March 2003. *Behavioral and Cognitive Neuroscience Reviews*, **2(1)**, 15–46

Memory for moving faces: Psychological and Neural Perspectives on the Role of Motion in Face Recognition

Dana A. Roark
The University of Texas of Dallas

Susan E. Barrett
Lehigh University

Melanie Spence, Hervé Abdi, & Alice J. O'Toole
The University of Texas of Dallas

In the real world, faces are in constant motion. Recently, researchers have begun to consider how facial motion affects memory for faces. We offer a theoretical framework that synthesizes psychological findings on memory for moving faces. Three hypotheses about the possible roles of facial motion in memory are evaluated. In general, although facial motion is helpful for recognizing familiar/famous faces, its benefits are less certain with unfamiliar faces. Importantly, the implicit social signals provided by a moving face (e.g., gaze changes, expression, and facial speech) may mediate the effects of facial motion on recognition. Insights from the developmental literature, which highlight the significance of attention in the processing of social information from faces are also discussed. Finally, a neural systems framework that considers both the processing of socially relevant motion information and static feature-based information is presented. This neural systems model provides a useful framework for understanding the divergent psychological findings.

Keywords: faces, face recognition, facial expressions.

Why are faces so intriguing? The answer has much to do with the variety and importance of the information they can convey. Most notably, a face provides a unique indication of the identity of a person. Indeed, the remarkable capacity humans have to recognize and discriminate the myriad of faces they see in a lifetime has been a subject of interest to psychologists for many years. We can remember hundreds if not thousands of individual faces, with memories that endure decades of separation (Bahrick, Bahrick, & Wittlinger, 1975). The ability to remember large numbers of individual exemplars within a category of objects distinguishes faces from other familiar visual object categories (see Farah, Wilson, Drain, & Tanaka, 1998).

The information faces provide about identity, however, is just the beginning. From the time we are born, faces are the preferred objects of our attention (Nelson, 2001). With a quick glance at a face, we can easily guess the age, race, sex, and ethnicity of a person (Bruce & Young, 1986). The face also provides us with a moment-to-moment window on the emotional state of a person. All of this information is effortlessly and instantaneously accessible to us from a single static image of a person. Yet, in the real world, we see faces *in motion*. The complexity and diversity of the information available from static features of faces has sometimes diverted the attention of psychologists away from the fact that faces tilt, nod, look away, laugh, grimace, and speak. Each of these motions serves to enrich the social content of everyday human interactions (Allison, Puce, & McCarthy, 2000). These facial motions guide our attention, engage our emotions, and prompt our actions. The goal of the present review is to understand how facial motions affect our ability to learn and remember faces.

The vast majority of psychological data on the perception and recognition of faces has relied on static images of faces as stimuli. Only in the last five years or so have researchers embarked on studies employing moving faces. The shift from using static to moving faces as stimuli is in part due to advances in computing power that enable easy presentation and manipulation of digital video in the context of an experiment. It has thus been possible to begin to address questions aimed at understanding face recognition and person identification in more naturalistic contexts. In addition, the presence of security cameras and the development of automatic face recognition algorithms have likewise motivated recent attempts to understand the effects of motion on memory for faces. How well do the results of experiments on static face recognition generalize to more naturalistic contexts? What additional performance factors must we consider when faces are in motion during a memory task? How well can we remember moving faces? Does motion facilitateor perhaps hinderthe process of extracting the important invariant, and inherently static features of the face? What role do the social communication signals embedded in facial motions play in our memory for a person? Remarkably little is known about these important questions.

