arose because participants sorted the faces based on head poses, rather than characteristic rigid head movements.

The results from Experiment 6 and 7 were relatively consistent, but it was unclear whether the extra eye information (Experiment 6) or mouth information (Experiment 7) improved sorting performance compared to basic PLDs (Experiment 5), and, if so, whether the eyes or mouth were more useful when sorting faces according to identity. The next section compares the effect of mouth movements (Experiment 7) and eye movements (Experiment 6) to baseline PLD matching (Experiment 5), to determine whether one facial area or type of movement is particularly useful overall.
6.9 Combined Analysis of Experiments 5, 6, and 7

The results from Experiments 5, 6 and 7 were combined and subjected to a mixed methods ANOVA, with two within subjects variables (presentation and familiarity) and two between subjects variables (stimulus type and movement type).

6.9.1 Results and Discussion of the Combined Analysis

6.9.1.1 Accuracy. The ANOVA on accuracy revealed no main effect of familiarity or presentation, $F(1,144) < 1$ for both comparisons. The main effects of stimulus type and movement type were both significant, stimulus type: $F(2,144) = 11.37, p < .0005, \eta_p^2 = .14$; movement type: $F(2,144) = 5.31, p = .006, \eta_p^2 = .07$.

Overall, PLDs with eyes and PLDs with mouths were sorted equally well, $p = .114$, and both were sorted more accurately than standard PLDs, PLDe: $p < .0005$; PLDm: $p = .026$. In general, clips in the combined movement condition were sorted more accurately than either the rigid movement, $p = .040$, or non-rigid movement, $p = .008$, conditions.

There were also a number of significant interactions, mostly related to familiarity. There was a significant interaction between familiarity and stimulus type, $F(2,144) = 5.28, p = .006, \eta_p^2 = .07$, and a near-significant three-way interaction between familiarity, movement type and stimulus type, $F(4,144) = 2.35, p = .057, \eta_p^2 = .06$. These interactions reflect the fact that the effect of familiarity and the interaction between familiarity and movement type were inconsistent across Experiments 5, 6 and 7 (see individual experiments for more detail). The familiarity by stimulus type interaction also reflects the fact that famous faces were sorted with equal accuracy across all three stimulus types ($ps > .05$), whereas unfamiliar faces were sorted more accurately from PLDe than from PLDs or PLDm ($ps < .0005$).

The ANOVA revealed a highly significant interaction between familiarity and movement type, $F(2,144) = 5.75, p = .004, \eta_p^2 = .07$. For famous faces, combined and rigid movement were sorted more accurately than non-rigid movement, combined: $p = .008$; rigid: $p = .029$, whereas for unfamiliar faces, combined and non-rigid movement were sorted more accurately than rigid movement, combined movement: $p = .005$; non-rigid movement: $p = .040$. There was a significant familiarity by presentation interaction, $F(1,144) = 5.61, p = .019, \eta_p^2 = .04$, reflecting the fact that famous faces were sorted more accurately from moving than static clips,
p = .047, whereas unfamiliar faces were sorted equally well from both moving and static clips, p = .230. The familiarity by movement type and familiarity by presentation interactions were qualified by a three-way interaction between familiarity, presentation and movement type, F(2,144) = 3.12, p = .047, \( \eta_p^2 = .04 \). Pairwise comparisons showed a movement advantage for famous faces showing non-rigid movement, p = .029, and a trend for a movement advantage for unfamiliar faces showing combined movement, p = .076. There was also a movement disadvantage for unfamiliar faces showing non-rigid movement, p = .035. No other comparisons were significant, ps > .05. Pairwise comparisons were also conducted on movement type. They showed no differences for famous moving clips or unfamiliar static clips, ps > .05. For famous static clips, the pairwise comparisons mirrored the familiarity by movement type interaction (combined and rigid static clips were sorted more accurately than non-rigid static clips), combined: p = .004; rigid: p = .004, but for unfamiliar moving clips, combined movement was sorted significantly more accurately than rigid movement, p < .0005, and somewhat more accurately than non-rigid movement, p = .057.

6.9.1.2 Views. The ANOVA on average number of views per clip showed no significant main effect or interactions related to familiarity. This confirms that the familiarity effects found for accuracy were not based on the participants watching the clips more often. However, there were significant main effects for presentation, F(1,144) = 17.39, p < .0005, \( \eta_p^2 = .11 \), stimulus type, F(2,144) = 5.91, p = .003, \( \eta_p^2 = .06 \), and movement type, F(2,144) = 4.33, p = .015, \( \eta_p^2 = .06 \). Participants watched the static clips significantly more times on average than the moving clips, which is interesting given that static and moving clips were sorted with equal accuracy. On average, participants viewed PLDe more times than PLDm, p = .002, and non-rigid more times than combined movement clips, p = .016. No other pairwise comparisons on the main effects were significant, all ps > .1.

The main effect of presentation was modified by two significant interactions. The effect of presentation varied across stimulus types, F(2,144) = 4.73, p = .010, \( \eta_p^2 = .06 \).

The lack of movement advantage for unfamiliar faces was not a result of the abnormal accuracy results in the non-rigid condition of Experiment 6. The combined analysis was re-run with the neutral frame experiment results replacing the normal non-rigid PLDe results, and unfamiliar faces still failed to show a movement advantage, p = .405.
= .06, reflecting the fact that static clips were viewed significantly more times than moving clips in Experiment 6 (PLDs with eyes), but not Experiments 5 or 7 (PLDs and PLDs with mouths). Presentation also interacted with movement type, $F(2,144) = 5.96, p = .003, \eta_p^2 = .08$. Static clips were viewed more times than moving clips for combined and non-rigid movement, $p$’s < .0005, but not for rigid movement, $p = .681$.

6.9.1.3 Time. The ANOVA on completion time revealed a significant effect of stimulus type, $F(2,144) = 4.31, p = .015, \eta_p^2 = .06$. Participants who sorted the PLDs with eyes took significantly longer to complete each task on average compared to participants who sorted PLDs with mouths, $p = .012$. This may be because the PLDe participants also watched the clips more times on average than the PLDm participants.

No other main effects or interactions were significant, all $ps > .1$.

6.9.1.4 Accuracy per view. The accuracy per view ANOVA mirrored many of the accuracy results. The main effect of movement type was significant, $F(2,144) = 9.59, p < .0005, \eta_p^2 = .12$. When results were averaged across all levels of familiarity, presentation and stimulus type, participants sorted combined movement significantly better than non-rigid movement, $p < .0005$, and somewhat more accurately than rigid movement, $p = .076$. The main effect of stimulus type approached significance, $F(2,144) = 2.94, p = .056, \eta_p^2 = .04$, but no pairwise comparisons reached significance.

There was no main effect of familiarity, but there were significant interactions between familiarity and stimulus type, $F(2,144) = 3.07, p = .050, \eta_p^2 = .04$; familiarity and movement type, $F(2,144) = 3.72, p = .027, \eta_p^2 = .05$; and familiarity, stimulus type and movement type, $F(4,144) = 3.34, p = .012, \eta_p^2 = .08$. Famous faces were sorted significantly better from PLDe and PLDm than PLDs, $ps < .05$, whereas unfamiliar faces were sorted equally well from all stimulus types, $ps > .6$. Across the different stimuli, famous faces were sorted equally well from combined and rigid movement, $p = 1$; both types of movement were sorted significantly better than non-rigid movement, combined: $p < .0005$; rigid: $p = .001$. However, unfamiliar faces were sorted better from combined than rigid or non-rigid movement, rigid: $p = .029$; non-rigid: $p = .042$. Like in the analysis of accuracy, the three-way interaction
reflects the variability in the familiarity by movement interaction across Experiments 5, 6 and 7, particularly the variable effects of movement type on unfamiliar faces. The familiarity by presentation interaction was not significant, $F(1,144) = 1.54, p = .217, \eta^2_p = .02$, but planned pairwise comparisons revealed that famous faces were sorted better from moving than static clips, $p = .008$, whereas unfamiliar faces were sorted equally well from both moving and static clips, $p = .153$.

Finally, unlike in the accuracy analysis, there was a main effect of presentation, $F(1,144) = 6.81, p = .010, \eta^2_p = .04$, and a significant interaction between presentation and movement type, $F(2,144) = 4.24, p = .016, \eta^2_p = .06$. Moving clips were sorted significantly better than static clips in the combined movement condition, $p = .001$, but not in the rigid or non-rigid movement conditions, $ps > .1$. The three-way interaction between familiarity, presentation, and movement type was not significant, $F(2,144) = .672, p = .512, \eta^2_p = .01$, but pairwise comparisons showed that there was a significant movement advantage for both famous and unfamiliar faces showing combined movement, famous and unfamiliar: both $ps = .005$, and a movement advantage for famous faces displaying non-rigid movement, $p = .038$. No other comparisons were significant, all $ps > .5$. The movement advantage for non-rigidly moving famous faces in the accuracy and accuracy-per view analyses is the only evidence of an advantage for non-rigid movement for famous faces, as reported by Lander and Chuang (2005). The lack of movement advantage for rigid movement further supports the idea that rigid clips were sorted consistently well because of static cues (such as head poses), but adding or isolating non-rigid movement allowed participants to effectively use movement cues to sort famous faces. No other interactions were significant, all $ps > .1$.

**6.9.2 Conclusions from the Combined Analysis**

Overall, these results reflect a similar pattern to the pilot sorting experiment and Experiments 5, 6 and 7. Accuracy is primarily affected by the familiarity of the person, and varies significantly depending on the type of movement shown in the clip. In general, famous faces are sorted more accurately from combined and rigid movement than non-rigid movement, whereas unfamiliar faces are sorted more accurately from combined and non-rigid movement than rigid movement. Interestingly, only combined movement gives rise to an overall movement advantage in the accuracy analysis.
The number of times participants view each clip is affected by the presentation of the clips – whether the clip is moving or static – and also varies significantly depending on the type of movement in the clip. In general, participants view static clips more often than moving clips to complete the sorting task.

When both accuracy and number of views are taken into account, both familiarity and presentation have a significant effect on the sorting task, and this is strongly influenced by the type of movement being displayed. Interestingly, neither accuracy nor accuracy per view replicated the effect of movement found by Hill and Johnston (2001) in their sorting experiment, or Watson et al. (2005) in their matching task. Both studies found that unfamiliar, moving faces were sorted or matched more accurately from rigid than non-rigid movement (although Watson et al., 2005, found an overall advantage for combined movement, whereas Hill & Johnston, 2001, found an overall advantage for rigid movement). In the current experiments, combined movement was generally sorted better than either rigid or non-rigid movement, regardless of familiarity. This is in line with Watson et al.’s (2005) results, but not Hill and Johnston’s (2001) findings. It is possible that the discrepancy in results arose due to the type of stimuli used in the current study, or due to the timing restrictions on the task. These factors will be discussed further in the general discussion.

Unlike the individual experiments, the combined analysis provided some limited support for Lander and Chuang’s (2005) results, finding a movement advantage for famous faces in the non-rigid, but not rigid movement conditions.

6.10 General Discussion of Experiments 5, 6, and 7

Participants are very good at sorting PLDs into groups based on identity. Across the three experiments in this chapter, sorting accuracy was above chance-level in the majority of conditions. Furthermore, Experiments 6, 7, and the combined analysis found a significant overall movement advantage. However, the effects of movement and familiarity differed from those found in Chapters 4 and 5, and the effects of movement type were quite different to those found in previous studies – in Experiments 5, 6, and 7, famous faces were sorted poorly from non-rigid clips, and unfamiliar faces were sorted best from combined rigid and non-rigid movement. Interestingly, the results indicate that participants can extract and use movement
information from multiple areas of the face: adding either eyes or the teeth and
tongue to basic PLDs improves performance to a similar extent.

6.10.1 Familiarity and the Movement Advantage in the Sorting Task

In general, unlike the matching tasks in Experiments 1a, 2a, 3, and 4, and the
pilot sorting task, there was no overall advantage for sorting famous faces in
Experiments 5 and 6, and only a minimal advantage for sorting famous faces in
Experiment 7. Furthermore, the sorting tasks failed to replicate the unfamiliar face
movement advantage found in Experiments 1a, 2a, and 4. This finding supports the
theory, advanced in Chapter 5, that the movement advantage for unfamiliar faces
arises primarily from structure-from-motion processes (or, in some cases, from social
cues, e.g., Thornton & Kourtzi, 2002). When the faces do not differ in structure
(because it is normalised), as in the current chapter, movement only has a small
impact on unfamiliar face matching. This appears to be at odds with other studies
that have found a movement advantage for normalised unfamiliar faces (e.g., Hill &
Johnston, 2001; Watson et al., 2005), but those studies did not include a static control
condition, which makes it difficult to disentangle the effects of static recognition and
the effects of movement. Furthermore, Watson et al. asked participants to match
individual movement sequences (akin to the task in Experiments 1a and 2a, Chapter
4), and the research presented in this thesis has established that matching from one
movement sequence to another, as in the current experiments, is a much harder task
(see Chapter 4).

It should be noted that Experiment 4 (Chapter 5) and some conditions of
Experiment 5, 7, and the combined analysis did find a significant movement
advantage for shape-normalised unfamiliar faces, which suggests that characteristic
movement patterns can give rise to a movement advantage in unfamiliar faces under
some circumstances. However, the movement advantage for unfamiliar faces in this
chapter was relatively small compared to that for famous faces, which is the reverse
of the results found in Chapter 5. This difference may have arisen due to different
task constraints. Experiment 4 found extremely poor matching performance for static
unfamiliar avatars, possibly because they were only shown once during each trial.
The fact that the sorting task allowed participants to view the clips multiple times
may have improved performance in the static condition compared to the
same/different task – in fact, the majority of static conditions in Experiments 5, 6,
and 7 were sorted at above-chance levels – which could have reduced the movement advantage.

Although there was no evidence of an overall movement advantage for unfamiliar faces, Experiment 7 and the combined analysis did find a significant movement advantage for famous faces. This is the first time a movement advantage has been shown for famous faces presented as shape-normalised stimuli, and this finding provides strong evidence that people can use characteristic patterns of movement to match famous faces. However, the movement advantage for famous faces did not appear in Experiment 5, which suggests that characteristic movement patterns for famous faces are easiest to extract and match when detailed mouth or eye movement is present. The role of mouth and eye movement information will be discussed further in section 6.10.3.

6.10.2 Rigid and Non-rigid Movement

Overall, combined movement (featuring both rigid and non-rigid movement) was sorted better than isolated rigid or non-rigid movement. The exact details of the effect of movement type for each study varied, but combined movement was sorted as well or better than rigid and non-rigid movement in every accuracy and accuracy-per-view comparison, regardless of familiarity or stimulus type. Combined movement was also the only movement type to show an overall movement advantage: combined movement clips were sorted better than static clips in Experiment 7 (accuracy and accuracy-per-view analyses), and in the combined analysis (accuracy-per-view analysis).

The effect of movement type also varied significantly with familiarity. In general, famous faces were sorted worst from non-rigid movement. Like the overall effect of movement type, the effect of movement type for famous faces varied throughout the three experiments, but a disadvantage for non-rigid clips appeared in both the accuracy and accuracy-per-view analyses in Experiment 5, 6, and the combined analysis. Surprisingly, though, the overall analysis revealed that famous faces showed a movement advantage for non-rigid movements – an effect that was not apparent in any of the individual analyses. No previous research has examined the effect of movement type for famous faces. However, this result is broadly consistent with previous studies: a review of research using personally familiar faces reveals that non-rigid movements consistently give rise to a movement advantage, of
equal or greater size than the movement advantage for rigid movement (Bruce & Valentine, 1988; Lander & Chuang, 2005; see also Lander et al., 2006; Rosenblum et al., 2007). The overall disadvantage of non-rigid movement probably arose because participants performed particularly poorly when asked to sort famous faces using static non-rigid clips (as is shown in the results from Experiment 5 and the combined analysis), but were relatively accurate when asked to sort moving clips, regardless of whether non-rigid, rigid, or combined movement was shown.

While non-rigid clips were sorted poorly from static frames, but showed a significant movement advantage, rigid clips of famous faces produced the opposite results. Rigid clips of famous faces were sorted very well overall – famous faces were sorted equally well from rigid and combined movements in all analyses, in all experiments – but, unlike combined or non-rigid movement, rigid head movements do not give rise to a movement advantage in any condition. In other words, there was no evidence that participants were using isolated rigid head movements as a cue to identity. This suggests that isolated rigid head movements provide good static cues for sorting, but they do not necessarily provide additional, movement-based identity information.

In general, famous faces were sorted better from combined rigid and non-rigid movement than either cue in isolation: although there was no significant difference between sorting performance for rigid and combined movements, combined movements gave rise to a movement advantage in Experiment 7 and the combined analysis. This indicates that combined movements provide better movement cues than rigid movements, probably because participants are using some non-rigid movement as a cue to identity. Alternatively, it is possible that characteristic movement patterns for famous faces consist of a combination of rigid and non-rigid movements (e.g., a head shake or pose paired with a characteristic smile or grimace), and presenting either cue in isolation impairs sorting performance (since people rarely produce isolated rigid or non-rigid movement in real life). If rigid cues are still present, participants may be able to fall back on pose information, explaining the high performance in the rigid condition; if non-rigid cues are available, people might use them to extract identity cues, explaining the limited movement advantage; but presenting either cue alone limits the movement advantage.

Unfamiliar faces showed a slightly different pattern of results for movement type than famous faces. The overall analysis revealed that combined movement was
sorted more accurately than either non-rigid or rigid movement, and, like famous faces, there was a significant movement advantage for unfamiliar faces showing combined movement in Experiment 5 and the combined analysis, and a near-significant movement advantage for combined movement in the accuracy analysis of Experiment 7. Unlike famous faces, there was no indication that non-rigid movement alone gave rise to a movement advantage, which suggests that, for unfamiliar faces at least, characteristic movement patterns consist of a combination of rigid and non-rigid movement cues, and presenting either cue alone eradicates any movement advantage.

The fact that unfamiliar faces did not show any movement advantage for non-rigid clips, but famous faces did, suggests that as we become more familiar with a face we are better able to extract and compare isolated non-rigid movements. Unfamiliar faces also failed to show the consistent benefit for rigid clips that was present for famous faces. This indicates that familiarity with a face may increase our sensitivity to characteristic rigid head poses (not rigid movements, as discussed above), which may explain why famous faces were matched so well from static images in Experiments 1a, 2a, 3, and 4 (Chapters 4 and 5).

As mentioned in the individual experiments, this study failed to replicate the rigid over non-rigid movement advantage for unfamiliar faces found by Hill and Johnston (2001) and Watson et al. (2005). There are many reasons why their results might differ from Experiments 5, 6, and 7: it is possible that participants in Hill and Johnston’s study were using the amount or distinctiveness of rigid movement as a cue to identity, which may have increased accuracy in their rigid condition – Hill and Johnston did not measure the distinctiveness or amount of movement in their clips, whereas the clips used in the current research were matched for these movement characteristics. It is also possible that the actors in their study produced less varied non-rigid movements than the individuals used to create the PLDs in this study, since their actors were reading from a script, rather than speaking candidly. Another, somewhat simpler explanation of the difference in results is that Hill and Johnston only tested accuracy, and only for moving clips – it is possible that their participants watched the clips in the rigid condition more often, leading to more accurate sorting, or that the inclusion of a static condition in this experiment changed participants’ responding strategies. Finally, imposing a time limit on the task may also have changed participant’s response strategies. Although they were instructed to
concentrate on accuracy, not time, participants may have been conscious of the time measure and rushed the task, which again may have encouraged a different response strategy to the participants in Hill and Johnston’s study.

Overall, Experiments 5, 6, and 7 found no conclusive evidence that either isolated rigid or non-rigid movement provide reliable identity cues for famous or unfamiliar faces. The movement advantage is most apparent when both types of movement are present, which indicates that characteristic movement patterns contain both rigid and non-rigid cues, and it is difficult to separate them for either famous or unfamiliar faces.