The primary purpose of this paper is to provide an exhaustive review of the psychological literature on memory for moving faces. Although the number of papers on this topic is limited, there are now enough studies to combine the available findings and to propose a more theoretical approach to the problem of memory for moving faces. In an earlier paper we introduced, in a highly abridged format, some of the core theoretical ideas that we discuss in this review (OToole, Roark, & Abdi, 2002). Our objective here is to provide a more comprehensive and critical review than was possible in a short synopsis. We hope this review will be helpful for psychologists interested in face perception and memory, and also for computational researchers interested in developing algorithms for face recognition, face tracking, and other applications that process moving faces. We also present a neural framework for understanding the functions of various face responsive brain regions in the context of the face memory literature. As might be expected from a new area of inquiry, the goals and methods of experimentation, as well as the results, have varied widely. Our attempt here is to integrate divergent findings, and identify the stimulus and processing factors that have emerged as relevant variables for recognizing moving faces.

This paper is organized into four parts. In the first part of the paper, we present a taxonomy of facial motions and their functions. We argue that most kinds of facial movements serve a social function and therefore, may impact how a person processes identity information from a face in motion. Next, we briefly sketch out three hypotheses about the role of facial motions in memory. Although the ideas behind the hypotheses have been mentioned frequently in the studies we re-

---

Send correspondence to: Dana A. Roark, or Hervé Abdi, or Alice O'Toole, School of Human Development, GR4.1 The University of Texas at Dallas Richardson TX 75083-0688 danar@utdallas.edu, herve@utdallas.edu, otoole@utdallas.edu. Note: For readers interested in viewing video examples of facial movements, please visit: `http://www.utdallas.edu/∼/otoole/index.htm`. Authors Notes This work was supported in part by a DOD/DARPA grant awarded to A.O. and H.A. Thanks are due to Josh Harms for help with figure preparation, to Dawn Hurst for help with references, and to Sarah Snow, Christina Blum, Yi Cheng, and Steve Lomber for helpful comments on the text and insightful discussions. Special thanks go to Jim Bartlett and to an anonymous reviewer for their perceptive comments and useful suggestions on an earlier draft of this paper.

view, the hypotheses have not been tested systematically. They nonetheless provide a useful framework for guiding readers through the exhaustive review of memory studies with moving faces that follows. We have decided to provide more detail about the methods and stimuli employed in these studies than one might expect in a review of this sort. We do this because the experimental design and stimulus factors that impact performance are still emerging in this new area of research. It is, thus, handy to have the details of the various studies easily available in a single review paper. For readers less interested in the details of the studies, we provide a comprehensive summary of the basic psychological findings following the review. This should enable a more casually interested reader to skip through the exhaustive review to the next section.

The second part of the paper is devoted to integrating the psychological data into the hypotheses. In doing so, we attempt to delineate the nature of evidence required to support each hypothesis in the context of an empirical study.

In the third part of the paper, we draw on findings from the developmental literaturelargely overlooked in previous discussions of moving facesto assess what is known about how infants extract information from moving faces. A developmental perspective, with its emphasis on attentional and social processes, may offer new insights into how face recognition operates in natural contexts. Although an exact mapping between the adult and developmental literatures is not possible, researchers studying adult face recognition should find this section interesting and informative. In part four, we will introduce a neural systems framework for processing faces. This framework, provided by Haxby, Hoffman, and Gobbini (2000), is relevant here because it considers both the processing of social information from moving faces and the processing of identity information from static facial features. In this final section, we map the relevant psychological findings onto the proposed brain areas from this model.

Facial Motions and Hypotheses

A taxonomy of facial motions

Facial motions can be divided technically, albeit somewhat artificially, into rigid and non-rigid movements. We define and characterize the different motions and their functions for reference in the review of the psychological studies.

Rigid rotations and translations of the head provide continually changing views of the face. These movements are called “rigid because the face itself does not deform or change shape as it moves (i.e., like the face of a statue). Common rigid rotations include head turning, nodding (ventral and dorsal flexion of the head), and shaking (left and right movements of the face). Such movements provide the observer both with a moving stimulus and with more perspective views of the head than would be encountered by a static viewer and subject. In other words, as a face turns, we see it from a number of different angles throughout the rotational sequence (e.g., frontal to profile, see Figure1). Because rigid head motions provide the viewer with additional perspective views of the face, care has been taken in psychological experiments to distinguish between recognition effects due to the number of views seen and those due to the motion of the head per se (e.g., Pike, Kemp, Towell, & Phillips, 1997; Lander, Christie, & Bruce, 1999). We note here that non-rigid movements can also provide, albeit to a much lesser degree, additional views of the face. However, non-rigid motions do not afford the viewer with the kind of sweeping rotation and flexion movements typical of rigid facial motions (see section below).