6.10.3 Eyes and Mouths

This series of experiments was the first systematic examination of the role of different facial areas in a movement-based identification task. Overall, the combined analysis found that adding information from either the eye or mouth regions improved sorting accuracy. This is not surprising – any extra movement cues were expected to provide identity information that could improve performance. Interestingly, the effect of facial area was only present in the accuracy analysis, and not when the results were adjusted for the average number of views per clip. Participants may have needed extra exposure to the clip (i.e., extra views) to benefit from the added eye or mouth movement cues.

The effect of facial area differed for famous and unfamiliar faces. For famous faces, all PLDs were sorted with equal accuracy, but when views were accounted for, PLDs with extra mouth information were sorted better than normal PLDs or PLDs with eyes. This is in line with previous research using personally familiar faces that has shown a movement advantage for stimuli that focus on mouth movement (Lander et al., 2006; Rosenblum et al., 2007), and indicates that the mouth might be the primary source of the non-rigid movement advantage for familiar faces (Lander & Chuang, 2005); perhaps the mouth provides the most reliable movement information for identity discrimination. The findings from the combined analysis may also explain Bruce and Valentine’s (1988) results – they found an equal movement advantage for both rigid and non-rigid movement, but they did not include detailed teeth and tongue movements (whereas Lander and Chuang used degraded video stimuli, which contained mouth cues).
Unlike famous faces, participants were more accurate when sorting unfamiliar PLDs with eyes than normal PLDs or PLDs with mouths. Once again, there is a good possibility that this effect arose due to participants adopting strange response strategies in the non-rigid condition of Experiment 6, since unfamiliar faces were sorted equally well from all PLDs in the accuracy-per-view analysis. It is interesting that unfamiliar faces did not show a similar effect as famous faces – even when the anomalous condition was taken into account, there was no indication that participants could use detailed articulatory information to improve sorting performance for unfamiliar faces.

These results indicate that familiarity with a face may make it easier to extract characteristic movement patterns from the mouth region, perhaps because the mouth displays timing cues related to speech and expression. However, it is important to note that both eyes and mouths gave rise to a movement advantage for famous faces, suggesting that both eye and detailed mouth movement are sufficient, but not necessary, for a movement advantage for famous faces. On the other hand, neither eye or mouth information is necessary for unfamiliar face sorting: even in the basic PLD condition, there was a limited movement advantage for unfamiliar faces (combined movement only). It is possible that the detailed information in the PLDe and PLDm was not helpful for unfamiliar faces, because participants were completing the sorting task for unfamiliar faces based on gross mouth and head movements and/or static cues, rather than the fine detailed movement cues they used for famous faces.

6.10.4 Advantages and Disadvantages of the Sorting Task

This chapter used a different task to Chapters 4 and 5, and introduced several new measures (views; accuracy-per-view) to examine the effect of movement on face matching. As mentioned in the introduction to this chapter, the sorting task has several advantages over traditional same/different matching tasks: participants can complete the task at their own pace and watch the clips multiple times, the sorting task eliminates the possibility of response bias, and it is much faster than the same/different experiments. The results from Experiments 5, 6, and 7 confirm that participants performed remarkably well, even in the static conditions, which suggests that allowing participants to compare the clips multiple times at their own pace probably helped matching performance compared to Experiments 1-4. However,
giving participants more freedom may have minimised the movement advantage, by allowing participants to concentrate on static cues and inflate static matching performance. This is a major drawback with the sorting task when only accuracy is taken into account: the fact that participants can watch clips multiple times may obscure a movement advantage, particularly if the number of times participants watch the clips differs systematically across conditions, or if the difference between conditions is small. For example, as mentioned in section 6.10.3, participants in Hill and Johnston’s (2001) sorting experiment may have watched the rigid movement clips more often than the combined or non-rigid clips, leading to a rigid movement advantage.

To minimise the effect of participants systematically watching some clips more often than others, this chapter introduced the views and accuracy-per-views measures. Throughout the three experiments and the combined analysis in this chapter, several findings suggested participants were compensating for difficult conditions by watching the clips more often – for example, the results from the overall analysis of views suggest that static clips and clips in the non-rigid movement condition were watched more often than moving clips and those in the combined and rigid movement conditions; and the accuracy and accuracy-per-view results for movement type and familiarity differed significantly in all three experiments. There are several reasons why participants may have watched the clips in one condition more than in another: they may require more time or exposure to extract and compare the information in the clips, or they may have been less certain about the clips in one condition than in another. Either way, the fact that moving clips overall were watched fewer times than static clips, but were sorted with a similar level of accuracy, indicates that movement does give extra identity information and/or make participants more confident about their sorting decisions. The fact that accuracy and accuracy-per-view can give rise to slightly different results indicates that participants may watch the clips in some conditions more often than others, possibly due to difficulty or uncertainty (or both), but these differences might not be apparent in a straightforward accuracy measure. As mentioned in the pilot study, not many studies have examined the effect of movement on measures other than accuracy, and those that have (priming and visual search tasks, e.g., Lander & Bruce, 2004; Pilz et al., 2006; Pilz et al., 2009; Thornton & Kourtzi, 2002) only used static test images, thus limiting the origins of the movement advantage to the effect of social signals and
possibly structure-from-motion. As such, this is one of the first studies to
demonstrate that movement does not simply increase accuracy when identifying or
matching faces, and the first study to establish that the presence of characteristic
movement patterns (separate from structure-from-motion) allows participants to sort
moving faces with less exposure than static faces. Consequently, the sorting task, and
in particular the views and accuracy-per-view measures, are an effective way of
measuring the movement advantage and taking into account the different ways
movement can help face matching.

6.10.5 Conclusions from Experiments 5, 6, and 7

Experiments 5, 6, and 7 confirmed that it is possible to match famous and
unfamiliar faces based on characteristic movement patterns. Famous faces are
matched relatively well from static rigid head poses, but show a significant
movement advantage when non-rigid or combined rigid and non-rigid movements
are present. Famous faces also appear to benefit more from the presence of detailed
articulatory information than eye movements. Similar to famous faces, unfamiliar
faces are sorted best when both rigid and non-rigid movement information are
present, but unlike famous faces, the movement advantage for unfamiliar faces is
very tenuous, and there is no clear advantage for static rigid poses or isolated non-
rigid movement. Taken together with the results from Chapter 5, this suggests that
the benefit of movement for unfamiliar faces is primarily based on structure-from-
motion, with a relatively small contribution of characteristic movement patterns.
Unlike famous faces, unfamiliar faces were sorted best when eye movements, rather
than mouth movements, were present. However, this is likely a result of a single
anomalous condition, in which participants may have adopted unusual strategies.

As reviewed above, the movement advantage for famous faces is roughly
consistent with previous research on personally familiar faces (Bruce & Valentine,
1988; Lander & Chuang, 2005). However, the benefit of non-rigid movement for
famous faces did not appear in any single experiment, and was only apparent in the
combined analysis. It is possible that famous and personally familiar faces are
processed in different ways – for example, famous faces may be recognised better
from iconic images (Carbon, 2008), whereas our experience of personally familiar
faces is less constrained and less stereotyped. Consequently, we may be better at
matching personally familiar faces compared to famous faces. However, very little
research has investigated whether famous and personally familiar face recognition differ, and none has examined the role of movement in famous and personally familiar face recognition. The next chapter presents the last two experiments of this thesis. Experiments 8 and 9 explore personally familiar and self-face matching and recognition, and examine the effects of speech type, exaggeration and timing cues.
Chapter 7

The Role of Movement in Personally Familiar and Self-Face Recognition
CHAPTER 7: THE ROLE OF MOVEMENT IN PERSONALLY FAMILIAR AND SELF-FACE RECOGNITION

7.1 Recognition of Personally Familiar and Self-Faces

Chapters 4, 5, and 6 presented a series of experiments examining the role of movement in famous and unfamiliar face matching. However, as discussed in Chapter 2, there are different types of familiar faces. This chapter examines the use of movement in personally familiar and self-face matching and recognition. The first section summarises evidence that personally familiar faces and self-faces evoke different behavioural and neural responses to famous and unfamiliar faces, and reviews research on personally familiar and self-recognition from the biological motion literature. Experiment 8 then investigates the use of rigid and non-rigid movement in unfamiliar, personally familiar and self-face matching and recognition, and whether matching is possible across different speech types. Experiment 9 investigates whether exaggeration helps or hinders unfamiliar, personally familiar and self-face sorting and recognition, and whether relative and absolute timing cues carry important identity information.

7.1.1 Personally Familiar Faces

While the majority of this thesis has used famous faces as stimuli, this chapter presents two experiments that used personally familiar faces. There were several reasons for this change – firstly, personally familiar faces offer greater control over the stimuli than famous faces. Secondly, it is unclear whether the results for famous faces can be generalised to all familiar faces, and the use of personally familiar faces offered a chance to compare the effect of movement across different types of familiarity.

As reviewed in Chapter 2, both famous and personally familiar faces are referred to as “familiar” faces. Many researchers do not discriminate between the two, and simply refer to findings for familiar faces (e.g., Hancock et al., 2000; Johnston & Edmonds, 2009). This is not surprising, since the majority of behavioural evidence suggests that famous and personally familiar faces are processed in a
similar manner. Both types of familiar face can be matched or named significantly better than unfamiliar faces (Burton, Wilson et al., 1999; also see Megreya & Burton, 2006), they elicit equivalent priming effects (Herzmann et al., 2004), and show similar effects of caricaturing (Kaufmann & Schweinberger, 2008). Both personally familiar and famous faces show composite effects indicative of holistic processing (Ramon & Rossion, 2012; Young et al., 1987), give rise to an inner-face advantage (de Haan & van Kollenburg, 2005; Ellis et al., 1979), and show relative viewpoint independence (e.g., Bruce, 1982; Eger et al., 2005).

However, several behavioural and neural findings indicate that famous and unfamiliar faces are not processed in an identical manner. Carbon (2008) found that famous face recognition is more image-dependent than personally familiar face recognition, and concluded that famous face recognition is not as robust as was previously thought – famous faces are not really processed as familiar people, but as iconic images, whereas personally familiar face recognition is relatively flexible and can readily accommodate new images. The idea that famous face recognition relies more heavily on stereotyped images is supported by Endo et al.’s (1992) study of famous and personally familiar faces. Famous faces were recognised faster from smiling than neutral images, but personally familiar faces were recognised faster from neutral than smiling or angry expressions. Once again, these findings indicate that our representations of famous and personally familiar faces may differ – our experience with famous faces might not be enough to derive completely expression-independent representations.

Personally familiar and famous faces also elicit different physiological and neural responses. Herzmann et al. (2004) measured skin conductance responses to assess the affective response to famous, personally familiar and unfamiliar faces. Personally familiar faces lead to larger, more frequent changes in skin conductance, which the authors interpreted as a reflection of higher personal significance of the personally familiar than famous faces. In the same study, Herzmann et al. (2004) recorded ERP responses to famous and personally familiar faces, and found that personally familiar faces elicited a significantly higher N250r than famous faces during a priming task. As mentioned in Chapter 2, the N250r is an ERP response thought to reflect the activation of individual face representations (Burton et al., 2011). Herzmann et al. (2004) suggested that the change in N250r amplitude reflected the fact that stronger, more widespread neural networks were involved in
the encoding of personally familiar faces compared to those for famous faces. This hypothesis is supported by several fMRI studies: Gobbini et al.'s (2004) study, which found more widespread activation for personally familiar than famous faces; and Taylor et al.'s (2009) study, which found that fusiform activation was bilateral for personally familiar faces, but isolated to the right hemisphere for famous faces.

Taken together, the behavioural and neural results suggest that there may be some differences between famous and personally familiar face processing – representations of personally familiar faces may be more robust than famous faces. However, it is unclear whether movement is one of the many factors that have the same effect on famous and personally familiar faces, or whether the movement advantage is larger (or smaller) for personally familiar faces than famous faces. A comparison of previous studies is hampered by the fact that famous face studies generally used unscripted, natural excerpts of movement (e.g., Knight & Johnston, 1997; Lander & Bruce, 2000; Lander et al., 2001; Lander et al., 1999), whereas studies of personally familiar faces either used highly scripted, staged movements (e.g., Bruce & Valentine, 1988; Rosenblum et al., 2007), or whole-body movements (e.g., Bruce et al., 2001; Burton, Wilson et al., 1999). Consequently, Experiments 8 and 9 examined the effect of movement on personally familiar and unfamiliar faces, using similar stimuli and methods to those in Experiments 1-7, with relatively natural speech movements. This allowed the pattern of results to be compared to those found for famous and unfamiliar faces in Experiments 1 – 7 (Chapters 4, 5, and 6).

It is difficult to determine whether people should show a larger movement advantage for personally familiar than unfamiliar faces. Currently, only three studies have directly compared the movement advantage for personally familiar and unfamiliar faces, and out of the three, two asked participants to generalise from moving CCTV images to high-quality photographs (Bruce et al., 2001; Burton, Wilson et al., 1999), and neither found a movement advantage. The third study (Cook et al., 2012) did not report participants’ accuracy for unfamiliar faces (this study will be discussed further in section 7.1.2). Since people are likely to have better three-dimensional representations and more experience of idiosyncratic movement patterns for familiar than unfamiliar faces, participants should be able to use structure-from-motion and characteristic movement patterns when matching personally familiar faces. In that case, participants should show a greater movement advantage for personally familiar than unfamiliar faces (similar to the results
reported in Experiments 6 and 7, Chapter 6). However, it is possible that, like famous faces, personally familiar faces can be matched well using static information, and dynamic cues are more helpful for unfamiliar faces (e.g., Experiments 1, 2, 4).

7.1.2 Self Recognition

Self-recognition can be viewed as a special case of personal familiarity. Like personal familiarity, there is a large amount of neurological evidence to suggest that self-recognition activates brain networks distinct from those involved in other face recognition. A recent meta-analysis identified four areas consistently involved in self-recognition: the left fusiform gyrus, bilateral middle and inferior frontal gyri, and right precuneus (Platek, Wathne, Tierney, & Thomson, 2008). Other researchers have also stressed the role of the right prefrontal cortex in self-processing (Keenan, Wheeler, Gallup, & Pascual-Leone, 2000).

Behavioural findings for static self-face recognition are less clear-cut than the neural results. Some studies have found a significant advantage for self-recognition compared to familiar and unfamiliar faces (Keenan et al., 1999; Tong & Nakayama, 1999), whereas other studies have found no behavioural differences between self and familiar faces (Kircher et al., 2001; Platek et al., 2006). However, when the stimuli are presented in motion, there is a consistent advantage for self-recognition over personally familiar and unfamiliar recognition. Studies involving whole-body point-light displays of walking and other whole-body movements have found that recognition or matching of oneself is better than recognition of familiar others, which is in turn better than recognition of strangers (Loula et al., 2005; Prasad & Shiffrar, 2009). Furthermore, self-recognition of point-light walkers is less viewpoint-dependent than recognition of familiar individuals in point-light walker displays (Jokisch, Daum, & Troje, 2006).

Why are people so good at recognising themselves from their movements? It is rare that we see ourselves from an external perspective, particularly when compared to highly familiar friends and colleagues. Therefore, it is unlikely that the self-recognition advantage arises from visual experience. Several authors have proposed that the ability to recognise one’s own actions arises from motor experience. Knoblich and Flach (2003) proposed that there is a close link between perception and action – viewing an action activates mental representations of that action, and viewing our own action results in a stronger activation than viewing someone else.
producing the same action. In other words, motor processes influence the visual analysis of human movement, resulting in an increased sensitivity to one’s own motion patterns (Jokisch et al., 2006; Loula et al., 2005; Prasad & Shiffrar, 2009).

Currently, only one study has tested these predictions with facial movement patterns. Cook et al. (2012) presented participants with shape-normalised avatars showing their own, their friend’s, or an unfamiliar person’s face movements while telling a joke. Participants were able to identify their friend and themselves relatively accurately, and equally well, when the faces were presented upright. Friend recognition, but not self-recognition, dropped to chance levels when the faces were inverted. In a second experiment, Cook et al. presented “anti-sequences” that preserved the timing characteristics of the movements, but reversed the rigid and non-rigid spatial characteristics of the movements (e.g., a smiling face tilted forward became a frowning face tilted backwards); they also presented stimuli with timing manipulations (staggered or slowed movements) that preserved the spatial characteristics, but not temporal cues, of the original sequences. The results for the anti-sequence condition replicated the inverted condition: self-faces were discriminated better than chance, while friends’ faces were not. Staggered and slowed movement, however, reduced self-face recognition (and personally familiar face recognition) to chance levels. Performance was not reported for unfamiliar faces in any condition – the results were calculated using an unusual d’ measure that only factored in unfamiliar faces when they were mistaken for self or friend faces. Consequently, it is difficult to tell whether movement-based self-recognition is significantly better than unfamiliar face recognition.

Cook et al.’s (2012) results are not the first to show that temporal characteristics, like the rhythm and speed of the movement, are important for self-recognition – for example, Flach, Knoblich, and Prinz (2004) showed that people were better at identifying auditory sequences of their own clapping rhythms when the original tempo was preserved. However, Cook et al.’s results are the first to show that one’s own facial movements can be identified based almost solely on temporal cues, and these cues can be extracted from local movement cues (as indicated by the fact that inversion did not impair performance). On the other hand, personally familiar face movements appear to be identified based on a combination of spatial and temporal cues (as shown by decreased performance in the anti-sequence and timing-disrupted conditions), and extraction of these cues requires the face to be
upright. The results for personally upright and inverted personally familiar faces strongly resembled the results from Hill and Johnston’s (2001) odd-one-out task (which presented unfamiliar faces upright and inverted), suggesting that unfamiliar and personally familiar faces both rely, at least partially, on configural or holistic processing of the face (Maurer et al., 2002).

Cook et al.’s (2012) findings imply that visual familiarity with a face (i.e., personally familiar faces) leads to significantly different outcomes than motor familiarity with a face (i.e., one’s own face). Experiments 8 and 9 therefore focused on comparing unfamiliar, personally familiar and self-face matching. For the first time in this thesis, participants were also asked to complete a recognition task (i.e., say who the face was). Based on the results from whole-body and facial movement recognition, self-faces were expected to be matched and recognised better than personally familiar or unfamiliar faces overall.

### 7.2 Experiment 8: Self, Personally Familiar and Unfamiliar Face Matching From Rigid and Non-Rigid Movements

Experiment 8 was designed to investigate the effect of different types of movement (rigid, non-rigid, and combined rigid and non-rigid) in unfamiliar, personally familiar, and self-face recognition. As discussed in Chapter 6, previous research has found that personally familiar faces can be identified from either rigid (Bruce & Valentine, 1988) or non-rigid movements alone (Bruce & Valentine, 1988; Lander & Chuang, 2005; Rosenblum et al., 2007), whereas unfamiliar faces are sorted or matched best from isolated rigid movements (Hill & Johnston, 2001; Watson et al., 2005). However, Experiments 5, 6, and 7 (Chapter 6) investigated the role of different types of movement in famous and unfamiliar face sorting, and found that combined movement was the only source of a movement advantage, and only famous faces were sorted better from moving than static images overall (although

---

unfamiliar faces showed a significant movement advantage for combined movement in isolated conditions). This suggests that characteristic movement patterns are carried in both rigid and non-rigid movements, and prior familiarity with the movement patterns can make it easier to match or sort faces based on movement alone.

Experiment 8 addressed a similar question to Chapter 6: do different types of movement (rigid, non-rigid, combined rigid and non-rigid movement) contribute to unfamiliar, personally familiar and self-face recognition? Unlike Chapter 6, though, the stimuli used in Experiment 8 were not shape-normalised – that is, participants could match the stimuli using structural cues, as well as characteristic patterns of movement. Experiment 8 also reverted back to a modified version of the original same/different matching task used in Experiments 2, 3, and 4. The task and stimuli differences, and the fact that personally familiar (not famous) faces were used in Experiment 8, meant that the predicted effect of movement type was unclear.

Given the extra shape and structural cues in the stimuli of Experiment 8, it was expected that unfamiliar faces would once again show a movement advantage (as in Experiments 1a, 2a, and 4). Unfamiliar face matching was also expected to mirror the findings in the combined analysis of Experiments 5, 6, and 7 (Chapter 6), with an overall advantage of combined movement compared to either rigid or non-rigid movement in isolation. However, if the movement advantage for unfamiliar faces is based primarily on structure-from-motion, there was also a possibility that rigid and combined movements (which both contain equal structure-from-motion cues) would be matched equally well and give rise to an equal movement advantage. Personally familiar faces were expected to be matched and recognised best from combined and rigid movements, and worst from non-rigid movements (as in Chapter 6 for famous faces; cf. Lander & Chuang, 2005). To date, no studies have examined the effect of movement type on self-face recognition or matching. Therefore, it was unclear whether movement type would have a significant effect for self-face matching and recognition; however, there is no reason to believe that increased sensitivity to one’s

---

13Although this is Experiment 8 in the thesis, it was the first experiment conducted during my candidature. Consequently, the methods (appropriate length of clips, sequential presentation of clips, and, most importantly, the treatment of static control trials) were based on several previous studies using biological motion, and a small pilot test of the stimuli, rather than on the results of Experiments 1 – 6.
own motor patterns (Knoblich & Flach, 2003) is isolated to rigid or non-rigid movements.

While several studies have tested the effect of movement type on movement-based face recognition, there has been little research on whether the way someone is speaking (which will be referred to as speech style for the remainder of the chapter) can also affect recognition. Lander et al. (2007) manipulated the “manner” of speech in a cross-modal matching task (face-to-voice or voice-to-face) – that is, they asked participants to speak conversationally, casually (in a rush), or clearly (Helfer, 1997), or to phrase the sentence as a question. Participants could match faces to voices and vice versa at above-chance levels, but participants performed better when the manner of speaking was the same in both items compared to when the manner of speaking was mismatched. Lander et al. (2007) also established that the effects of manner were not due to absolute speed or duration of the speech – instead, they proposed that the detrimental effect of changing manner arose because the patterns of stress and intonation that contain meaning (prosody) probably contain cues to identity, above and beyond the characteristic or idiosyncratic timing cues that are carried across all speech tokens.

Lander et al.’s (2007) findings indicate that matching or mismatching the style of speech can have a significant impact on identification performance, even for faces we have never seen before (Lander et al., 2007, used unfamiliar faces). Previous studies have investigated whether different types of non-rigid movement can affect movement-based recognition (e.g., Lander & Chuang, 2005), but the effect of speech style has not been examined in a purely visual matching task. It is possible that familiarity reduces the effect of changing speech style – as someone becomes more familiar with a person, they may be better at extracting the less variable movement patterns, and focus on the characteristic movement patterns that occur across speech styles. This may be particularly true for self-faces, where we have experience producing different speech styles as well as perceiving them (either visually or through listening to our own speech).

To address the effect of movement type and speech style on unfamiliar, personally familiar and self-face recognition, Experiment 8 tested recognition and matching of moving and static PLDs. The PLDs were shown speaking normally (during a casual conversations with the experimenter), telling a scripted joke (similar to Hill & Johnston, 2001), or reading some dramatic dialogue. Participants
completed a same/different identity matching task, similar to that used in
Experiments 2, 3, and 4, and a recognition task. In the matching task, the two PLDs
in each trial either showed different excerpts of the same speech type (two jokes, two
dialogue excerpts, or two candid speech excerpts), or excerpts of two different speech
types (joke-dialogue, candid-joke, dialogue-candid). As in Chapter 6, the type of
movement was manipulated: participants matched or identified PLDs with combined
rigid and non-rigid movement, rigid movement only, or non-rigid movement only.
Based on Lander et al.’s (2007) findings, it was expected that mismatching the
speech style within a trial would have a detrimental effect on matching performance,
although this effect was predicted to decrease with familiarity.

7.2.1 Methods of Experiment 8

7.2.1.1 Participants. Fifteen actors\textsuperscript{14} (9 female, $M$ age: 32 years, range: 22 to
60), including the author, had their facial movements recorded for this study. Actors
were acquaintances of the authors who took part in the study for financial
compensation. Of the 15 actors, 11 also took part in the perceptual testing (5 female,
$M$ age: 27 years, range: 25 to 60). The remaining four actors served as familiar and
unfamiliar stimuli for the other participants. All participants reported normal or
corrected-to-normal vision.

7.2.2 Stimuli and Materials. Movement sequences were captured using an 8-
camera Vicon motion capture system (Oxford Metrics), calibrated to less than 1mm
error. The motion capture system recorded data at 200Hz from 28 markers placed on
the face (Figure 21 illustrates the configuration used). Any errors or gaps in data
collection were corrected using a spline algorithm.

Each participant was recorded telling a joke, reading some dramatic dialogue
(excerpt from Hitchhikers Guide to the Galaxy, the Original Radio Plays) and
engaging in casual conversation. Each speech style was repeated twice, to ensure that
videos presented in the final experiment never showed exactly the same movement
sequence. For each participant, two clips of between 4 and 10 seconds were selected
to represent each speech style (i.e., six clips in total).

\textsuperscript{14} For the sake of clarity, the participants who took part in the recordings will be referred to as
“actors” in Experiments 8 and 9. None had formal acting training.
For each clip, four PLD stimuli were created, displaying natural (rigid and non-rigid) motion, rigid motion only, non-rigid motion only, and a single static image. The PLDs were created using Vicon BodyBuilder (Oxford Metrics) and Matlab (Mathworks). To create the combined rigid and non-rigid movement videos, the motion capture data were processed in Matlab to present the movement from a frontal view. To create the rigid and non-rigid motion videos, the movement of the forehead, temples and nose bridge was isolated using Vicon BodyBuilder (Oxford Metrics), and used to approximate rigid head movement. For the rigid movement videos, all non-rigid markers (i.e., all markers other than the four mentioned above) were “frozen” in a neutral pose (i.e., neutral expression, with a closed mouth, extracted from the beginning of each clip). The neutral pose was then superimposed onto the rigid head movements in each clip. For the non-rigid videos, the rigid head movements were removed from all the marker data, leaving the non-rigid face movements intact. Examples of the PLDs used in Experiment 8 are included in Appendix A. To create the static images, the 32\textsuperscript{nd} frame of each combined motion video was extracted using Matlab. In all videos and static images, the first frame was centred on the screen using the nose-bridge marker.

The PLDs were rendered as black dots on a white background, with a 500ms black screen before and after the clip. After editing, each video measured 560 x 420 pixels, and each PLD measured approximately 5.5 cm across onscreen. The
experiments were run using SuperLab 4.07 (Cedrus). All participants were tested on a MacBook Pro 15 inch laptop, screen resolution 1440 x 900.

7.2.3 Design and procedure. Participants were tested between 4 and 6 weeks after the initial capture session. They were informed that they would be shown PLDs of themselves, their familiar partner (who they knew the identity of), and an unfamiliar person. Familiar pairs consisted of people of the same gender who had known each other for at least one year (min 14 months, max 28 years). The unfamiliar faces were also matched for gender and approximate age. All eleven participants completed the matching task, but due to computer problems, only ten participants completed the recognition task. The matching task was run before the recognition task, because it was assumed that extra exposure to the stimuli would not help people name the PLDs, but may help people match them. Participants received a 10 min break between tasks.

7.2.3.1 Same/different matching task. The matching task had three within-subjects variables: familiarity (self, familiar, and unfamiliar), movement type (natural, rigid, non-rigid, and static) and speech style (matched or mismatched). In matched speech style trials, two clips of the same type were played (e.g., two jokes), whereas for mismatched trials, the videos depicted two different speech styles (e.g., joke and dialogue). In total, each participant completed 144 trials, half of which were “same” trials.

In the matching task, each trial consisted of two PLDs, both moving or both static, presented simultaneously on the screen (see Figure 21), separated by 70 pixels (approximately 3.5cm) horizontally. The participant was asked to indicate via key press whether the images were of the same person or different people. “Different” trials included all possible identity pairs (i.e., self with familiar, self with unfamiliar, familiar with unfamiliar) \(^{15}\), but only one different pairing was presented in each condition to ensure even numbers of “same” and “different” trials – for example, for combined movement, matched speech style condition, either the self-familiar or self-unfamiliar pair was shown. The choice of “different” pair for each condition was

\(^{15}\) Note this is different to Experiments 2, 3, and 4 (Chapters 4 and 5), in which the “different” pairs always consisted of the same pairs of people.
counterbalanced. Each trial repeated itself every 10 s until a response was recorded, and participants could respond at any time throughout the trial.

Order of trial presentation was randomised for all variables but movement type, which was blocked. Combined, rigid and non-rigid motion blocks were counterbalanced between participants. The static images were always presented last, because the results and feedback from pilot participants indicated that participants were trying to extract characteristic static cues (e.g., head shape, head tilt) from the static trials, and then use these cues, instead of movement-based information, to match the moving clips. By presenting the static images last, it was hoped that participants would be encouraged to pay attention to the movement, rather than static cues. Participants viewed six practice trials (combined rigid and non-rigid movement, depicting faces that did not appear elsewhere in the experiment) before beginning the first block.

7.2.3.2 Recognition task. The recognition task had two variables: familiarity (self, familiar, and unfamiliar) and movement type (natural, rigid, non-rigid, and static). Speech style also varied (casual, joke, and dialogue), but since there were only two trials per speech type, and no theoretical reason to believe that speech type would affect recognition, speech type was excluded from analysis. There were 72 trials in total per participant.

The videos used in the recognition task were identical to those in the matching task. In each trial, participants were shown a single PLD, centred on the screen, and asked to indicate via key press whether the PLDs depicted their own face, their familiar partner, or an unfamiliar person. Once again, participants could respond at any point during a trial. However, the video was only presented once, not looped as in the matching task.

As in the matching task, the order of presentation of trials was randomised for all variables but movement type, which was blocked. The order of presentation of combined, rigid and non-rigid movement trials was counterbalanced, but static trials were always presented last. Participants did not receive any practice trials, as they already had experience with the stimuli from the matching task.

7.2.2 Results for Experiment 8

7.2.2.1 Matching task. Three dependent variables were analysed for the matching task: $d'$, hits, and bias (see Chapter 4, section 4.3.2, for an explanation of
the bias statistic used throughout this thesis). Because there is only one self-face for each person (as opposed to multiple familiar or unfamiliar faces), all the “different” trials contained stimuli from different levels of familiarity (this is in contrast to Experiments 2-4, where all the different trials contained either famous or unfamiliar faces, but never both). Errors in the different trials could have resulted from the misperception of either identity, so it is not practical to assign the different trials to a particular level of familiarity. Consequently, false alarm rates (FA) were pooled across all levels of familiarity for each level of movement and speech style (e.g., FA were calculated for all the combined movement, matched speech style trials, regardless of familiarity). To obtain a d’ and bias score for each condition, the pooled FA rates were combined with individual hit rates for each level of familiarity (e.g., the pooled FA for combined movement, matched speech style was combined with the hit rate for combined movement, matched speech style for self, familiar and unfamiliar trials). Hit scores of 0 or 1 were replaced with 0.167 and 0.833, respectively; FA of 0 were replaced with 0.083 (MacMillan & Kaplan, 1985). In this experiment, a d’ of 0 represents chance performance, while a d’ of 3.04 represents perfect performance in a condition.

Separate 3-way within-subjects ANOVAs (3 x familiarity; 4 x movement type; combined, rigid, non-rigid, static; 2 x speech style; matching/mismatched) were carried out on each measure. Due to sphericity violations and the small sample size, multivariate tests (Wilks’ Lambda) are reported. All pairwise comparisons were Bonferroni-corrected.

7.2.2.1.1 Signal detection theory analysis. The d’ scores for the matching component of Experiment 8 are shown in Figure 22. Two-tailed t-tests of the d’ scores revealed that all conditions were matched significantly above chance levels, all ps < .01. The ANOVA showed no significant main effects or interactions, familiarity: $F(2,9) = 0.29, p = .753, \eta_p^2 = .06$; movement type: $F(3,8) = 1.23, p = .361, \eta_p^2 = .31$; speech style match: $F(1,10) = 1.50, p = .248, \eta_p^2 = .13$; all interactions $ps > 0.3$. Planned pairwise comparisons on familiarity and movement type were non-significant, familiarity: all $ps = 1$; movement type: all $ps > .3$. 

239
Figure 22: \(d'\) results from the matching task in Experiment 8. 

**a)** matching speech styles; **b)** mismatched speech styles. White bars represent self-faces; light grey bars represent personally familiar faces; dark grey bars represent unfamiliar faces. Error bars represent +/- 1 standard error of the mean.

### 7.2.2.1 Hits analysis

The ANOVA on hit rates also showed no main effect of familiarity or movement type, familiarity: \(F(2,9) = 0.87, p = .917, \eta_p^2 = .02\); movement type: \(F(3,8) = .835, p = .512, \eta_p^2 = .24\), but there was a significant main effect of speech style match, \(F(1,10) = 6.89, p = .025, \eta_p^2 = .41\).

Participants were more accurate at matching PLDs when they showed the same speech style, \(M = .869\), than when the speech styles were different, \(M = .780\). There was a near-significant interaction between speech style match and movement, \(F(3,8) = 3.91, p = .055, \eta_p^2 = .59\), driven by the fact that combined movement was matched
significantly worse in the mismatched than in the matched condition, $p = .004$, but all other movement types were matched equally well regardless of whether the speech style was matched or not, $ps > .3$. No other interactions approached significance, $ps > .1$. Planned pairwise comparisons on the familiarity and movement type variables were non-significant, familiarity: all $ps = 1$; movement type: all $ps > .8$.

**7.2.2.1 Bias analysis.** Participants were relatively neutral in their response bias throughout the experiment, $c = .02$, SD = 0.37. The ANOVA on bias showed no main effect of familiarity or movement type, familiarity: $F(2,9) = 0.10, p = .902$, $\eta_p^2 = .02$; movement type: $F(3,8) = .251, p = .859$, $\eta_p^2 = .09$, but there was a significant effect of speech type match, $F(1,11) = .954, p = .011$, $\eta_p^2 = .49$. Trials with a matching speech type elicited a slight bias towards a “same” response, $c = -.05$, whereas trials with mismatched speech types elicited a small bias towards a “different” response, $c = .10$. This may explain the significant effect of speech style match in the hits analysis – participants were simply more biased towards reporting an identity match in the matching speech style condition. No interactions were significant, $ps > .05$. Planned pairwise comparisons on the familiarity and movement type variables were non-significant, all $ps = 1$.

**7.2.2.2 Recognition task.** Three analyses were carried out on the recognition data. First, accuracy scores were calculated and subjected to a two-way within-subjects ANOVA (3 x familiarity; 4 x movement type; results collapsed across speech style). Secondly, to compare the current results to Cook et al. (2012), d’ results were calculated in the same way as in their study. Hits were correct responses to the self and familiar face stimuli (i.e., “self” response to self face, “familiar” response to familiar face), and false alarms (FA) were incorrect responses to the unfamiliar face stimuli (i.e., FA for self stimuli occurred when participants incorrectly responded “self” to the unfamiliar faces, and likewise for the familiar faces). Like the accuracy analysis, hits and FA were collapsed across speech types, to increase the number of trials contributing to each d’ statistic. Hit or FP scores of 0 or 1 were replaced with 0.08 and 0.92, respectively (Macmillan & Kaplan, 1985). In this experiment, a d’ of 0 represents chance performance, while a d’ of 2.79
represents perfect performance in a condition. The d’ scores were subjected to a two-way within-subjects ANOVA (2 x familiarity; 4 x movement type). Finally, the distribution of “self”, “familiar” and “unfamiliar” responses for each participant was collapsed across speech type and movement type and subjected to a chi-squared analysis to determine whether participants were responding randomly in the recognition task.

### 7.2.2.2.1 Accuracy

The accuracy scores for the recognition component of Experiment 8 are shown in Figure 23. Participants performed poorly overall – two-tailed t-tests revealed that accuracy in the majority of conditions was not significantly different from chance (33.3%). Unfamiliar faces in the rigid movement condition were recognised better than chance, \( p = .034 \), whereas familiar faces in the combined movement condition and self-faces in the static condition were recognised significantly worse than chance levels, \( ps < .05 \).

![Recognition accuracy](image)

**Figure 23:** Proportion correct in the recognition task in Experiment 8. White bars represent self-faces; light grey bars represent personally familiar faces; dark grey bars represent unfamiliar faces. Chance performance was 0.33, indicated by the dashed line. Error bars represent +/- 1 standard error of the mean.

---

16 The maximum d’ score is lower than in the same/different tasks because a different formula was used to calculate d’ for same/different tasks and the current experiment. Formulas for same/different and standard d’ were from MacMillan and Creelman (2005).
The ANOVA on accuracy revealed a main effect of familiarity, $F(2,18) = 8.78$, $p = .002$, $\eta_p^2 = .49$. Pairwise comparisons revealed that unfamiliar faces were recognised significantly more accurately than self-faces, $p = .003$, and marginally more accurately than familiar faces, $p = .076$. There was also a main effect of movement type, $F(3,27) = 3.92$, $p = .019$, $\eta_p^2 = .30$, but no pairwise comparisons were significant. The interaction between familiarity and movement type was not significant, $F(6,54) = 1.11$, $p = .368$, $\eta_p^2 = .11$. Despite the significant results, it is impossible to draw any conclusions from the accuracy analysis about familiarity or the role of movement in face recognition. Participants were generally very inaccurate, and chance level performance indicates that they were generally unable to perform the task of recognising themselves or their friend from a PLD.

### 7.2.2.2.1 Signal detection theory analysis.
Like the accuracy scores, the d’ scores for the recognition task were extremely low – participants performed at chance levels in all conditions, $p_s > .1$. The ANOVA revealed no effect of familiarity, $F(1,9) = 4.50$, $p = .519$, $\eta_p^2 = .05$. There was a significant main effect of movement type, $F(3,27) = 4.18$, $p = .015$, $\eta_p^2 = .32$, but no pairwise comparisons on movement type were significant, all $p_s > .1$. There was no significant interaction between familiarity and movement type, $F(3,27) = 1.19$, $p = .904$, $\eta_p^2 = .02$. Once again, there was no evidence that participants could recognise either their own face or their friend’s face when presented as a PLD.

### 7.2.2.2.3 Chi-squared analysis.
A visual inspection of the data revealed that several participants were highly consistent in their pattern of misidentification of PLDs – that is, some participants were consistently identifying their own face as familiar and the familiar face as their own. If participants were consistently misidentifying faces, it would suggest that people were not simply guessing during the recognition task (as the chance-level performance would suggest). If participants were consistent (if incorrect) with their identification, it would suggest that people are quite good at remembering and recognising the shape or characteristic movements of a particular PLD across a large number of trials, but that this ability may not translate to overt naming ability. In other words, people seem to be able to correctly indicate that clips A and B depict the same person, and that clips C and D perhaps depict another, but they cannot say which person either face belongs to. This is in accord with the high performance across all conditions in the matching task.
To investigate whether participants showed a consistent pattern of responding across trials, the distribution of “self”, “familiar”, and “unfamiliar” responses was examined for each PLD. Responses for each participant were collapsed across speech type and movement type, and subjected to three chi-squared analyses, one per identity: self, familiar, or unfamiliar (i.e. for familiar faces, the chi-squared analysis checked to see whether there were more responses in one category than would be expected if people were responding randomly, regardless of whether the responses were correct). The results of the chi-squared analyses are shown in Figure 24. Twelve out of the thirty chi-squared analyses were significant ($p < .05$), and another four were approaching significance ($p < .1$). Some of these results (e.g., P8, P10) reflect a bias towards the “unfamiliar” response, but, out of the ten participants, the majority (seven) showed a tendency to (mis)identify at least one face consistently. For example, P1 consistently responded “unfamiliar” to his own face and “familiar” to the unfamiliar face, whereas P4’s responses were weighted towards “self” to the familiar face and “familiar” to the unfamiliar face.