Rigid translations of the head also occur when a person approaches or passes by us. The relative size and viewing angle of a face can change quickly in these situations. In the real world, therefore, translations are generally combined with size and perspective changes and are caused by the relative motions of the viewer and the person viewed. Rigid motions play an important role in the context of social interaction. For instance, to begin a conversation we turn our head to look at the person we wish to address. Looking away from someone

Figure 1. Rigid head motions provide more views of the face than can be observed from a static head. The 9 still images shown here are extracted from a video sequence, in which the head moves 180-degrees, from left to right profile.

can indicate boredom or the intent to end a social interaction. A head turn is also an effective way to redirect the attention of others. Also, we can evaluate the continually changing attention of others to objects and people in a scene by noting changes in direction of their heads.

Rigid motions of the head can serve non-verbal linguistic functions as well. A nod of the head up and down can mean, "I agree!" Shaking the head, left to right, can mean, "I disagree!" And, more complex combinations of head rotations can indicate disbelief or aversion. Putting our head down can indicate embarrassment, sadness, or shyness just as a head tilted to the side can convey curiosity or inquisitiveness.

Non-rigid head movements are temporary "deformations" of the shape of a face or the relative locations of facial features. Perhaps more so than rigid head motions, non-rigid movements vary considerably in form and are especially effective for adding nuances of meaning to our conversations and communications with others. These can be grouped into three main categories: speech production movements, facial expression movements, and changes in the direction of eye gaze.

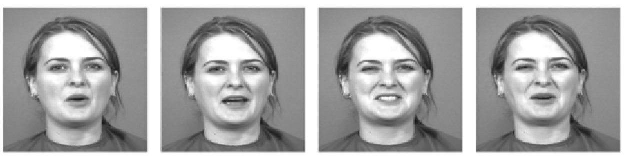

Figure 2. Facial speech is one example of non-rigid facial motion. Viewing lip and mouth movements can facilitate speech perception.

Figure 3. Facial speech is one example of non-rigid facial motion. Viewing lip and mouth movements can facilitate speech perception.

The movements of the mouth and face during the production of speech are arguably the most commonly observed types of facial movement. These movements provide information useful for understanding speech (Bernstein, Demorest, & Tucker, 2000), as illustrated compellingly by the now classic "McGurk Effect." Specfically, McGurk and MacDonald (1976) found evidence for multimodal integration of visual and auditory information for phoneme recognition by placing visual and auditory information about phonemes in competition. They found that subjects hearing a person pronouncing "ba," while watching a face saying "ga," often report perceiving the sound "da"a fused version of "ba" and "ga." The effects of visually observing facial speech have been extended to show that viewing lip movements during speech can improve the intelligibility of a verbal message under a variety of conditions (Calvert, Brammer, & Iversen, 1998; Campbell, Dodd, & Burnham, 1998). See Figure 2 for an example.

A second type of non-rigid facial movement occurs in the form of facial expressions. The perception of facial expressions has been investigated intensely in the cognitive, social, clinical, and neuropsychology literatures (see Adolphs, 2002, for a comprehensive review). Expressive face movements involve a combination of face parts (i.e., eyes, mouth, nose, cheeks) via complex patterns of time-coordinated muscle contractions. Facial expressions convey information about a persons emotional state. This information is useful for guiding the tone and tenor of social interactions. For example, when we are speaking to someone, who raises an eyebrow or gives a puzzled look, we interpret this as an invitation to further explain ourselves. See Figure 3 for examples.

Figure 4. Eye-gaze changes are a type of non-rigid facial motion. They are important for indicating the focus of a persons attention and for re-directing the attention of others.