### 7.2.3 Discussion of Experiment 8

Participants were extremely good at matching PLDs representing their own face, a personally familiar face, and an unfamiliar face. However, they were unable to identify the PLDs when they were shown in isolation. The results from the matching task of Experiment 8 do not show any evidence of a movement advantage, for any level of familiarity or movement type. The lack of movement advantage may be explained by the unusually high performance in the static condition. Unlike previous experiments in this thesis, the movement conditions were blocked, and static images were always presented last. This procedural change was implemented because participants in pilot versions of the experiment indicated that they tried to learn the static characteristics of the faces, and then use these static cues to match the moving stimuli. However, this strategy appeared to bring about the opposite result: presenting the static condition last allowed participants to use the moving stimuli to encode structure or shape-based information, which they were then able to apply in the static matching. This resulted in extremely high hit rate – 83.3% in the static condition – when compared with the hit rates from other matching
Figure 24: Response frequency in the recognition task of Experiment 8. Categories show the person in the PLD. Bars indicate participants’ responses: White “self”, light grey “familiar”, dark grey “unfamiliar”. Stars indicate the results of the chi-squared analyses: 1 star, $p < .1$; 2 stars, $p < .05$; 3 stars: $p < .0005$. 
experiments using PLDs: Experiment 2a, 69.6%; Experiment 2b, 57.7%; Experiment 3 (PLD only), 64.8%; Experiment 4 (PLD only), 33.9% \(^{17}\).

Alternatively, it is possible that the lack of movement advantage arose due to other changes to the methods in this study – particularly the fact that participants could match the faces simultaneously, and they had unlimited time to watch the clips. However, the hit rate in the combined movement condition of Experiment 8 – 76.8% – was roughly in line with performance in other PLD matching tasks when two moving images were present: Experiment 2a, 77.8%; Experiment 3 (PLD only), 76.2%; Experiment 4 (PLD only), 80.3%; with the exception of Experiment 2b, 57.6% \(^{18}\). Overall, then, it appears that the lack of movement advantage probably arose because participants learnt to match static structural or shape cues. This makes it hard to quantify the effect of movement in personally familiar and self-recognition, but it does support the hypothesis that moving faces provide enough information to build a robust static representation of the face and head (Roark et al., 2003).

There was only one significant result in the analysis of the matching task: a detrimental effect of mismatching speech styles was present in the hits analysis. This result lends support to Lander et al.’s (2007) proposal: some variation in movement patterns of a face and head can be accounted for by the manner in which the person is speaking, and some variation is a result of the individual’s characteristic movement patterns. The fact that matching was still well above chance levels when speech style was mismatched suggests that people are sensitive to both forms of variation, regardless of familiarity, but it is slightly easier to match a face when the two images show the same speech style.

As noted above, there were no overall effects of familiarity, and no interactions between familiarity and speech style matching: the effect of viewing clips with matched or mismatched speech styles was the same when a face was unfamiliar, familiar, or even one’s own. This suggests that visual (familiar faces) and/or motor (self-faces) experience with a range of speaking styles does not provide any extra benefit when trying to generalise across speech styles, at least on tasks where

\(^{17}\) Due to the different methods for calculating FA in Chapters 4 and 5, compared to this chapter, it is not practicable to directly compare d’ scores.

\(^{18}\) These figures were calculated from the average of all clip lengths, paired with moving comparison images, in Experiment 2. Combined movement is offered as a comparison because Experiments 2, 3, and 4 only showed combined rigid and non-rigid movements.
performance is already very good. It is possible that the effect of familiarity would be more pronounced in a slightly more difficult task.

It is particularly interesting that unfamiliar faces were matched as well as faces with which people had extensive personal experience, even in the mismatched speech type condition. However, it is unclear whether this high level of performance was due to the use of structural information, characteristic movement patterns, or static cues. Numerically, unfamiliar faces were matched better from rigid movement than in any other condition, which indicates a possible role of structure-from-motion. However, the small sample size and sheer number of pairwise comparisons meant that no differences reached significance. This leaves open the possibility, supported by the results of Experiment 4 (Chapter 5), that participants could match the unfamiliar faces based on characteristic movement patterns. Like unfamiliar faces, it is unclear whether the high levels of performance for familiar and self-faces were due to the use of structure-from-motion, characteristic movement patterns or static cues.

One thing that was clear from the results is that whatever cues were being used in the matching task were not useful in the recognition task. Performance was at or below chance levels for all conditions except one – the unfamiliar, rigid movement condition. Familiarity with a face and the movement type did not appear to help recognition; despite significant effects in the accuracy analysis, no individual categories even approached above-chance performance. Despite extremely poor naming performance, the chi-squared analyses indicated that the way participants were responding was not random. Rather, it appears that participants were consistently matching the movement stimuli in the recognition task to pre-defined, arbitrarily named people – for example, a participant might have noticed a particular pattern of movement and labelled a face “familiar” every time it appeared, even though they did not overtly recognise the movement as belonging to their familiar partner. This is a significant finding because it suggests that people can extract characteristic movement patterns and match them to those held in memory, even after some delay and intervening stimuli.

It is unclear why participants in the recognition task performed so poorly – in other studies, participants have been able to name friends from facial PLDs (Bruce & Valentine, 1988; Rosenblum et al., 2007), and identify themselves from facial avatars (Cook et al., 2012) or whole-body PLDs (Jokisch et al., 2006; Loula et al.,
Chapter 7: Personally Familiar and Self-Face Recognition

2005; Prasad & Shiffrar, 2009). The stimuli used in Experiment 8 may have presented less identifying movement information than a whole-face avatar (Cook et al., 2012) or a detailed PLD with over 50 points (Bruce & Valentine, 1988), or a PLD with a large number of points dedicated to showing the movements of the lips, teeth and tongue (Rosenblum et al., 2007). To address this concern, the stimuli used in Experiment 9 were created using the same technique as Experiment 7 (inserting the greyscale teeth and tongue from the original video into the PLD), in order to provide detailed mouth information.

Overall, Experiment 8 did not provide any evidence that personally familiar or self-face recognition were significantly better than unfamiliar face recognition, in either a matching or a recognition task. Unlike Experiment 6, there was no clear effect of movement type, and no overall movement advantage. Experiment 9 addressed some of the shortcomings of Experiment 8, and reverted to using shape-normalised stimuli in a sorting task to examine matching performance. Experiment 9 also addressed the role of temporal information and exaggeration in unfamiliar, personally familiar and self-face recognition.

7.3 Experiment 9: Natural Exaggeration and Temporal Variation in Self, Personally Familiar and Unfamiliar Face Matching

Experiment 9 was designed to extend on the findings of Experiment 8, by examining the effect of natural exaggeration and temporal constraints on unfamiliar, personally familiar and self-face sorting and recognition.

Exaggeration has often been used investigate the nature of our mental representation of faces. Studies using caricatured static images have shown that exaggerating spatial deviations from the average can make faces easier to recognise – for example, Rhodes et al. (1987) and Benson and Perrett (1994) found that caricatures of line drawings are recognised faster than veridical representations or anti-caricatures (in which deviations from the average are minimised). Photographic caricatures have a similar, but more inconsistent effect: in some cases, slight positive caricatures are rated as better likenesses of a person, and identified faster than veridical images (Benson & Perrett, 1991; but see Rhodes, Byatt, Tremewan, & Kennedy, 1997; Kaufmann & Schweinberger, 2008). The effects of caricaturing are not restricted to spatial or structural changes to the face – Lee and Perrett (2000)
found that caricaturing colour intensity resulted in higher accuracy when identifying famous faces.

Given the inconsistent results with photographic caricatures, it seems unlikely that our mental representation of faces is caricatured. Instead, Rhodes et al. (1987) interpreted the effectiveness of caricatures as evidence for a norm-based model of face encoding – that is, faces are encoded in relation to a grand average face (derived from our experience of all faces), and elements of a face that are further away from the norm are considered distinctive, and accessed more easily than typical features. In other words, it is possible that caricatures make it easier to access a stored representation of a person, by emphasising the elements of their face that easily distinguish them from other, similar-looking people (Kaufmann & Schweinberger, 2008).

It is also possible to caricature movement. Like static faces, there are multiple ways to exaggerate moving stimuli – the most common of these are spatial and temporal exaggerations. Imagine that the average person shakes their head by moving 2 cm over 5 s, whereas Person A moves their head 5 cm over 3 s. Spatial exaggeration involves exaggerating differences in the change in position, but not time. Since Person A already moves further than average, spatial exaggeration would accentuate this – the exaggerated Person A might move 7 cm over the same 5 s time period. Temporal exaggeration involves exaggerating differences in the time it takes to produce elements of a movement, but not the spatial coordinates of the movement. Person A moves faster than average, so temporal exaggeration would highlight this – the exaggerated Person A would still move 5 cm, but over 1.5 s.

Both spatial and temporal exaggeration can influence our perception of movement. Exaggerating the spatial characteristics of a movement can help action categorisation (e.g., styles of tennis serve; Pollick et al., 2001) and increase emotional intensity ratings for facial expressions (Hill et al., 2005; Pollick et al., 2003). Exaggerating the temporal characteristics of facial movement does not have any clear effect on emotional intensity ratings for facial expressions (Hill et al., 2005; Pollick et al., 2001), but temporal exaggeration of body movement might help identify an individual. Hill and Pollick (2000) familiarised their participants with unexaggerated arm movements from six individuals, then tested recognition using the original, unexaggerated movements and movements that had been temporally exaggerated. Positive exaggerations, which increased differences from the average,
increased recognition performance, whereas negative exaggerations, which reduced differences from the average (i.e., anti-caricatures), were recognised worst overall.

In general, these findings suggest that caricaturing or exaggerating movement information can help us categorise the movement or identify the person producing it. Like static face caricatures, an advantage for exaggerated movements suggests that we may store information about characteristic movement patterns (for emotions, tennis serves or even identities) in a norm-based way – that is, the movements are encoded in relation to average spatial and temporal characteristics, and exaggerating the task-relevant aspects of movement – emotional, category-specific or individual – enhances our categorisation or recognition performance. This does not suggest that we have a stored movement pattern for every possible biological movement, merely that exaggerating the task-relevant aspects of the movement (e.g., exaggerating the aspects of movements that discriminate between happy and neutral, or between Person A and Person B) might help performance in a range of movement-based tasks (Hill et al., 2005).

One aspect of caricaturing and exaggeration that is unclear is the effect of familiarity. Benson and Perrett (1991) found a correlation between familiarity and ratings of likeness for famous faces: for more familiar faces, people rated higher levels of caricature as better likenesses of the person. Kaufmann and Schweinberger (2008) found that positive caricatures of famous faces were recognised better than anti-caricatures (unlike Benson & Perrett, 1991, veridical images were considered better likenesses than caricatures), but there was no effect of caricature for unfamiliar faces. These results seem to indicate that caricatures are more important for familiar than unfamiliar faces, perhaps because we are more familiar with the distinctive elements of familiar faces. However, Kaufmann and Schweinberger (2012) found that caricaturing faces during the learning phase of an experimental familiarity study increased subsequent identification performance with different images of the same people. This implies that caricaturing is also important for experimentally familiar faces, and possibly for unfamiliar faces that appear repeatedly throughout the course of an experiment. Caricaturing can help us learn a new face, possibly by emphasising the distinctive elements of its shape. Currently, no studies have compared familiar and unfamiliar faces using moving caricatures, and no studies have examined the effect of movement exaggeration on self-recognition.
To test the effects of exaggeration on unfamiliar, personally familiar and self-face recognition, participants completed two tasks: a modified version of the sorting task used in Experiments 5, 6, and 7, which measured the ability to match faces; and a recognition task, which measured the ability to identify a person based on their movement alone. Participants were expected to perform well in the sorting task, especially when sorting personally familiar and self-faces. Previous research on bodies (Jokisch et al., 2006; Loula et al., 2005; Prasad & Shiffrar, 2009) and the one study that researched self-recognition using moving faces (Cook et al., 2012) suggested that participants should be more accurate at recognising themselves than personally familiar people overall, and it was expected that this result would be reflected in both tasks used in Experiment 9. The results in Experiment 9 were expected to diverge from the recognition task in Experiment 8 (which showed poor recognition overall) because the stimuli in Experiment 9 (i.e., shape-normalised stimuli, limited clip length, relatively detailed movement cues in the mouth area) more closely resembled those used by Cook et al. (2012) than those used in Experiment 8.

Most previous studies on movement exaggeration have used techniques based on measuring and artificially altering the trajectory of points in a PLD. Experiment 9 took a different approach: the stimuli were created using natural exaggerations that could occur in everyday life. Natural exaggeration of body movements has been used to investigate emotion recognition – Atkinson, Dittrich, Gemmell, and Young (2004) found that emotions were identified more accurately and rated as more intense from exaggerated body movements than normal movements. To date, no other studies have applied this technique to faces, or used natural exaggerations in identity-based experiments. To generate naturally exaggerated movements, actors were overtly asked to exaggerate their movements (Atkinson et al., 2004) and covertly encouraged to exaggerate their movements by introducing noise to the recording environment (speech-in-noise, Kim, Davis, Vignali, & Hill, 2005). Both methods of exaggeration were expected to change the spatial and temporal properties of the movements (see Kim et al., 2005, for a description of the spatial changes during speech-in-noise). However, since only one study has examined the effects of movement exaggeration on identification (Hill & Pollick, 2000, temporal exaggeration of arm movements), it was unclear whether either form of exaggeration (exaggerated movements or speech-in-noise) would help or hinder sorting and recognition.
Asking actors to exaggerate might enhance distinctive elements of their characteristic movement patterns (e.g., Atkinson et al., 2004). If the distinctive aspects of characteristic movement patterns are encoded or exaggerated in our memory of the way a person moves, or exaggeration helps us access our memory of that person (Rhodes et al., 1987), then exaggerated movements that emphasise these distinctive elements should help performance, regardless of familiarity. On the other hand, if characteristic movement patterns are stored veridically – that is, the exact spatial or temporal properties are important – exaggeration could mask characteristic patterns of movement, by distorting normal spatial or timing cues. Alternatively, people may tend to exaggerate their movements in a similar way, which could result in movements that are less distinctive or individual than the original versions. In either of these cases, normal movements were expected to be matched and recognised better than exaggerated movements for familiar faces and possibly self-faces, but not unfamiliar faces (which would have no stored characteristic movement patterns to compare to).

It is unclear where speech-in-noise falls in these predictions. Speech-in-noise might be a better natural exaggeration than asking people to deliberately exaggerate their movement, simply because it occurs frequently in real life – people may already be familiar with the exaggerated movements they and their friends produce when talking in a noisy environment. In that case, speech-in-noise might be the best of both worlds: exaggerating the distinctive elements of our movements, but also lining up with our representation of how a person moves in real life. If this is the case, speech-in-noise was expected to result in better performance than either normal or exaggerated speech. On the other hand, it is possible that speech-in-noise acts as an anti-caricature, forcing participants to constrain their characteristic movement patterns in favour of clear, stereotyped speech production. In that case, speech-in-noise would be expected to elicit poor performance for all levels of familiarity.

In addition to testing the effect of exaggeration, Experiment 9 examined the effect of restricting the amount of temporal variability in speech movements. Many studies of movement and face recognition have concluded that relative timing information is an important cue – for example, performance in matching and identification tasks is impaired if the rhythm of the movement is staggered (e.g., Lander et al., 1999; Rosenblum et al., 2002) or unnatural (e.g., Lander et al., 2006). Some studies have also suggested that absolute timing cues, such as speed, are also
important: speeding or slowing movement can also hinder face recognition (Lander & Bruce, 2000, 2004; Lander et al., 1999; cf. Lander et al., 2007). Research suggests that timing cues, both relative and absolute, are also extremely important during self-recognition – for example, Cook et al. (2012) found that self-recognition was reduced to chance levels when the movements were staggered or slowed (also see Flach et al., 2004). Self-recognition can also be impaired when timing cues are reduced naturally. Loula et al. (2005) and Sevdalis and Keller (2010) found that more constrained movements, such as walking and clapping, were harder to identify as one’s own than dancing, which contains less constraints on timing and movement.

To investigate the role of absolute and relative timing cues in unfamiliar, personally familiar and self-face recognition, Experiment 9 systematically manipulated the availability of timing cues by asking participants to speak (no timing restrictions), sing a familiar melody (relative timing cues restricted, but no absolute timing restrictions), and sing the melody in time with a metronome (relative and absolute timing restrictions). If characteristic identity information is contained in the rhythm and tempo of our speech (relative and absolute timing cues), participants were expected to perform best when viewing normal speech, followed by singing, followed by singing to the metronome. Placing more constraints on the movement was expected to have a particularly strong impact on self-recognition (Cook et al., 2012; Loula et al., 2005; Sevdalis & Keller, 2010). If, on the other hand, temporal variations are not informative in movement-matching, or are contained in individual phonemic units as opposed to spread across words and sentences, constraining rhythm and timing should not have a significant impact on performance. All movement conditions were expected to elicit better performance than the static condition.

Like Experiments 5, 6, and 7, Experiment 9 was concerned with the role of characteristic movement patterns, rather than the contribution of structure-from-motion information to recognition. Furthermore, the extremely high rates of static face matching in Experiment 8 suggested that people were using shape-based cues to match their own, their friends’, and unfamiliar faces. Consequently, the stimuli in Experiment 9 were all shape-normalised PLDs with mouths (PLDm, produced in exactly the same way as those in Experiment 7).
7.3.1 Methods of Experiment 9

7.3.1.1 Participants. As in Experiment 8, participants who took part in the recordings will be referred to as actors, whereas those who took part in the perceptual experiments will be referred to as participants. Twenty-one actors (11 female, $M$ age: 26.1 years, range: 18 to 56) took part in this study. Actors were acquaintances of the author who took part in the study for financial compensation. Recordings from three actors were excluded because they had distinctive teeth. Of the remaining 18 actors, 10 also took part in the perceptual testing ($6$ female, $M$ age: 31.4 years, range: 24 to 56). The remaining eight actors served as familiar and unfamiliar stimuli for the other participants. All participants reported normal or corrected-to-normal vision.

7.3.1.2 Stimuli and Materials. The video stimuli for this experiment were captured using a FlipVideo (Mino) camera. The resulting videos were 640 x 480 resolution, recorded at 25 fps. Actors were filmed speaking to the experimenter, who was positioned behind the camera. The videos displayed the actor’s head and shoulders on a blank background from a distance of approximately 2 m. Each recording began with the actor holding a neutral facial expression, mouth closed, in the centre of the frame. Sound was recorded, but edited out in the stimulus creation process.

Each actor was recorded saying the words to the song *Twinkle Twinkle Little Star*. Actors recorded the song in three different exaggeration conditions: normal speech, speech-in-noise, and exaggerated speech. In normal speech, actors simply spoke or sang the lyrics to the experimenter. In the speech-in-noise condition, a two-speaker sound-system playing multi-talker noise was used at approximately 70 dB was

---

19 One of the participants also took part in Experiment 8. However, repeating the analyses without her data did not change the results, so all participants were included in the final analysis.

20 The recordings used in Experiment 9 were a small part of a larger recording session, which included a variety of singing and speaking conditions and lasted approximately 45 minutes. Consequently, it is unlikely that participants remember details of any idiosyncratic movements they produced while being recorded.

21 A recording of many people producing unintelligible speech sounds (also referred to as babble speech), which replicates the effect of being in a very crowded room. The same babble speech was used in Kim et al. (2005).
directed towards the actor. Actors were instructed to make themselves understood above the noise – if the experimenter could not clearly understand what they were saying, the recording was repeated. In the exaggerated speech condition, actors were not required to speak over noise, but were instructed to exaggerate their movements by making them “bigger” and “more obvious”. Participants who did not understand these instructions were told to imagine they were trying to speak to someone on the other side of a large room, but they were not able to raise their voice.

In each of the exaggeration conditions, participants were recorded speaking in three different styles (referred to as “speech type” in the results): speaking the lyrics, singing the lyrics, and singing the lyrics in time with a metronome set to 120 beats per minute. In each condition, the recordings were repeated several times to ensure the actors were not speaking in the rhythm of the song and were singing in time with the metronome. This resulted in nine excerpts per actor, each ranging from 10 s to 20 s in duration. The excerpts were edited in iMovie (Apple) to ensure that all actors’ heads were centred onscreen, and all measured approximately 9-10 cm across in the initial frame. Four 2 s (50 frame) clips were extracted from each excerpt, beginning on the tenth frame (several participants paused at the beginning of the recordings, and this ensured all excerpts contained at least 1 s of speaking). The clips did not overlap, and the end of one clip was separated from the beginning of the next clip by at least 10 frames.

Shape-normalised PLDm stimuli were created using the same custom written tracking program that was used to create the avatars in Experiments 4-7, Chapters 5 and 6 (Saragih et al., 2010, 2011). In this case, the same PLDs with mouths (PLDm) were used as in Experiment 7 – that is, the inner mouth region from the original video (depicting the teeth and tongue, but not the lips) was converted to greyscale.

Note that the centring and size-normalising process was not repeated for each clip, due to the constraints of the tracking program used. If participants leaned towards or away from camera (or side to side), the changes in face size and screen position were reflected in the clips, and consequently in the PLDs. One solution was to resize the final PLDs – this had the side effect of changing the relative size of the dots, which participants then used as a default identity cue (this was tested in trial experiments, using two participants not involved in any other element of the recording or experimentation). The final PLDs therefore showed some slight size and screen position differences (the maximum change was less than 2.5 cm), but as these reflected natural movement patterns, they were left in the stimuli.
and inserted into each shape-normalised PLD. Examples of the PLDm clips used in Experiment 9 are included in Appendix A. Four static clips were created for each level of exaggeration. To create the static clips, four frames were selected semi-randomly from each speech type (speaking, singing and singing with the metronome), the only constraint being that all three speech types had to be represented in the collection of static images. These static frames were converted into 2 s clips using Matlab. Static images were not produced separately for each speech type for two reasons: first, because there was no reason to believe that speaking, singing and singing with a metronome would produce different characteristic static poses, whereas different levels of exaggeration may offer different static cues (for example, people might consistently open their mouth wider or tilt their head more in the exaggerated speech condition); and second, adding a static condition for each speech type would have made the experiment unfeasibly long.

In total, there were 36, 2 s video clips for each actor (four each speaking, singing, singing with metronome, and static, for each of three exaggeration levels). Due to a difference in the Matlab scripts used to render the videos in Chapter 6 and this chapter, each video in Experiment 9 measured 1124 x 624 pixels, with a 2 cm grey border around the black background (the background itself was the same size as used in Chapter 6). The PLDm measured between 5.5 and 8cm across. All videos were presented at 25 fps. The sorting experiments were run using a custom-written program on Lenovo T500 laptop computers, running Windows XP, with screen resolution set to 1280 x 800 pixels. The recognition experiments were run using SuperLab 4.05 (Cedrus). All participants were tested on a MacBook Pro 15 inch laptop, screen resolution 1440 x 900.

7.3.1.3 Design and procedure. Participants were tested between 2 and 10 months after the initial capture session. They were informed that they would be shown PLDm of themselves, their familiar partner (who they knew the identity of), and an unfamiliar person. Familiar pairs consisted of people of the same gender who had known each other for at least one year (min 12 months, max 7 years). The unfamiliar faces were also matched for gender and approximate age. The recognition task was run before the sorting task, to check the possibility that consistent responses in the chi-squared analyses in Experiment 8 arose due to prior exposure to the
stimuli. Participants received a 10 min break between tasks. For the sake of consistency with Experiment 8, the sorting task is presented before the recognition task in the methods and results.

7.3.1.3.1 Sorting task. The sorting task had three within-subjects variables: familiarity (self, familiar, and unfamiliar), exaggeration (normal speech, exaggerated speech, and speech-in-noise), and temporal variation, also referred to as speech type (speaking, singing, metronome, and static).

The sorting task itself was modified to suit a self-face experiment. As it is impossible to sort images of oneself into four different identity categories, there were only two identities in each sorting task: one was the self, familiar, or unfamiliar face used in the recognition task; the other was an unfamiliar foil, matched for gender and approximate age 23. The face-foil pairings remained consistent across participants (i.e., if Face A was familiar to one participant and unfamiliar to another, it was still paired with the same foil, Face B, which was unfamiliar to both participants). The rationale behind these pairings was similar to the sorting task in Chapter 6. If participants are more sensitive to characteristic patterns in their own movements (or a familiar person’s movements), they should be able to discriminate between their own (or a familiar face’s) movements and the movements of an unfamiliar face more easily than discriminating between the movements of two unfamiliar faces. In other words, participants should be able to recognise or match the familiar movement patterns, and hence sort the four clips of themself or a familiar face with relative ease (and by default, sort the clips of the unfamiliar foil). However, when faced with two unfamiliar faces, participants should have more trouble sorting the two faces based on their movement patterns. In total, participants completed 36 sorting tasks, one for each pairing.

Apart from the fact that participants only had to sort two faces, rather than four, the procedures for the sorting task were identical to those in Chapter 6. Participants were presented with eight boxes (four each of two different people), and asked to sort them into two groups according to identity. When they double-clicked on a box, the video (or still image) would appear for 2 s, and then disappear. To sort

---

23 Note that the small size and screen location variations were also taken into account when pairing the actors for the sorting task: if the familiar person leaned backwards or to the left in their clips, an unfamiliar foil was chosen that also leaned backwards or to the left.
the boxes, participants simply dragged the clips from one side of the screen into numbered columns on the other side of the screen. Participants were advised that they could watch the clips as many times as they wanted, but they must watch each clip at least once before sorting it into a group. Participants were also told that they could change their mind and re-sort the boxes into different groups as many times as necessary until they were happy that each group contained four clips of the same person. Each task had a 10 min time limit, and participants could monitor the time on their screen. No participants exceeded the time limit in any condition. When they were finished, participants clicked a button to indicate this and stop the timer. They could then move onto the next task in their own time.

Order of presentation of each task was pseudo-randomised across all variables except exaggeration, which was blocked and counterbalanced across participants. The only constraint to the randomisation was that the same identity or speech type did not appear three times in a row.

7.3.1.3.2 Recognition task. The recognition task also had three variables – familiarity, exaggeration, and speech type – with the same levels as the sorting task. There were four trials per condition, for a total of 144 trials per participant.

The videos used in the recognition task were identical to those in the matching task, except that only one unfamiliar face was presented. The procedure was almost identical to Experiment 8. In each trial, participants were shown a single PLDm, centred on the screen, and asked to indicate via key press whether the PLDm depicted their own face, their familiar partner, or an unfamiliar person. Unlike Experiment 8, Participants could not respond until the video was over.

Order of presentation of the videos was randomised for all variables except exaggeration, which was blocked. The order of exaggeration conditions was counterbalanced across participants. Participants did not receive any practice trials, but they were shown a sample stimulus (depicting someone not shown in the experiment) prior to the experiment.

7.3.2 Results of Experiment 9

7.3.2.1 Sorting task. As in Chapter 6, four dependent variables were analysed in the sorting task: accuracy, average number of views per clip, time to complete, and accuracy-per-view. Separate three-way within-subjects ANOVAs (3 x familiarity; 3 x exaggeration; normal, speech-in-noise, exaggerated; 4 x speech type;
metronome, singing, speaking, static) were carried out on each measure. Due to sphericity violations and the small sample size, multivariate tests (Wilks’ Lambda) are reported for all measures. Bonferroni corrections were applied to all pairwise comparisons.

7.3.2.1.1 Accuracy. Accuracy in the sorting task was calculated the same way as in Chapter 6. This resulted in a maximum score of 24 and a minimum score of eight. Chance performance for the 2-identity sorting task was calculated by randomly allocating the eight videos into two groups 100,000 times and calculating the average score. This gave chance performance in the Experiment 9 sorting task as 10.3. T-tests (two-tailed) revealed that the majority of conditions were above chance levels ($p < .05$), with the exception of four static conditions (self/normal, familiar/exaggerated, unfamiliar/exaggerated, unfamiliar/speech-in-noise; $p > .05$), one speaking condition (unfamiliar/exaggerated, $p = .159$) and one metronome condition (unfamiliar/exaggerated, $p = .091$).

Accuracy scores for the sorting task in Experiment 9 are shown in Figure 25. The ANOVA on accuracy showed significant main effects of exaggeration, $F(2,8) = 9.86, p = .007, \eta^2_p = .71$, and speech type, $F(3,7) = 8.68, p = .009, \eta^2_p = .79$. Pairwise comparisons on exaggeration showed that exaggerated speech clips were sorted significantly less accurately than either normal speech, $p = .013$, or speech-in-noise clips, $p = .008$. Normal speech and speech-in-noise clips were sorted equally well, $p = 1$. Pairwise comparisons on speech type showed that static clips were sorted significantly less accurately than any moving clips, all $ps < .05$, but there were no differences between speaking, singing, and metronome clips, all $ps = 1$. The main effect of familiarity was not significant, $F(2,8) = 2.35, p = .158, \eta^2_p = .37$, but planned comparisons showed that, on average, self and familiar faces were sorted marginally better than unfamiliar faces, $p = .065$.

There were significant interactions between familiarity and exaggeration, $F(4,6) = 6.30, p = .024, \eta^2_p = .80$, and familiarity and speech type, $F(6,4) = 11.92, p = .016, \eta^2_p = .95$. Self-faces were sorted less accurately from speech-in-noise than normal speech, $p = .036$; familiar faces were sorted less accurately from exaggerated speech than either normal speech, $p = .006$, or speech-in-noise, $p = .002$; unfamiliar faces were sorted less accurately from exaggerated than normal speech, $p = .008$.

The majority of pairwise comparisons in the familiarity and speech type interaction did not reach significance. There was no effect of speech type on familiar
Figure 25: Accuracy for the Experiment 9 sorting task. 

- a) Normal speech; 
- b) Speech-in-noise; 
- c) Exaggerated speech. 

White bars represent self-faces, light grey bars represent familiar faces, dark grey bars represent unfamiliar faces. Error bars represent +/- 1 standard error of the mean.
faces, and no overall advantage for familiar moving clips (metronome, singing, and speaking) compared to static clips, \( p = .132 \). Both self-faces and unfamiliar faces were sorted significantly more accurately from moving clips on average than static clips, self: \( p = .003 \); unfamiliar: \( p = .007 \), but the exact pattern of results differed slightly. Self-faces were sorted more accurately from metronome singing than static frames, \( p = .010 \), while unfamiliar faces were sorted more accurately from singing than from static frames, \( p = .027 \). No other interactions were significant, \( ps > .1 \).

7.3.2.1.2 Views and time. There were no significant effects of average number of views per clip or average time to complete each task, all \( ps > .1 \).

7.3.2.1.3 Accuracy per view. The accuracy-per-view results are shown in Figure 26. The overall pattern of results was broadly consistent with the accuracy analysis. The main effect of speech type was significant, \( F(3,7) = 14.27, p = .002, \eta^2_p = .86 \), and the main effects of exaggeration and familiarity were marginally significant, exaggeration: \( F(2,8) = 4.30, p = .054, \eta^2_p = .52 \); familiarity: \( F(2,8) = 3.97, p = .064, \eta^2_p = .50 \). For speech type, pairwise comparisons showed that static frames were sorted worse than either speaking, \( p = .021 \), or metronome singing, \( p = .002 \), but equally as well as singing, \( p = .183 \). When averaged together, there was an overall movement advantage: the average metronome, singing, and speaking conditions were significantly better than the static conditions, \( p = .004 \). For exaggeration, exaggerated speech was sorted significantly worse than speech-in-noise, \( p = .034 \), but not normal speech, \( p = .160 \). However, when averaged together, normal speech and speech-in-noise were sorted better than exaggerated speech, \( p = .022 \). As in the accuracy analysis, individual comparisons on familiarity failed to reach significance, \( ps > .05 \), but when averaged together, self and familiar faces were sorted better than unfamiliar faces, \( p = .024 \).

The interaction between familiarity and exaggeration was significant, \( F(4,6) = 6.74, p = .021, \eta^2_p = .82 \). Familiar faces were sorted worse from exaggerated clips than from normal speech, \( p = .036 \), or speech-in-noise, \( p = .006 \). Unfamiliar faces showed a similar trend: exaggerated clips were sorted worse than normal and speech-in-noise clips averaged together, \( p = .045 \), but no individual pairwise comparisons were significant, \( ps > .1 \). Self-faces were sorted equally well from normal, exaggerated and speech-in-noise clips, \( ps > .1 \). The interaction between speech type and exaggeration was also significant, \( F(6,4) = 18.96, p = .007, \eta^2_p = .97 \).
Figure 26: Accuracy-per-view for the Experiment 9 sorting task. *a)* Normal speech; *b)* Speech-in-noise; *c)* Exaggerated speech. White bars represent self-faces, light grey bars represent familiar faces, dark grey bars represent unfamiliar faces. Error bars represent +/- 1 standard error of the mean.
Static clips and singing were sorted equally well (or, in the case of static clips, equally poorly) from all levels of exaggeration, all \( ps > .1 \), but speaking clips were sorted worse from exaggerated speech than normal speech, \( p = .014 \), or speech-in-noise, \( p = .011 \), and metronome clips were sorted worse from exaggerated than speech-in-noise, \( p = .047 \). The familiarity and speech, or the familiarity, speech, and exaggeration interactions were not significant, \( ps > .7 \). However, individual analyses of each level of familiarity showed that self and unfamiliar faces showed a movement advantage: self: \( p = .014 \); unfamiliar: \( p = .016 \). There was no movement advantage for familiar faces, \( p = .372 \).

### 7.3.2.2 Recognition task

As in Experiment 8, recognition performance was analysed in three ways: accuracy, \( d' \), and chi-squared analyses to look at whether responding at each level of familiarity was random (see section 7.1.2.2 for details of how each measure was calculated). All three variables – familiarity, exaggeration, and speech type – were subjected to analysis in Experiment 9, unlike Experiment 8 where only movement type and familiarity were analysed. All ANOVA test statistics are Wilks’ Lambda.

#### 7.3.2.2.1 Accuracy

The accuracy scores for the recognition component of Experiment 9, collapsed over exaggeration, are shown in Figure 27. As in Experiment 8, participants performed poorly overall – responding was at chance level (33.3%) in the majority of conditions. Unfamiliar faces showing exaggerated movement were recognised better than chance in the metronome, singing, and static conditions, \( ps < .05 \), and unfamiliar faces showing normal speech were recognised better than chance in the static condition, \( p = .015 \).

The ANOVA on accuracy revealed no significant main effects or interactions, familiarity: \( F(2,8) = 1.77, p = .232, \eta^2 = .31 \); exaggeration: \( F(2,8) = 1.98, p = .201, \eta^2 = .33 \); speech type: \( F(3,7) = 1.19, p = .382, \eta^2 = .34 \); all interaction \( ps > .3 \). The results from Experiment 9 thus mirror Experiment 8 – participants were generally unable to recognise themselves or their friend from a PLDm.
Chapter 7: Personally Familiar and Self-Face Recognition

7.3.2.2.2 Signal detection theory analysis. Like the accuracy scores, the d’ scores for the recognition task were extremely low – participants performed at chance levels in all conditions, $p$s $>.1$, except familiar faces with exaggerated speech in the metronome condition, $p = .005$. The ANOVA revealed no significant main effects, familiarity: $F(1,9) = .111, p = .746, \eta_p^2 = .01$; exaggeration: $F(2,8) = 1.79, p = .228, \eta_p^2 = .31$; speech type: $F(3,7) = 1.30, p = .348, \eta_p^2 = .36$; all interactions $p$s $>.1$. No planned pairwise comparisons on movement type were significant, all $p$s $>.1$. Once again, there was no evidence that participants could recognise either their own face or their friend’s face when presented as a PLD.

7.3.2.2.3 Chi-squared analysis. As for Experiment 8, each participant’s response to each individual face (collapsed across exaggeration and speech type) was subjected to a chi-squared analysis. The individual response frequencies and chi-squared results are depicted in Figure 28. Fourteen out of the thirty chi-squared analyses were significant ($p < .05$), and another five were approaching significance ($p < .1$).

Figure 27: Proportion correct in the recognition task in Experiment 9, averaged across exaggeration condition. Chance performance was 0.33, indicated by the dashed line. White bars represent self-faces; light grey bars represent personally familiar faces; dark grey bars represent unfamiliar faces. Error bars represent +/- 1 standard error of the mean.
Figure 28: Response frequency in the recognition task of Experiment 9. Categories show the person in the PLD. Bars indicate participants’ responses: White “self”, light grey “familiar”, dark grey “unfamiliar”. Stars indicate the results of the chi-squared analyses: 1 star, $p < .1$; 2 stars, $p < .05$; 3 stars: $p < .0005$. 

265
Some of these results arose due to a response bias (e.g., an “unfamiliar” response bias for P2 and P7 when viewing familiar and unfamiliar faces). Overall, however, there was evidence that the majority (seven) of the ten participants was responding consistently, if incorrectly, when asked to identify at least one person from a PLDm.

The most obvious example of this is P9, who consistently mislabelled self-faces as familiar, familiar faces as unfamiliar, and unfamiliar faces as herself. Unlike Experiment 8, participants in Experiment 9 completed the recognition task first. They had no chance to learn the individual clips prior to the recognition task, which indicates that the consistent pattern of responding does not require pre-exposure to the stimuli.

7.3.3 Discussion of Experiment 9

Overall, the results from Experiment 9 confirm that participants can sort, but not recognise self and unfamiliar faces from characteristic movement patterns. Personally familiar faces were sorted well, but there was no evidence of a movement advantage. Unlike studies on emotion (e.g., Atkinson, 2004; Hill et al., 2005; Pollick et al., 2003), or studies on arm movements (Pollick et al., 2001) exaggeration does not help face recognition – in fact, personally familiar and unfamiliar faces were sorted worse from exaggerated clips than normal or speech-in-noise clips.

Eliminating relative or absolute timing cues through singing and metronome singing still preserves enough movement information to sort faces well above chance levels. However, as in Experiment 8, there was no evidence that participants could identify a person from their characteristic movement patterns, even when the movement was their own.

7.3.3.1 Self, personally familiar, and unfamiliar faces: sorting and recognition. In the sorting task, personally familiar and self-faces were sorted equally well, and better than unfamiliar faces on average. However, a movement advantage was only present for self and unfamiliar faces. The effects of familiarity and movement are therefore similar to those found in Experiments 1a, 2a, 3, and 4 (Chapters 4 and 5), in which famous faces were matched more accurately than unfamiliar faces overall, but only the unfamiliar faces showed a movement
advantage. As in Experiments 1-4, it is possible that any beneficial effects of movement for familiar faces were minimised due to high performance in the static condition – numerically, static personally familiar faces in this experiment were sorted better than static unfamiliar or self-faces. Once again, though, there is no obvious reason why participants should be better at sorting static familiar faces than self or unfamiliar faces. It is possible that participants were using idiosyncratic teeth or tongue information to match the familiar faces. However, unlike Experiments 1-7, familiar faces in Experiment 9 also served as self and unfamiliar faces for different participants – making it extremely unlikely that the effect of familiarity arose from idiosyncrasies in the stimuli. Furthermore, there is no evidence that the presence of teeth and tongues helped static famous face sorting in Experiment 7 (Chapter 6). As in Experiments 1-4 (cf., Chapter 6), the advantage for static familiar faces indicates that participants may be more sensitive to characteristic static cues in faces they are familiar with, even when the face is seriously degraded. It appears that this advantage does not arise because famous people are more stereotyped in their movements or static poses – the advantage can also appear for personally familiar faces. Furthermore, the high level of experience we have with moving personally familiar faces does not increase the movement advantage. Like famous faces, matching or sorting performance for personally familiar faces may reach an asymptote from static cues alone, and movement does not add significantly more information. However, the results from this experiment, along with Experiments 1a, 2a, and 4, and previous work by Thornton and Kourtzi (2002), Pilz et al. (2006), and Pilz et al. (2009) indicate that movement can significantly improve performance for unfamiliar faces compared to static images, via structure-from-motion, characteristic movement patterns and social signals.

Unlike personally familiar faces, the results from Experiment 9 suggest that characteristic movement patterns can be used when sorting one’s own face. However, while previous research using whole-body PLDs has found a clear advantage for self-recognition and matching compared to familiar or unfamiliar people (e.g., Loula et al., 2005; Prasad & Shiffrar, 2009), Experiment 9 did not replicate this pattern. Self-faces were sorted quite accurately and showed an overall movement advantage, but there was no distinct advantage for matching or recognising self-faces compared to personally familiar or unfamiliar faces. It is possible that the advantage for self-recognition is smaller for faces, compared to
whole-body stimuli: Cook et al. (2012) also found that personally familiar and self-faces were recognised equally well when they were presented upright, showing veridical movement. However, it is important to note that the results of the recognition task in Experiment 9 differ significantly from Cook et al., because Experiment 9 did not find any effect of familiarity when people were asked to name (rather than sort) the actors in the PLDm, and recognition performance was at chance levels for all levels of familiarity. The results from the recognition tasks in Experiments 8 and 9 are discussed further in the conclusions of Chapter 7.