Eye gaze changes, a third category of non-rigid facial motion, are frequent and socially informative. The muscles that control eye movements are some of the busiest in the body. Humans can move their eyes every 150 ms, enabling the active and efficient exploration of our visual environments. Changes in the direction of eye gaze are important for guiding our social interactions. Specifically, direction of gaze is an important cue to the object of a persons attention and shifts in eye gaze help orient us to changes in the environment. Moreover, eye gaze direction, compared to the direction in which another persons head is pointing (rigid motion), is a better indicator of a persons focus of attention (Perrett, Harries, Mistlin, Hietanen, Benson, & Bevan, 1990). Even momentary changes in gaze are quite effective for redirecting the attention of others (see Langton, Watt & Bruce, 2000, for a review). See Figure 4 for an example.

Although the labels of "rigid" and "non-rigid" facial motion work well for classification purposes, in most cases, different kinds of motions occur simultaneously in moving faces. To illustrate this point, the reader is encouraged to try talking to someone, or laughing without moving her head, her eyes, etc.! In laboratory settings, however, it is possible, and common, to isolate or control the two types of motion in an experiment, by creating stimuli that include a single type of motion. Yet more often than not, real-world human interactions include both rigid and non-rigid motions. Work in our lab indicates that examples of pure rigid or pure non-rigid motions are difficult to capture during a videotaped conversation. Usually, head movements accompany both facial speech and expression sequences in a very natural way.

Further, it is important to note that the non-rigid speech, expression, and gaze-change movements also occur in combination, in meaningful temporal sequences. Think about someone asking, "What did you have for lunch yesterday?" In responding to this question, your eyes may gaze upward initially while you try to recall yesterdays lunch. Your mouth may form the words, "Let me think," and you will probably smile in response to such an off-the-wall question. Clearly, in naturalistic settings, non-rigid facial motions intermix, providing the viewer with more complex and detailed information than would be available from viewing the individual movements in isolation. See Figure 5 for an example.

In summary, facial motions can be categorized into rigid and non-rigid motions. Social signals permeate most kinds of facial motions and thus, visual processing of the movements should yield information useful for interacting with people in the world. Moreover, as we will emphasize later, the social information embedded within these motions likely plays an important role in our ability to remember facial identity.

Before we turn our attention to the psychological studies, we will briefly outline three hypotheses that are implicit in the current body of literature concerning recognition of moving faces.

Overview of the hypotheses

By what mechanism(s) might motion help face recognition? By what mechanism(s) might it hinder recognition? Three non-exclusive hypotheses have been proposed in the literature[1]. By "non-exclusive," we mean that the validity of any one of the hypotheses does not exclude the validity of the others. In fact, these hypotheses may serve complementary roles as they apply to different parts of the process of perceiving, encoding, and remembering a face. Because recognition of a moving face occurs in different contexts (i.e., different facial motions, various viewing conditions, and either newly-learned or highly familiar faces), the respective contribution of each hypothesis is likely to be determined by the conditions of the recognition task at hand.

The *supplemental information hypothesis* posits that identity-specific facial motion is encoded in addition to the invariant structure of a face. The facial motion considered in this hypothesis consists of idiosyncratic patterns of facial motions, sometimes called "dynamic facial signatures." For example, a person might have a characteristic way

[1] The labels we use for the hypotheses were introduced formally by OToole et al., (2002). The ideas behind the supplemental information and representation enhancement hypotheses were introduced in examining face recognition with highly familiar faces. We connect the studies to the hypotheses in the next section.

Figure 5. Combinations of rigid and non-rigid movements occur in most natural situations. These combined motions unfold in natural temporal sequences and with just a few frames it is possible to capture examples of both types motion. The woman here is tilting her head, smiling, laughing, looking down, and shifting her gaze.

of smiling or grimacing or may inject certain idiosyncratic facial gestures into her speech. These characteristic facial motions can help us identify a person. A certain amount of experience with a face may be needed in order to learn an individuals characteristic facial motion style, restricting the usefulness of this mechanism to relatively familiar faces.

The *representation enhancement hypothesis* posits that facial motion can benefit recognition by enhancing the quality of the three-dimensional information available from faces. Unlike the previous hypothesis, which emphasizes the contribution of motion per se, the mechanism posited here operates by using facial motion to enhance the perception of the static structure of a face. And this in turn, is helpful for recognition. The benefits associated with representation enhancement are perceptual and so do not depend on having prior experience with a face. Thus, representation enhancement may apply equally well to familiar and unfamiliar faces.

The *motion as a social signal hypothesis* proposes that the social communication information embedded in a moving face may affect recognition, but not necessarily in a beneficial way. On the one hand, social information available from a moving face may help to attract and maintain the attention of the viewer and thus increase the likelihood that the face will be remembered later. On the other hand, the attentional demands needed to interpret the social information in a moving face could distract the viewer from processing the identity of a face. The effects of this mechanism, therefore, may be context dependent.

In the next section we provide a detailed review of the psychological studies of memory for moving faces. We note here that none of these psychological studies was designed to seek evidence for the exclusive contribution of any one hypothesis to their results. Readers less interested in the details of these studies can skip through to the section entitled “Conclusions from the Psychological Studies.”

Psychological studies of memory for moving faces

There are important differences in recognition memory performance for familiar and unfamiliar faces. Therefore, we divide this discussion of the behavioral literature on memory for moving faces into studies employing familiar or famous faces and studies employing unfamiliar faces.

As is well known from previous studies using static faces, recognition accuracy for relatively unfamiliar faces is surprisingly fallible (Hancock, Bruce, & Burton, 2000). Indeed, even relatively minor changes in viewpoint (e.g., Troje & Blthoff, 1996, OToole, Edelman, & Blthoff, 1998) and illumination (e.g., Braje, Kersten, Tarr, & Troje, 1999; Hill and Bruce, 1996) result in measurable decreases in performance when people are asked to remember, or even to match, pictures of newly learned faces (e.g., Henderson, Bruce, & Burton, 2001). By contrast, with just a brief glance across a dimly lit room, we can recognize the face of a friend. There are relatively few studies (Burton, Bruce, & Hancock, 1999). However, it is clear anecdotally that we can easily recognize the faces of people we know, even when viewing conditions are poor or when the faces are seen from a distance. There are at least two reasons for the limited availability of data on familiar face recognition. First, because we recognize familiar faces so accurately, performance, in most cases, is close to ceiling. A second problem is that familiar face recognition experiments, in theory, require the construction of a "customized" stimulus set for each participant. In practice, this is rarely done (but see Burton, Wilson, Cowan, & Bruce, 1999). Consequently, famous faces, rather than personally familiar faces, are used in most experiments of familiar face recognition. We note however, that famous faces are, strictly speaking, not the same as personally familiar faces, and indeed there may exist subtle differences between the two types of stimuli.

The performance differences that characterize familiar and unfamiliar face recognition with static images apply also to recognition of moving faces (Burton, Wilson, et al., 1999). We will see that although motion can benefit familiar face recognition, the effects of motion are less clear for unfamiliar faces.

Familiar Faces

Research with familiar/famous faces has generated cogent results that reveal an important role for motion cues in recognition. Participants in these studies are generally asked to name famous faces from spatially degraded motion displays. Spatially degrading the pictorial information is necessary for two reasons. First, as noted previously, recognition of familiar faces is generally at, or near, ceiling. Degrading the images in a video sequence eliminates these ceiling effects. Second, spatially degrading the stimuli serves to maximize subjects reliance on motion cues by minimizing the spatial information available from the image-based frames. Indeed, it is important to bear in mind that videos are simply sequences of high quality visual images. Thus, the studies we discuss have tried to focus on the benefits of motion per se, above and beyond the picture recognition that might be performed on the individual static frames.