### 7.3.3.2 Exaggeration in self, personally familiar and unfamiliar face sorting

The effect of exaggeration and speech-in-noise on sorting performance was surprising. Participants performed worst when sorting exaggerated movements, and equally well when sorting normal movements and speech-in-noise. However, when accuracy was adjusted for the number of views per clip, a slightly different pattern emerged. The detrimental effect of exaggerated clips was only apparent for familiar faces, and to a lesser extent, unfamiliar faces: self-faces were sorted equally well, regardless of whether the clips were normal, exaggerated, or speech-in-noise. There are several interpretations of this finding. First, it is possible that natural exaggeration does not enhance the idiosyncratic, individuating elements of characteristic movement patterns, but instead changes the movement patterns so much that they are unrecognisable. However, this does not explain why self-faces were not affected by exaggeration – self-face sorting was relatively accurate regardless of the level of exaggeration – or why familiar faces showed a stronger negative effect of exaggeration than unfamiliar faces.

Alternatively, it is possible that characteristic movement patterns (and characteristic static poses) based on visual experience with a face – that is, personal familiarity – are primarily veridical. If this were the case, exaggeration, or any change to the spatiotemporal characteristics of the movement, could mask familiar movement patterns (and change the characteristic static poses associated with the person). In other words, in the natural speech and speech-in-noise conditions, familiar face matching was helped by prior experience with the characteristic movement patterns (and static poses arising from them), but in the exaggerated speech condition, familiar faces were matched the same as unfamiliar faces – with no reference to prior experience. This is in line with Cook et al. (2012), who found that...
altering the spatial or temporal characteristics of personally familiar avatars reduced recognition levels to chance.

The fact that unfamiliar faces do not have any pre-existing characteristic movement patterns could explain why unfamiliar faces only showed a weak effect of exaggeration: the only visual experience participants had with the unfamiliar faces occurred during the experiment. This experience may have been sufficient to extract and match some characteristic movement patterns, but not enough to lead to a significant matching advantage for normal speech or speech-in-noise individually. It is possible that self-faces were not affected by exaggeration for the same reason unfamiliar faces were not – a lack of stored characteristic movement patterns – but this seems unlikely given our extensive experience producing characteristic patterns of movement. It is more likely that the self-face sorting results reflect the fact that self-face recognition relies more heavily on temporal characteristics than spatial characteristics (Cook et al., 2012). Exaggeration may have emphasised differences in spatial cues, but since participants were still speaking or singing the same words and melody, participants may have been relying on the (relatively consistent) timing cues to sort images of their own face from unfamiliar faces.

7.3.3.3 Timing cues in self, personally familiar and unfamiliar face sorting.
In general, all of the moving conditions – speaking, singing, and singing with a metronome – were sorted equally accurately. In other words, across all levels of familiarity and exaggeration, there was no clear effect of restricting the timing cues available in the stimuli. This was particularly surprising for self-faces, since sensitivity to one’s own movements is thought to be closely linked to sensitivity to one’s own temporal patterns (Cook et al., 2012; Flach et al., 2004). The fact that participants could still sort clips when the absolute and relative timing cues were restricted suggests two possibilities. Firstly, it is possible that participants could still extract and match some timing cues, probably at the phoneme level. The idea that characteristic timing cues may exist even in a single syllable or phoneme is supported by the fact that several cross-modal studies have found that participants can match an unfamiliar face to a voice based on a single word (Lachs & Pisoni, 2004a, b).

Secondly, participants may have been relying purely on spatial variation to sort the faces (i.e., variation in the direction or distance the face moved, as opposed to
variations in the speed or timing of the movement). It seems unlikely that movement-based recognition is based purely on spatial cues – previous studies have found that timing cues are particularly important for familiar and self-face recognition (e.g., Cook et al., 2012; Lander et al., 2006), and the results from the exaggerated condition of this experiment suggest that spatial variations have little impact on self-face sorting. However, actors in the recording phase may have compensated for the lack of timing variation in the singing and metronome conditions by increasing the amount of rigid head movement, which may have helped sorting independently of the level of exaggeration. Further research is necessary to examine whether isolating non-rigid movement cues would increase the effects of timing restrictions in perceptually-based experiments, and whether forcing people to curtail timing (or spatial) variations elicits more idiosyncratic movements in the spatial (or temporal) domain during production-based experiments.

7.4 Conclusions from Experiments 8 and 9

Experiments 8 and 9 found mixed results. Experiment 8 found no evidence for a movement advantage in self, personally familiar or unfamiliar face matching or recognition. Unlike Experiments 5, 6, and 7 (Chapter 6), there was no indication that the type of movement (rigid or non-rigid) had any effect on familiar or unfamiliar face matching. It is possible that the lack of effect in Experiment 8 arose due to the methodology or stimuli used – participants may have learned to recognise the static cues in the faces (decreasing the movement advantage), or they may have been able to match the faces based on structural information present in the non-shape-normalised PLDs.

Unlike Experiment 8, Experiment 9 found a strong effect of movement and familiarity that mirrored the results found for famous and unfamiliar faces in Experiments 1a, 2a, and the avatar condition in Experiment 4 (Chapters 4 and 5): unfamiliar, but not familiar faces showed a movement advantage, even when the stimuli were shape-normalised. This indicates that movement plays a similar role – that is, it has a relatively small impact – in movement-based matching of famous and personally familiar faces. It is surprising that the results from the sorting task in Experiment 9 do not match the results of the sorting tasks in Experiments 6 and 7, and the combined analysis in Chapter 6 – that is, a movement advantage for familiar
faces. One possible explanation is that movement is more important for familiar faces when people are attempting to discriminate between a number of different people (as in the sorting task in Experiments 5, 6, and 7, Chapter 6), as opposed to discriminating between two people (as in the matching task used in Experiments 1-4, Chapters 4 and 5; and the sorting task used in Experiment 9).

Notably, Experiment 9 is the third experiment in this thesis to find an overall movement advantage for shape-normalised unfamiliar faces, which confirms that it is possible to extract and match characteristic movement patterns for unfamiliar faces, even when they are highly degraded. The fact that unfamiliar faces showed a movement advantage in Experiment 9, in the avatar condition of Experiment 4, and also in the combined movement condition of Experiment 5 (Chapter 6), supports the idea that characteristic movement patterns can be matched quite well without prior familiarity. This confirms that, although it may add significant information on top of characteristic movement patterns, participants do not rely solely on structure-from-motion information to match unfamiliar faces. Furthermore, the movement advantage for shape-normalised unfamiliar faces in Experiment 4 was not simply a consequence of the same/different matching task.

Likewise, the fact that shape-normalised self-face sorting showed a movement advantage in Experiment 9 suggests that characteristic movement information can carry identity cues for self-recognition, independent of structure-from-motion information. Interestingly, exaggeration had a minimal effect on self-sorting, unlike familiar or unfamiliar faces. This suggests that our stored characteristic movement patterns for self faces (with which we have motor experience, but limited visual experience) may differ significantly from those for familiar faces (for which we only have visual experience), possibly because self-face sorting relies more on timing cues than on spatial information (see also Cook et al., 2012).

One of the most interesting results from Experiment 8 and 9 is the fact that participants could not identify themselves, their friend and an unfamiliar person in a three-alternative forced choice task, even when structural cues were available and there was unlimited time to make a decision (Experiment 8), or when the stimuli provided detailed articulatory information (Experiment 9). This indicates that our ability to match or sort faces based on movement dissociates from our ability to identify a person based on their movement. It is unclear why participants could not name themselves or their friends from the PLDs or shape-normalised PLDm.
Previous studies have shown ample evidence that participants can name familiar faces from PLDs in a six- or seven-alternative forced choice task (Bruce & Valentine, 1988; Rosenblum et al., 2007). It is possible that the results from Experiment 9 differ from Bruce and Valentine (1988) and Rosenblum et al. (2007) because their stimuli were not shape-normalised, but that does not explain why Experiment 8, which included structural cues, also failed to find any evidence of movement-based recognition. The pattern of results is even more puzzling when the results of Experiment 9 are compared to Cook et al. (2012). Experiment 9 and Cook et al. (2012) used the same three-alternative forced choice procedure, and both used shape-normalised stimuli, similar clip lengths (Cook et al.’s clips were 3.7 s long, compared to 2 s in this experiment), showing relatively natural speech-based movements. It is possible that asking people to speak the words of a song (as opposed to telling a joke) minimised the presence of characteristic movement patterns. However, Experiment 8 also found no evidence that participants could recognise their own facial movements when they were shown longer clips depicting themselves and their friends telling jokes, speaking casually to the experimenter, or reading dialogue from a radio play. This suggests that the type of speech and length of the clip were not factors in participants’ inability to perform the recognition task.

Despite the extremely poor performance in the recognition task, the results from Experiments 8 and 9 provided some evidence that participants were not guessing when asked to recognise themselves and their friends. The majority of participants showed highly consistent response patterns for at least one of the three faces in the recognition task, which implies that they were extracting some characteristic movement patterns in a very few trials and generalising them across the rest, but they were not able to link them to stored movement patterns. The fact that participants showed similar response patterns in Experiments 8 and 9 indicates that the consistent patterns of response do not require prior familiarisation with the clips, and nor do they require shape information to be present in the stimuli.

Overall, the results from Experiments 8 and 9 indicate that participants are capable of extracting and matching characteristic movement patterns from unfamiliar and self-faces. There was no indication that matching was affected by the presence of rigid or non-rigid movement, but the presence of structural cues in the Experiment 8 stimuli may have encouraged participants to rely on structure-from-motion or static form-based cues rather than specific movement patterns. The way someone moves
their face varies depending on the requirements of the task (e.g., casual conversation, compared to singing, or speaking in a noisy environment), but results from mismatched speech-style trials (Experiment 8) and different types of speech (Experiment 9) indicate that people are remarkably good at generalising across these variations and extracting invariant identity cues (characteristic movement patterns), particularly for self and unfamiliar faces. In the recognition task, participants showed some evidence that they were extracting individuating information from the stimuli and applying this consistently across the recognition task, which confirms that movement-based matching can occur over a longer time-frame than a simple sequential matching task. However, the results from Experiments 8 and 9 offer no evidence that viewing these characteristic movement patterns allows us to access and match our stored representations of movement patterns; either visually for familiar face recognition, or via a perception-action coupling for self-recognition.
Chapter 8

Movement and Familiarity in Face Recognition

Conclusions
CHAPTER 8: MOVEMENT AND FAMILIARITY IN FACE RECOGNITION:

CONCLUSIONS

This thesis presented nine experiments investigating the role of movement and familiarity in face recognition. The main question addressed throughout these experiments was whether familiar and unfamiliar faces show a quantitatively different movement advantage, and whether the movement advantage for familiar and unfamiliar faces arises from different mechanisms – that is, do familiar or unfamiliar faces benefit more from structure-from-motion and characteristic movement patterns, or do both mechanisms contribute equally to the movement advantage in all faces. As well as these theoretical contributions, the thesis makes methodological contributions to our understanding of when movement advantages occur. Experiments 1 and 2 (Chapter 4) investigated whether participants could match a video or still image to a facial point-light-display (PLD), and whether the duration of the PLD or the presence of movement in the sample clip contributed to the movement advantage for famous and unfamiliar faces. Experiments 3 and 4 (Chapter 5) compared matching performance for PLDs and shape-normalised avatars, once again using famous and unfamiliar faces. Experiments 5, 6, and 7 (Chapter 6) specifically examined characteristic movement patterns, and whether the type of movement (rigid, non-rigid, or both) and location of movement (eyes or mouths) had an impact on the movement advantage for famous and unfamiliar faces. Finally, Experiments 8 and 9 (Chapter 7) investigated the movement advantage for personally familiar and self-faces, focusing on whether natural variations in speech type, exaggeration or timing cues had a similar effect on matching and recognition performance for unfamiliar, familiar and self-faces. The results from Experiments 1-9 are summarised in Table 17. Each chapter included a general discussion that examined the theoretical implications of the empirical findings, and those discussions will not be repeated in detail. This chapter examines the overall pattern of results across all nine experiments, and discusses their implications for models of face recognition, their impact on real-world face recognition, and future directions for research on movement and face recognition.
8.1 Familiarity and the Movement Advantage: Structure-from-Motion and Characteristic Movement Patterns

The majority of experiments found that, in at least some conditions, participants were more accurate when asked to match or sort moving than static clips. In general, the results of Experiments 1-9 indicate that movement does carry individuating cues, and people are able to use those cues to match, but not necessarily name, familiar and unfamiliar faces. Overall, there was more evidence for a movement advantage for unfamiliar faces than for familiar faces. As illustrated in Table 17, there was a movement advantage for unfamiliar faces in Experiments 1a, 2a, 4, 5, 7, and 9; for familiar faces in Experiments 4, 6, 7; and for self-faces in Experiment 9. This section examines why the movement advantage arose in those experiments – in other words, which movement-based mechanisms gave rise to the movement advantage for familiar and unfamiliar faces. Roark et al. (2003) summarised previous research on movement and face recognition, and proposed three non-exclusive movement-based mechanisms to explain the movement advantage. The representational enhancement hypothesis suggests that movement facilitates recognition by providing enhanced information about the three-dimensional structure of the face and head. In other words, head movements give rise to structure-from-motion, which builds a better three-dimensional representation of the face and head. The supplemental information hypothesis suggests that people move in idiosyncratic manner, and that these characteristic movement patterns can act as an alternate route to identification. Finally, the social signals hypothesis suggests that movement carries social information, which may attract and maintain attention to a face, thereby helping subsequent recognition; or may distract people from the identity of the face, thereby hindering subsequent recognition. The contribution of structure-from-motion and characteristic movement patterns will be addressed separately in the following sections. Social cues, which were not tested directly, will be discussed where relevant to the results.
Table 17: 
*Summary of results from Experiments 1-9*

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Familiarity effect</th>
<th>Movement advantage</th>
<th>Other effects</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chapter 4</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E 1a</td>
<td>Famous &gt; Unfamiliar</td>
<td>Unfamiliar faces (trend only)</td>
<td>Moving &gt; Static sample clip</td>
</tr>
<tr>
<td>E 1b</td>
<td>No familiarity effect</td>
<td>No movement advantage</td>
<td>None</td>
</tr>
<tr>
<td>E 2a</td>
<td>Famous &gt; Unfamiliar</td>
<td>Unfamiliar faces</td>
<td>Moving &gt; Static sample clip</td>
</tr>
<tr>
<td>E 2b</td>
<td>No familiarity effect</td>
<td>No movement advantage</td>
<td>None</td>
</tr>
<tr>
<td>Combined analysis</td>
<td>Famous &gt; Unfamiliar (overlapping clips only)</td>
<td>Unfamiliar faces (overlapping clips only)</td>
<td>Moving &gt; Static sample clip</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MTS = same/different task</td>
</tr>
<tr>
<td><strong>Chapter 5</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E 3</td>
<td>Famous &gt; Unfamiliar</td>
<td>No movement advantage</td>
<td>PLD &gt; avatars (accuracy only)</td>
</tr>
<tr>
<td>E 4</td>
<td>Famous &gt; Unfamiliar</td>
<td>Unfamiliar &gt; Famous</td>
<td>PLD &gt; avatars (movement advantage)</td>
</tr>
<tr>
<td><strong>Chapter 6</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E 5</td>
<td>No familiarity effect</td>
<td>Unfamiliar faces, combined movement</td>
<td>Famous: non-rigid disadvantage</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unfamiliar: combined movement advantage</td>
</tr>
<tr>
<td>E 6</td>
<td>Unfamiliar &gt; Famous</td>
<td>Famous faces (accuracy-per-view only)</td>
<td>Famous: non-rigid disadvantage</td>
</tr>
<tr>
<td>-----------</td>
<td>---------------------</td>
<td>--------------------------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>E 7</td>
<td>Famous &gt; Unfamiliar</td>
<td>Famous faces overall, unfamiliar faces (combined movement, accuracy trend only)</td>
<td>Famous: no clear pattern (movement type)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unfamiliar: no clear pattern (movement type)</td>
<td></td>
</tr>
<tr>
<td>Combined analysis</td>
<td>Famous &gt; Unfamiliar</td>
<td>Famous faces overall; fam/unfam combined movement</td>
<td>Famous: non-rigid disadvantage</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Famous non-rigid movement</td>
<td>Unfamiliar: combined movement advantage, slight rigid disadvantage</td>
</tr>
</tbody>
</table>

**Chapter 7**

<table>
<thead>
<tr>
<th>E 8</th>
<th>No familiarity effect</th>
<th>No movement advantage</th>
<th>No effect of movement type</th>
</tr>
</thead>
<tbody>
<tr>
<td>E 9</td>
<td>Self &amp; personally familiar &gt; Unfamiliar</td>
<td>Self &amp; unfamiliar faces</td>
<td>Exaggeration impairs familiar, but not self-face sorting. No effect of timing restrictions</td>
</tr>
</tbody>
</table>
8.1.1 Structure-from-motion

As reviewed in Chapter 3, there is a large amount of evidence that structure-from-motion contributes to face recognition: rigid head movement improves face learning and matching, compared to static images (Farivar et al., 2009; Pike et al., 1997; Schiff et al., 1986; cf, Christie & Bruce, 1998; Lander & Bruce, 2003) which supports the idea that structure-from-motion can help build a better representation of a previously unknown face and head. However, there has been no research comparing the effect of structural information on familiar and unfamiliar faces. Theoretically, structure-from-motion cues should help all face recognition, regardless of familiarity. Experiments 3 and 4 tested this hypothesis directly, by comparing matching performance for famous and unfamiliar faces using shape-normalised avatars (i.e., no structural cues) and PLDs containing structural information. Experiment 4 (Chapter 5) found that PLDs (non shape-normalised) gave rise to a significantly larger movement advantage than shape-normalised avatars, which indicates that structural cues provide important individuating information beyond characteristic movement patterns. Notably, there was no interaction between familiarity and the type of stimulus, which implies that the role of structure-from-motion is similar, regardless of whether the face is famous or not.

Experiments 3 and 4 (Chapter 5) made two significant contributions to the literature on structure-from-motion and face recognition. Several studies have examined the role of structure-from-motion in unfamiliar face recognition (Farivar et al., 2009; Pike et al., 1997; Schiff et al., 1986), but Experiments 3 and 4 are the first to study the impact of structure-from-motion on familiar face recognition. Consequently, Experiment 4 is also the first research to demonstrate that familiar faces derive a significant movement advantage from structure-from-motion, separate from the benefit derived from characteristic movement patterns. In general, previous research that has examined structure-from-motion (either directly or indirectly) has used staged rigid head movements (e.g., rotational movements, Farivar et al., 2009; Pike et al., 1997; Schiff et al., 1986; or nodding and shaking the head, Christie & Bruce, 1998; Lander & Bruce, 2003; Lander & Davies, 2007). As such, Experiment 4 is the first study to show that unscripted, natural movements can give rise to a movement advantage based on structure-from-motion.
The results from Experiment 4 support the representational enhancement hypothesis, and confirm that structure-from-motion can contribute to famous and unfamiliar face matching. However, it is difficult to quantify the benefit that structure-from-motion provides, because the stimuli in Experiment 4 also differed in two other ways – the shape-normalised stimuli had a face superimposed on them, and included eye movement information. Both of these factors could have distracted participants from the identity-matching task or otherwise lowered performance in the avatar condition, which may have increased the difference between stimuli with and without shape cues. The presence of eyes seems an unlikely explanation given that Chapter 6 showed no analogous effects when eyes were added to the PLD stimuli.

Interestingly, given the fact that structure-from-motion had a significant impact in Experiment 4, there was little evidence that participants were using structural cues to match faces in Experiments 1, 2, or 3. If participants were using structure-from-motion cues to match faces, Experiments 1 and 2 should have found no effect of clip length, and there should have been evidence of a movement advantage from non-overlapping movement sequences (Experiments 1b and 2b; see section 8.2 for further discussion). Theoretically, participants relying on structure-from-motion should also show a movement advantage when one image is static and one image is moving rigidly (e.g., Farivar et al., 2009; Pike et al., 1997), as participants should be able to extract structure-from-motion cues from the moving stimulus and match the structural cues to the static stimulus. This hypothesis was tested in Experiments 1-4 (Chapters 4 and 5), which included all permutations of moving and static images (i.e., moving/moving; moving/static; static/moving; static/static). In this thesis, only one experiment – Experiment 4 – found a significant movement advantage for PLDs when one was moving and one was static (the moving/static and static/moving conditions for famous and unfamiliar PLDs). One possible explanation for these findings is that the clips used in Experiments 1-3 did not contain enough rigid head movement to support identification from structure-from-motion. However, Experiments 3 and 4 used identical movement sequences, and Experiment 4 showed a clear movement advantage when participants were generalising from static to moving PLDs (and vice versa). Furthermore, this explanation does not account for the fact that PLDs provided a clear advantage over avatars in Experiment 4 and a moderate advantage over avatars in Experiment 3, assuming that at least some of the benefits of PLD in these experiments arose from structure-from-motion.
Another possibility is that participants had trouble matching a static and a moving image. The results from Experiments 1, 2, and 3 resemble those of previous studies such as Bruce et al. (2001, Experiment 3), Bruce and Valentine (1988), Christie and Bruce (1998, Experiment 2A), Lander and Davies (2007), and Schiff et al. (1986), who failed to find a movement advantage for experimentally familiar faces learnt in motion, but tested with static images. Chapter 3 raised the possibility that some studies of experimentally familiar faces failed to find a movement advantage because participants were asked to recall faces after protracted periods (e.g., 30 mins, Christie & Bruce; or 24 hours, Schiff et al., 1986). However, the fact that participants did not show a movement advantage for familiar or unfamiliar faces when matching from a static image to a PLD (Experiments 1-3), or from simultaneously presented CCTV images and photographs (Bruce et al., 2001, Experiment 1), suggests that the results of the previous studies might have occurred because participants were unable to extract and compare structural cues, rather than problems with memory retrieval. In other words, it may be difficult, although not impossible (e.g., Experiment 4, Farivar et al., 2009; Pike et al., 1997), to encode structural information from a moving image and subsequently compare it to a static image (or vice versa), regardless of whether the comparison is immediate or delayed. This idea will be discussed further in section 8.2.

The presence of superficial image features and social cues may have also contributed to the lack of movement advantage for static sample images in Experiments 1-3. Participants might be able to extract and use structure-from-motion cues in some circumstances – for example, when the head is simply rotating (Farivar et al., 2009; Pike et al., 1997; Schiff et al., 1986), or when both images are degraded (Experiment 4) – but the presence of the social cues in the non-degraded faces in Experiments 1-3 may have prevented or inhibited it. In other words, it is possible that participants in Experiments 1-3 were trying to interpret the social signals present in the static faces – for example, trying to identify the facial expression, or whether the person was familiar – or focusing on superficial cues (e.g., hairstyle), rather than concentrating on the structure of the face. The idea that social cues can impair structural encoding is in line with results from Schiff et al. (1986), who found that eyewitness identification was better when people were familiarised with neutral images of a person, compared to freeze-frames extracted from a mock crime-scene. The presence of social cues in non-degraded images may also explain why Christie
and Bruce (1998) and Lander and Bruce (2003) failed to find a movement advantage for experimentally familiar faces familiarised with rigid movements (nodding and shaking movements, which may be seen as communicative).

Overall, then, Experiments 1-4 established that structure-from-motion can contribute to famous and unfamiliar face matching, but it is difficult to determine how important structure-from-motion is when matching from a moving image to a static image, or vice versa. Future research may take advantage of the wide range of 3D cameras and monitors now available, to test whether providing three-dimensional cues in static images (i.e., providing more structural information for participants) improves movement-based face recognition or matching.

### 8.1.2 Characteristic Movement Patterns

Unlike structure-from-motion, the Roark et al. (2003) model suggests that familiarity should have a significant effect on the use of characteristic movement patterns. According to the O’Toole et al. (2002) and Roark et al. (2003), prior exposure to a face (i.e., some level of familiarity) is needed to determine what patterns of movement are characteristic or idiosyncratic to a person, and to use this information as a cue to identity. As reviewed in Chapter 3, very few studies have directly compared the movement advantage for familiar and unfamiliar faces (see Table 4, Chapter 3). Of those that have, none have attempted to isolate characteristic movement patterns from structure-from-motion. Consequently, Experiments 4-7 and 9 are the first studies to test whether characteristic movement patterns give rise to a movement advantage for both familiar and unfamiliar faces, and, if so, whether familiar faces derive more benefit than unfamiliar faces.

This thesis provided clear evidence that people can match or sort famous and unfamiliar faces based on characteristic or idiosyncratic patterns of movement. Participants showed a movement advantage when stimuli were shape-normalised (Experiments 4, 5, 6, 7, and 9), which suggests that some of the benefits of movement arose because participants could match characteristic movement patterns. Since the clips in Experiments 1-7 were matched for amount and distinctiveness of movement, the ability to match or sort shape-normalised stimuli was likely based on common or idiosyncratic sequences of movement (e.g., shaking of the head or raising the eyebrows) rather than amount of movement or how distinctive each movement pattern was.
The movement advantage for famous faces in Experiments 6, 7, and the combined analysis of Chapter 6 is noteworthy because Experiments 4-7 are the first studies to examine the movement advantage for shape-normalised famous faces. Most previous studies of famous faces have been unable to distinguish whether the movement advantage arose due to characteristic movement patterns, structure-from-motion, or both (e.g., Knight & Johnston, 1997; Lander et al., 2001; Lander & Chuang, 2005). Some studies examined the contribution of characteristic movement patterns by disrupting, reversing, speeding or slowing the movement of the famous faces (Lander & Bruce, 2000, 2004; Lander et al., 1999), but no studies have asked participants to match or identify famous faces based purely on their movement patterns. The movement advantage for famous faces confirms that participants can use characteristic movement patterns as a cue to identity, even when participants do not know the faces are famous and when all form-based and structural cues have been made uniform.

As mentioned above, the Roark et al. (2003) model only posits a limited role for characteristic movement in unfamiliar face recognition. This hypothesis is supported by the fact that people generally perform at chance levels when asked to name or match unfamiliar whole-body PLDs within the context of self, familiar and unfamiliar recognition tasks (Prasad & Shiffrar, 2009; Loula et al., 2005). However, Hill and Johnston (2001) demonstrated that participants can match or sort unfamiliar faces using shaped-normalised stimuli (i.e., based on characteristic movement patterns only). Results from this thesis, specifically Experiments 4, 5, 7 and 9, confirm Hill and Johnston’s findings that characteristic movement patterns can be used to match and sort unfamiliar faces. Participants can sort and match shape-normalised faces well above chances levels, even without prior exposure to non-degraded images of the faces.

Although both famous and unfamiliar faces show a movement advantage from shape-normalised stimuli, both Experiment 6 and the combined analysis of Chapter 6 found a significant interaction between familiarity and whether the clips were moving or static: famous faces gave rise to a significantly larger movement advantage than unfamiliar faces. These results are the first empirical evidence that characteristic movement patterns can give rise to a larger movement advantage for familiar than unfamiliar faces.
Based on the results from Experiments 5-7, it is possible to draw some preliminary conclusions about the type of characteristic movement information participants are able to extract and compare for famous and unfamiliar faces. There were some commonalities between famous and unfamiliar faces: evidence from Experiments 5, 6, and 7 indicates that characteristic movement patterns for famous and unfamiliar faces probably consist of a combination of rigid and non-rigid movements, and isolating either type of movement is detrimental to matching performance. These findings contradict those of Hill and Johnston (2001), who found that rigid head movement patterns were important for unfamiliar face sorting. This discrepancy may have arisen because Hill and Johnston only measured accuracy, whereas Experiments 5-7 also measured the number of times participants viewed each clip.

There were also some differences between familiar and unfamiliar faces. Although there was an overall advantage for combined rigid and non-rigid movement, participants could use isolated non-rigid movement as a cue to identity for famous faces (in line with Lander & Chuang, 2005; Bruce & Valentine, 1988; Lander et al., 2006; Rosenblum et al., 2007), but not unfamiliar faces. This finding is analogous to the inner-face advantage for familiar faces (Ellis et al., 1979), and may arise because we spend a large amount of time attending to the non-rigidly moving areas of the face (e.g., the eyes and mouth) when interacting with or watching familiar faces.

In line with this, the importance of different facial regions also varied across famous and unfamiliar faces. For famous faces, adding either eye or mouth information resulted in a significant movement advantage, which suggests that people can extract and compare movement patterns from multiple areas of a familiar face. Interestingly, the addition of detailed mouth information (teeth and tongues) helped participants sort famous faces more accurately than did eye information.

The results of Experiments 5, 6, and 7 indicate that we can also extract movement cues from multiple areas of an unfamiliar face. Unlike famous faces, though, there was no overall benefit of adding teeth and tongues. Furthermore, adding detailed eye movement information (which improves overall sorting accuracy) may impair the ability to extract and compare characteristic movement patterns from unfamiliar faces, as evidenced by the fact that Experiment 6 was the only sorting task that did not show a movement advantage for unfamiliar faces in any
condition. Once again, this may be due to social cues: eyes convey a myriad of signals about a person’s attention and state of mind (Emery, 2000), which may distract people from attending to characteristic patterns of movement.

Overall, these results offer some support for the supplemental information hypothesis proposed by Roark et al. (2003). When characteristic movement patterns are isolated using shape-normalised stimuli, famous faces show a slightly larger movement advantage than unfamiliar faces (although this result is not apparent in all studies). However, these results imply that characteristic movement patterns can be useful even for unfamiliar faces – this suggests that characteristic or idiosyncratic movements can be extracted and/or learnt after relatively little exposure to a face. In other words, familiarity is helpful, but not necessary, to match characteristic patterns of facial movement.

8.1.3 Familiarity and the Movement Advantage: Other Findings

8.1.3.1 Self Faces. The majority of this thesis examined familiarity using famous faces. However, This thesis also conducted two studies using self-face stimuli (Chapter 9). Due to the difficulty of recruiting participants, and apparent individual differences in movement-based matching performance, only limited conclusions can be drawn about the role of movement in self-face recognition. Nonetheless, Chapter 7 provided some evidence that movement plays an important role in self-face matching.

The role of movement in self-recognition has, to this point, primarily been examined using whole-body PLDs (e.g., Jokisch et al., 2006; Loula et al., 2005; Prasad & Shiffrar, 2009; Sevdalis & Keller, 2010). Experiments 8 and 9 are the first studies to use facial PLDs in self-face matching and recognition tasks. The results from Experiment 9 support the conclusion that a perception-action coupling enhances sensitivity to our own movements (Knoblich & Flach, 2003), although probably less so for face than for whole-body movements: Cook et al. (2012) and Experiment 9 both found equal performance for self and familiar faces displaying upright, veridical movement, whereas Prasad and Shiffrar (2009), and Loula et al. (2005) found a significant advantage for self compared to familiar body recognition. The fact that natural movement exaggeration does not impair self-face sorting, but it does impair familiar and unfamiliar face sorting, implies that either motor-based expertise builds stronger, more flexible representations than visual expertise; or that
primarily spatial manipulations (exaggeration) have little effect because self-
movement recognition is based on timing-based cues (cf. Cook at al., 2012), albeit
not the absolute rhythm or speed of speech (Experiment 9). Future research may
want to extend on these findings by testing the effect of artificial (i.e., computer-
generated) exaggerations on self-recognition using both face and whole-body stimuli.

8.1.3.2 An overall advantage for familiar face matching. Roark et al.’s
(2003) model suggests that familiar faces should benefit from movement in two
separate ways: via structure-from-motion, and via characteristic movement patterns.
However, apart from Experiments 4, 6, and 7, this thesis found relatively little
evidence for a familiar face movement advantage, particularly when participants
were asked to discriminate between two faces in a same/different matching task or 2-
face sorting task: Experiments 1, 2, 3, 5, 8, and 9 all failed to find a movement
advantages for familiar faces. This is surprising, given that previous research has
consistently found a movement advantage for famous faces using naming tasks (e.g.,
Knight & Johnston, 1997; Lander & Bruce, 2000, 2004; Lander et al., 2001; Lander
et al., 2001; Lnader & Chuang, 2005).

While there was no overall movement advantage for familiar faces in the
majority of experiments, Experiments 1a, 2a, 3, 4, 7 and 9 all found some advantage
for familiar face matching compared to unfamiliar face matching overall. This
indicates that the lack of a movement advantage did not occur because of poor
matching performance for familiar faces. Rather, participants were relatively
accurate when asked to match moving and static images of familiar faces, but
movement did not improve matching performance significantly compared to the
static trials. This is in line with previous studies, such as Burton, Wilson et al. (1999)
and Bruce et al. (2001), who found an overall matching advantage for personally
familiar compared to unfamiliar faces, even in the absence of a movement advantage.

This leads to the question raised in several chapters: why was static matching
of familiar faces so good? Some of the possible reasons for the familiar/famous over
unfamiliar face matching advantages were discussed in Chapters 4 and 5. These
include poorly matched static stimuli, increased attention to trials that showed
famous faces, and pre-existing representations of familiar faces. Overall, it is
unlikely that familiar faces were matched better than unfamiliar faces because the
static images of familiar faces were easier to match. While the image dissimilarity
analysis in Experiment 4 found that mismatched famous avatars were more dissimilar than mismatched unfamiliar avatars, image dissimilarity cannot explain why participants had significantly more hits in the famous static face condition than unfamiliar static face condition. Furthermore, a marginal overall advantage for personally familiar faces over unfamiliar faces was also found in Experiment 9, where the personally familiar stimuli also acted as self and unfamiliar stimuli (meaning that the difference could not be something about the stimuli per se). Stimulus discrepancies also fail to account for why previous research has found a similar advantage for familiar face matching (e.g., Bruce et al., 2001; Burton, Wilson et al., 1999). The attention-based explanation also fails to account for the famous face matching advantage: participants were better on average at matching famous than unfamiliar avatars in Experiment 4, despite the fact that both sample and test stimuli were degraded and they were not told that some of the faces were famous prior to the experiment.

As such, it is probable that participants were good at matching static familiar faces because they have stored pre-existing representations of them. Traditional research on static face recognition suggests that familiar face representations (often referred to as face recognition units, or FRUs) primarily encode invariant features, such as the structure or shape of the face and its features (Bruce & Young, 1986). However, the overall familiar face matching advantage arose in shape-normalised stimuli (Experiments 4 and 9) as well as standard PLDs (Experiments 1a, 2a, 3, and 4), which indicates that static shape- or feature-based cues were not the sole basis for the advantage. It is possible that familiarity with a person’s face entails more than simply familiarity with the structure of the face and head. As Burton et al. (2011) suggested, our mental representations of a familiar person might also capture typical elements of our encounters with them, such as size, lighting, and other image-based factors. The results of the current studies suggest that familiar face representations may also incorporate typical person-based factors such as common poses (for example, facial expressions or head tilt/turn), and this could give rise to a significant familiarity advantage when participants are asked to match degraded images of those faces that incorporate characteristic poses or expressions. The results from Experiments 5-7 and the combined analysis (Chapter 6) imply that rigid head poses are particularly useful for static famous face matching. Rigid clips were sorted consistently well in the famous face condition, regardless of whether they were
moving or static. In fact, rigid clips were sorted as well as combined movement, and
better than non-rigid movement, in Experiments 5, 6, and combined analysis. On the
other hand, for the unfamiliar faces there was little evidence of an advantage for rigid
over non-rigid clips, moving or static (with the exception of Experiment 7, in the
accuracy-per-view analysis).

The idea that we may be able to recognise people from characteristic poses
may explain why famous face recognition remains relatively accurate despite severe
image degradations such as blurring and stretching (Hole et al., 2002). Furthermore,
the artistic world frequently exploits the idea of characteristic poses: actors or
comedians imitating famous people often adopt distinctive poses that are designed to
capture and accentuate their subject’s typical head and body positions.

Evidence from Experiments 2 and 4 indicate that familiarity with a face might
help people match faces in a very specific way: familiarity may help reject
mismatched faces, independent of response bias, as opposed to simply making it
easier for people to identify when two images depict the same person. It is important
to note, though, that this pattern was not apparent in all experiments (e.g.,
Experiment 3).

In conclusion, the experiments presented in this thesis confirm that both
structure-from-motion and characteristic movement patterns play a role in familiar
and unfamiliar face processing. Characteristic movement patterns can also contribute
to self-face matching. Interestingly, the role of movement in face recognition was not
consistent across all experiments, probably because the presence of a movement
advantage for familiar faces was influenced by the presence of characteristic static
cues. The following sections examine the experiments that failed to find a movement
advantage (Experiments 1b, 2b, 3, and 8), and assess the contribution of various
methodological factors to the overall pattern of results.

8.2 Exceptions to the Movement Advantage

While the majority of experiments in this thesis found a movement advantage,
there were several notable exceptions – Experiments 1b, 2b, 3, and 8. The lack of
movement advantage in Experiments 1b, 2b, and 3 was most likely a result of
extremely poor performance overall – participants were barely above chance in the
majority of conditions in Experiments 1b and 2b, and performed notably worse in
Experiment 3 when compared to Experiment 4. The lack of movement advantage in Experiment 8 was the opposite of the problems in Experiments 1b, 2b, and 3 – participants performed remarkably well in the matching task, in both static and moving conditions, which minimised any chance of finding a movement advantage. The results of Experiment 8 probably arose because participants learnt the static cues throughout the experiment, and were able to match the faces simultaneously (these findings are discussed in detail in Chapter 7). This general discussion will focus on why participants performed poorly overall and showed no movement advantage in Experiments 1b, 2b, and 3.

Out of the nine experiments in this thesis, only three experiments – 1b, 2b, and 3 – asked participants to generalise across image formats and movement sequences. They were also the only experiments that failed to find a movement advantage due to overall poor performance. Based on these results, there are two main explanations for the failure to find a movement advantage in these experiments: first, participants could not generalise from one movement sequence to another; and second, participants could not generalise from one image format to another.

It is unlikely that Experiments 1b, 2b and 3 failed to find a movement advantage because participants could not generalise across movement sequences: several experiments, in this thesis and in previous research, have shown that participants can match two or more images of moving faces when the movement sequences do not overlap (e.g., Experiments 4, 5, 6, 7, 9; Hill & Johnston, 2001; Lander & Davies, 2007; Roark et al., 2006), so it seems unlikely that this caused the lack of movement advantage. This leaves the possibility that participants were unable to generalise across formats. The idea that changing the format of an image can impair face recognition is not completely new: Liu and Chaudhuri (1997) found that old/new recognition performance for experimentally familiar faces learnt and tested as photographic negatives was equal to those learned and tested as non-degraded images, but changing the format (negative to positive or vice versa) between learning and test significantly impaired performance. Similarly, Liu, Collin, Rainville, and Chaudhuri (2000) found that participants were better at recognising experimentally familiar faces when the learning and test images contained overlapping spatial frequencies, even when the images were severely blurred. Liu and Chaudhuri (2000) suggest that these findings can be explained by the encoding specificity principle – if the image characteristics present in the initial, encoded image are incongruent with
the test image (or the sample and test images in a matching task, for the experiments in the current thesis), it is more difficult to match than if the two images contain the same characteristics, regardless how degraded those characteristics are. The encoding specificity principle might explain why participants were unable to generalise from a non-degraded sample image to a PLD in Experiments 1a and 2a, and why there was no movement advantage in Experiment 3.

Although the encoding specificity principle may explain the poor performance in Experiments 1b, 2b, and 3, it is difficult to explain why all face matching is not hindered by image format changes. Liu, Seetzen, Burton & Chaudhuri (2003) found a slight advantage when participants were asked to recognise or match from CCTV footage to a high quality photograph (incongruent formats), compared to recognising or matching from CCTV footage to similarly degraded static images (congruent formats). Furthermore, in Experiments 1a and 2a of this thesis, in which the sample and test images contained overlapping movement sequences, participants were able to successfully (although not perfectly) generalise from non-degraded images to PLDs. Why are participants able to generalise in some studies and not others? The answer may lie in task difficulty. Participants in Liu et al.’s (2003) study were still relatively accurate in the incongruent condition – roughly 64% correct – compared to the incongruent conditions in Liu et al.’s (2000) study – roughly 50% correct. Comparisons of Experiments 1 and 2 show a similar pattern: in conditions with overlapping movement sequences, participants had a mean $d'$ of 1.28, whereas conditions without overlap had a mean $d'$ of 0.51. This implies that the encoding specificity principle is primarily important in difficult tests of recognition – such as when generalising from a negated image to a non-degraded image (Liu & Chaudhuri, 1997), from high-pass to low-pass filtered images (Liu et al., 2003), or from one movement sequence to another (Experiments 1b, 2b, and 3). In other words, generalising across image formats and movement sequences is difficult; it leads to poor performance and eradicates any movement advantage (Experiments 1b, 2b, and 3). On the other hand, generalising across formats is possible in an easy task (overlapping movement sequences, Experiments 1a and 2a), and participants can generalise across movement sequences when they are not distracted by changes to the image format (Experiments 4-7). Consequently, future research should take the encoding specificity principle into account when examining movement-based face recognition, and investigate under what circumstances changes to image formats
impair face recognition or matching performance. This may build a more comprehensive understanding of the role of movement in face recognition.

The encoding specificity principle might also apply when one image is moving and one image is static – participants may have trouble generalising due to the change in format, from static to moving, particularly when other format changes occur simultaneously (e.g., Experiments 1ab, 2b, 3; Lander & Davies, 2007). As discussed in Chapter 3, many studies on experimentally familiar faces have only tested the role of movement in the learning phase, and have presented static images in the test phase (see Table 2, Chapter 3). This has two problems – first, as discussed in Chapter 3, these studies only examined the role of structure-from-motion and social signals, not characteristic movement patterns; and second, changing the format between learning and test may have minimised any movement advantage (this may explain the mixed results for experimentally familiar faces). Results from Christie and Bruce’s (1998) experiment support this hypothesis – when participants were tested using static images, there was a significant movement disadvantage – that is, static images in the familiarisation phase resulted in better recognition of static test images, compared to moving images in the familiarisation phase. Once again, tests using three-dimensional static images may help us determine whether poor performance generalising from static to moving images occurs because participants cannot gather enough structural information from the non-degraded static image (in which case, 3D static images would improve matching), or whether a change in image format, from static to moving, disrupts performance (in which case there would be no, or minimal effect of showing the static images in 3D).

Overall, matching performance in general seems to be disrupted by differences in format when the task is relatively difficult or demanding, and movement does not facilitate generalisation across formats. Across the nine experiments in this thesis, several other methodological factors contributed significantly to the movement advantage (or lack thereof). These factors are discussed in section 8.3.

8.3 Methodology and the Movement Advantage

As well as the theoretical question addressed in this thesis – the effect of familiarity on the movement advantage in face recognition – the results discussed in sections 8.1 and 8.2 highlight the fact that methodological factors such as the task
and stimuli used in each experiment can also have a significant effect on the movement advantage. These methodological factors may also help to explain why the previous literature on movement and face recognition has found inconsistent results (e.g., compare Pike et al., 1997; Lander & Davies, 2007 and Bruce & Valentine, 1988; Christie & Bruce, 1998). To investigate the role of methodological factors, this thesis systematically compared the movement advantage across multiple tasks, stimulus types, and stimulus durations. The rationale and results for different tasks and stimuli have been discussed at length in individual chapters (Chapters 3, 4, 5, and 6), and this section offers a summary of these findings and a brief overview of some important methodological factors that should be taken into account when designing future experiments.

8.3.1 Does the Task Matter?

All of the experiments in this thesis used some variety of matching task: match-to-sample, same/different, or sorting (two or four faces). Matching tasks were chosen because they offer the opportunity to directly compare performance for familiar and unfamiliar faces. In general, matching tasks that asked participants to discriminate between two people – that is, Experiments 1-4, 8, and 9 – showed a similar pattern of performance (with the exception of the experiments discussed in section 8.2). Familiar faces (famous or personally familiar) were discriminated better than unfamiliar faces overall, but showed a small or non-existent movement advantage; unfamiliar faces were discriminated poorly from static images, but perhaps partly because of this, showed a significant movement advantage. Direct comparisons between the match-to-sample and same/different matching tasks (Experiments 1 and 2) revealed similar levels of performance, which indicates that matching performance was not significantly influenced by the fact that trials in the match-to-sample task were longer and contained more movement information (i.e., an extra PLD) compared to the same/different task. The fact that the two-face sorting task (Experiment 9) showed similar results suggests that the option to view each video or static image multiple times, additional time to make a response, and a larger number of movement sequences also has little impact on the results.

Introducing other faces into the task changed the pattern of results, as illustrated by the four-face sorting task used in Experiments 5, 6, and 7. Most notably, the overall matching advantage for familiar faces disappeared (with the
exception of a marginal accuracy advantage for famous faces in Experiment 7), but a movement advantage for famous faces arose in Experiments 6, 7, and the combined analysis of Chapter 6. The movement advantage for unfamiliar faces was substantially reduced compared to the two-face matching tasks.

It is unlikely that task difficulty fully explains the difference between the two- and four-face matching tasks: increasing task difficulty by asking participants to generalise across image formats eradicated the movement advantage in Experiments 1b, 2b, and 3. Instead, these results indicate that the movement advantage for familiar faces is more prominent when participants are required to discriminate between a number of faces, as opposed to a pair of faces. This may be because participants were relying on characteristic static poses to discriminate faces in the two-face tasks (see section 8.1.3.2), but there simply were not enough discriminating static cues to sort four faces. Consequently, participants shifted their focus to the characteristic movement patterns, giving rise to a movement advantage. This may explain why previous studies using famous faces have found a consistent movement advantage (e.g., Knight & Johnston, 1997; Lander & Bruce, 2000; Lander et al., 2001; Lander & Chuang, 2005). Since participants were “exposed” to a much larger “pool” of faces when asked to identify the famous person (i.e., they were asked to discriminate between all the celebrities they could recall, or from a list of multiple celebrities, rather than just matching pairs of faces), participants may have relied more on movement cues rather than static pose information.

Unlike familiar faces, the movement advantage for unfamiliar faces is most prominent when participants are asked to discriminate between two faces. This may be because static matching across two unfamiliar faces is difficult (Hancock et al., 2000), and participants concentrate on movement-based cues to supplement their poor performance. However, since participants have no experience with them, participants may have difficulty discriminating between multiple characteristic movement patterns. This leads to a reduced movement advantage compared to familiar faces, and may result participants focusing on unusual or idiosyncratic cues (such as the static eye cues in Experiment 6) to complete the task.

Several experiments in this thesis also used a recognition test to determine whether people could name themselves, their friend or an unfamiliar person based on movement. Although participants were able to match PLDs quite effectively, they could not accurately name people from the same stimuli. There was evidence that
participants were relatively consistent in their (mis)identification, indicating that movement patterns can be stored and matched over a relatively long time period (around 15 mins per experiment), but no indication that performance in matching experiments is related to performance in a naming task. Puzzlingly, participants in previous studies have been able to identify themselves their friends from whole-body and facial PLDs (Bruce & Valentine, 1988; Cutting & Kozlowski, 1977; Loula et al., 2005; Rosenblum et al., 2007) and shape-normalised avatars (Cook et al., 2012) – exactly what participants in Experiments 8 and 9 failed to do. It is unclear why these discrepancies arose: perhaps the PLDs in Experiments 8 and 9 presented movement information that was too sparse to support overt identification. As mentioned in Chapters 5 and 6, very little research has investigated whether the type of stimulus has an effect on movement-based face recognition, and future studies may wish to extend on the findings of Chapters 5 and 6 by examining whether the type of stimulus and the location of movement information (e.g., eyes, mouth, or a combination of the areas) affects naming as well as matching performance.

Future experiments should formally test the relationship between matching and naming using moving stimuli, as this issue has implications for eyewitness testimony – it may be more reliable to ask witnesses to match a moving image of a suspect to CCTV footage, rather than overtly name a person based on degraded videos. It would be particularly interesting to compare naming and/or matching performance for moving and static faces with facial features occluded, to determine whether the presence of characteristic movement information is particularly helpful when people are asked to identify someone who they encountered wearing a disguise (e.g., a mask over the eyes or mouth).

8.3.2 Do the Stimuli Matter?

Research on face recognition has found a movement advantage using a variety of stimuli: non-degraded faces (e.g., Pike et al., 1997; Pilz et al., 2006; 2009; Thornton & Kourtzi, 2002; Roark et al., 2006); negated, thresholded, blurred and pixelated faces (Knight & Johnston, 1997; Lander & Bruce, 2000; Lander et al., 2001; Lander & Chuang, 2005; Lander & Davies, 2007); PLDs (Bruce & Valentine, 1988; Rosenblum et al., 2002; Rosenblum et al., 2007) and shape-normalised avatars (Cook et al., 2012; Hill & Johnston, 2001; Knappmeyer et al., 2003; Watson et al., 2005). This indicates that the movement advantage is not confined to any one
method of presentation, but, as discussed in Chapter 3 and by several other researchers (e.g., Knight & Johnston, 1997; O'Toole et al., 2002), it is possible that the movement advantage is more prominent when static cues are harder to access (i.e. in degraded images).

This thesis examined the movement advantage using a number of image degradations: PLDs, shape-normalised avatars, shape-normalised PLDs, and shape-normalised PLDs with eyes or mouths. The results from Experiments 1-7 imply that the movement advantage for unfamiliar faces may be more prominent when structural cues are included in the stimuli. However, presenting PLDs with structural information may inflate static matching performance for familiar faces, and thus minimise the movement advantage (although this may also be associated with the use of non-neutral static frames, which will be discussed below). On the other hand, Experiments 4-7 and 9 found a movement advantage for both familiar and unfamiliar faces (although frequently not at the same time), which indicates that shape-normalised stimuli also carry helpful movement cues. The fact that both avatars and PLDs (normal and shape-normalised) can support a movement advantage in some circumstances suggests that the type of stimulus is less important than the task requirements discussed above. However, evidence from Knappmeyer et al. (2003) indicated that the form cues present in a full-face avatar (e.g., the face texture, shape and features) interact with movement cues when making identity judgements – even small changes in form (i.e., preserving or normalising the texture) had a significant effect on the movement advantage. Participants in Experiments 3 and 4 performed relatively poorly when asked to match full-face avatars, possibly due to the same effects found by Knappmeyer et al. (2003). Presenting a slightly more realistic facial form (compared to a PLD, for example) may encourage participants to focus on those superficial form cues – after all, they are the cues we use most often in everyday recognition. As such, it may be difficult to simply disregard them completely, particularly for familiar faces (even when we do not know the faces are familiar, as in Experiment 4). In other words, the visual similarity between the facial forms (i.e., shape and features) may overshadow the differences in their movement, and compromise the movement advantage.

Unsurprisingly, Experiments 5-7 showed that adding extra movement information to the stimuli (eyes and mouths) improves matching performance. Future research may wish to investigate whether eye or mouth movement alone can support
recognition or matching (using methods adapted from Cunningham et al., 2005; Nusseck et al., 2008), or whether adding extra cues increases the movement advantage for familiar and unfamiliar faces in an additive fashion (e.g., both eyes and mouths, shoulders to accentuate rigid movements, etc.). This research may be supported by new methods of isolating, recreating and manipulating facial movement, such as the face tracking technology used in this thesis or the “markerless” avatar creation technology used by Cook et al. (2012).

While the type of stimulus (PLD vs. avatar) may have some influence on the movement advantage, the content of the stimuli is also important. Static faces naturally vary in their distinctiveness – that is, how different they are compared to our facial norms – and this facilitates recognition performance (Valentine & Bruce, 1986). Lander and Chuang (2005) established that distinctiveness also contributes to the movement advantage: they found that “more distinctive” movements gave rise to a movement advantage, whereas “less distinctive” movements did not. In Experiments 1-7, the famous and unfamiliar face stimuli were carefully selected based on ratings of the amount and distinctiveness of movement in the clips – the first time these factors have been specifically controlled in a study of the movement advantage in famous faces. The fact that amount and distinctiveness of movement were controlled confirms that the findings from Experiments 4-7 and the combined analysis of Chapter 6 were based on participants discriminating characteristic patterns of movement, rather than using rules of thumb such as “face A moves a lot more than face B” or “face A moves in a very normal way, face B’s movements are unusual”.

It is unlikely that controlling distinctiveness and amount of movement can fully account for the lack of movement advantage for famous faces in Experiments 1-3, or for unfamiliar faces in Experiments 1b, 2b, and 3. However, if previous studies of the movement advantage in famous faces (e.g., Knight & Johnston, 1997; Lander & Bruce, 2000; Lander et al., 1999; Lander et al., 2001) included some particularly distinctive movements, it is possible that they inflated the movement advantage, which may go some way to explaining the differences between the previous studies and the results presented in Experiments 1-7.

The content of the static control frames is also important. As discussed in Chapters 4 and 5, studies that used static control frames may have minimised the presence of characteristic pose cues in their static stimuli, but not their moving clips.
This in turn may have confounded the effect of movement with the effect of static pose cues. Some empirical evidence for this assertion was provided in the control task in Experiment 6: when participants were shown neutral static frames, rather than frames extracted randomly from the centre of the clips, their performance in the static condition dropped significantly. Once again, this does not suggest that the use of neutral compared to random static frames is the sole reason that previous studies found a movement advantage for famous faces and Experiments 1-3 did not. The fact that participants still show a movement advantage when presented with multiple static frames (Lander et al., 1999) suggests that movement conveys an advantage beyond the addition of extra static cues. However, small stimulus-based factors such as the selection of controlled or matched moving clips static images and non-neutral static frames may have a significant impact on the presence of a movement advantage, particularly for famous faces. As such, these factors should be considered carefully in future research on the role of movement in face recognition.

8.4 Conclusions: Qualitative and Quantitative Differences in the Movement Advantage for Familiar and Unfamiliar Faces

Evidence from previous research and the nine experiments presented in this thesis indicates that the effect of movement on familiar and unfamiliar face recognition is not straightforward: it is affected by small variations in task and stimuli, which can eliminate the movement advantage for familiar or unfamiliar faces, or reverse the overall pattern of results (e.g., Experiments 1-4 and 5-7). Furthermore, an examination of performance across all experiments indicates that there are significant individual differences in people’s ability to use movement to match or identify a face. Despite these qualifications, this thesis found evidence for a quantitative difference in the movement-based matching of familiar and unfamiliar faces. The direction of the quantitative difference (i.e. do familiar or unfamiliar faces derive a larger benefit of movement?) depends primarily on the task and stimuli, which suggests that participants were using static and dynamic cues in a flexible manner.

As suggested in Chapter 5, the use of static cues and movement information may be strategic, and proceed in a hierarchical manner. This strategy may be akin to heuristics such as “take-the-best” or “fast-and-frugal trees” (Gigerenzer &
Gaissmaier, 2011), where people search through the available cues in order of their validity (or in a pre-determined order), then stop their search and make a decision (in this case, same/different identity) based on the first cue that discriminates between the alternatives. In real life, static information (e.g., shape, pose, texture) is probably the most commonly used cue to identity, and participants may carry this preference through even when the stimuli are moving, by attempting to match structural and pose cues. This may be an effective strategy when discriminating between two highly familiar faces, and could explain why static matching for familiar faces was relatively good in Experiments 1a, 2a, 3, 4, and 9. Reliance on static cues is likely to be less effective when attempting to discriminate between four familiar faces or two unfamiliar faces (and perhaps two highly degraded famous faces, as in Experiment 4 with PLDs). In this case, participants may resort to using movement cues, which could explain the movement advantages found in Chapters 4, 5, 6, and 7. In other words, movement may play a “catch-up” role when static recognition is impaired, as suggested by O’Toole et al. (2002) and Knight and Johnston (1997).

Participants may also use different movement cues preferentially: for example, when identical movement sequences are present, participants may concentrate on matching them, rather than extracting structure-from-motion or overarching characteristic patterns of movement. If identical movement sequences are not present, participants may attempt to match the clips based on structure-from-motion and characteristic movement patterns, or both, depending what is available.

The basic assumption behind this flexible use of cues is that participants have limited attentional resources (Chun, Golomb, & Turk-Browne, 2011), and employ strategies that facilitate fast, frugal (less cues) decision-making rather than the systematic evaluation and comparison of all available information (Gigerenzer & Gaissmaier, 2011). In practice, this may mean that when participants allocate attention to static cues, they do not have the capacity to also compare movement sequences. Likewise, if participants are attending to matching identical movement sequences, they do not have the capacity to also match faces based on structure-from-motion cues. This hierarchical or flexible cue use, constrained by limited attention, might explain why there was no movement advantage in Experiments 1b, 2b and 3. If participants were focusing on generalising from a non-degraded image to a PLD or avatar, they may have sufficient resources left to match identical movement sequences (Experiments 1a, 2a), but not to encode and match characteristic poses,
structure-from-motion or characteristic movement information (Experiments 1b, 2b, 3). When participants were not required to generalise across formats, but still had a two-face matching task, participants could attend to both static and movement-based cues, explaining the matching and movement advantage for famous faces in Experiment 4.

A flexible or strategic use of static cues and movement information can also account for many previous results. An overall preference for static cues could explain why many studies using non-degraded faces have failed to find a movement advantage (e.g., Bruce et al., 2001; Bruce & Valentine, 1988; Christie & Bruce, 1998; Lander & Bruce, 2003; Burton, Wilson et al., 1999). Likewise, the flexible use of either or both movement mechanisms may explain why some research has found that structure-from-motion information is useful for face recognition (e.g., Farivar et al., 2009; Pike et al., 1997; Schiff et al., 1986), whereas other studies support the idea that characteristic motion patterns support movement-based face recognition (e.g., Hill & Johnston, 2001; Kamachi et al., 2003; Lander & Bruce, 2000, 2004; Lander et al., 1999; Lander et al., 2006).

In sum, the flexible use of static and dynamic cues can account for the quantitative differences in the movement advantage for familiar and unfamiliar faces found across Experiments 1-9. However, there was little evidence that familiar and unfamiliar faces used movement in qualitatively different ways. There was some evidence that familiarity with a face increases performance when matching non-rigid motion (Chapter 6), and that eye and mouth information is more important for familiar than unfamiliar face matching – this supports the idea of an “inner-face” advantage for familiar faces (Ellis et al., 1979). However, further research is necessary to determine whether this represents a true qualitative difference between familiar and unfamiliar face processing, or whether the difference is merely quantitative.

Most importantly, though, there was no evidence of a qualitative difference in the use of structure-from-motion or characteristic movement patterns. The current results suggest that both structure-from-motion and characteristic movement patterns make unique contributions to movement-based face recognition, for both familiar and unfamiliar faces. This is surprising, given that the model of movement and face recognition by O’Toole et al. (2002) and Roark et al. (2003) implied that unfamiliar faces would only derive a minimal benefit from characteristic movement patterns.
compared to familiar faces. However, their model was focused on face recognition, and the majority of the studies in this thesis examined face matching. As mentioned above, future research may examine the relationship between recognition and matching, to determine whether the conclusions from the current research also apply to movement-based recognition.

In conclusion, this thesis found evidence that familiar and unfamiliar faces both use similar movement cues to supplement performance in identity-matching tasks, but the relative size of the movement advantage varied significantly depending on the task requirements and stimuli. One outstanding question in this field is how static recognition and movement-based recognition are related. Future research in this field may address this question by examining individual differences – in other words, is the movement advantage more pronounced for people who perform well in static face recognition tasks, or does movement play a “catch-up” role for people who are poor at recognising static images of faces? Such research, along with the studies presented here, would build on current models of face recognition and further our understanding of face processing and recognition in real-life scenarios.
REFERENCES


Bower, G. H., & Karlin, M. B. (1974). Depth of processing of pictures of faces and


APPENDIX A: STIMULI

Examples of stimuli used in Experiments 1-9. See attached CD-ROM.
APPENDICES ON CD-ROM CAN BE VIEWED AT UWS LIBRARY.