In studies of face recognition with familiar faces, several image-processing techniques have been employed to achieve the goal of reducing the pictorial or feature- previous studies. We connect the studies to the hypotheses in the next section. based information in the images. The techniques are rather different in nature and can affect static face recognition to varying degrees, depending on the technique itself and the way it is implemented. Thus, while the goal of these manipulations is to shift a subjects reliance from the pictorial information to the motion information, the techniques are not necessarily equivalent in other ways. We define these techniques briefly.

Photographically negated images are gray scale images in which the dark-light intensity scale has been reversed. Recognition of facial images is extremely difficult from photographic negatives, even for highly familiar faces (Galper & Hochberg, 1971). Negation thus assures a substantial impairment in the participants ability to recognize faces from the pictorial information in the image, simultaneously lowering performance from ceiling levels and shifting the participants reliance toward the motion information.

Image thresholding involves converting a standard gray scale image to a one-bit-perpixel pure black and white image (i.e., only black or white). Very different effects can be achieved by setting the pixel threshold at different levels. *Image pixelation* reduces the resolution of the image by averaging pixels values within square "blocks" of the original image. This produces an image made of large square regions/pixels of uniform luminance. Although this technique reduces the resolution of the face image, it also adds high spatial frequency noise at vertical and horizontal orientations, due to the vertical and horizontal lines that appear at the edges of the uniform intensity pixel blocks. A more direct way of reducing image resolution is by *blurring* the image. In short, blurring acts as a low pass spatial frequency filter of the image. The effects of image thresholding, pixelation, and blurring on face recognition are complex and depend on the way the techniques have been implemented, and on the particular parameters used. These effects are too numerous to review here. Suffice to say that under most circumstances, all of these manipulations make recognition more difficult.

Finally, the simplest technique used to degrade images is *inversion*, for which face images are simply turned upside-down. Inversion has long been known to decrease face recognition accuracy (Yin, 1969; for a review see Valentine, 1988). See Figure 6 for examples of these techniques.

Figure 6. In recognition experiments using familiar faces, image degradation techniques are commonly used to shift a subjects reliance from the pictorial information to the motion information. The top left image is in its original form. Five image manipulation techniques are shown here. They include inversion (top right), blurring (middle left), pixelation (middle right), negation (bottom left), and thesholding (bottom right).

Studies of Familiar Face Recognition

Knight and Johnston (1997) conducted one of the first recognition experiments using dynamic famous faces. Participants viewed either 5-second video clips from taped television footage of celebrities or 5-second photographic stills extracted from the video clip. Because the moving faces were extracted from videotaped television shows, the faces included *both* rigid and non-rigid motions. Participants were asked to provide the identity of 80 faces (40 famous, 40 unknown). Knight and Johnston limited the availability of feature cues from the faces by showing photographically negated faces to half of the participants and positive images to the other half of participants. In addition, within each sequence inverted and upright faces were randomly interleaved.

Knight and Johnston (1997) reported no motion benefit for either of the inverted conditions. However, when participants viewed *upright negated* faces, the moving sequences were better recognized than the still images. This recognition advantage for moving faces disappeared when participants viewed upright positive images. Knight and Johnston suggest that motion improved recognition performance in only the upright negative (vs. the upright-positive) condition because the motion information was “largely redundant” when the viewing conditions were less challenging. They propose that seeing a face move may provide evidence about its three-dimensional structure, which compensates for the degraded featural information inherent to negative images. They also suggest that well-known faces may have characteristic facial gestures, which remain recognizable in moving sequences, despite the negated format.

In a series of four experiments, Lander, Christie, and Bruce (1999) successfully replicated and extended Knight and Johnstons (1997) results, finding a more generalized recognition benefit for dynamic faces. First, Lander et al. compared subjects recognition accuracy for faces of film and media personalities viewed either as dynamic 2.5-second video sequences or as a series of 3 static freeze-frames selected from the moving sequences. Similar to Knight and Johnsons design, subjects viewed 36 faces (24 famous, 12 unfamiliar) as either upright negated images or inverted positive images. Participants were asked to name (or provide identifying semantic information about) the faces presented. Lander et al. found that subjects recognized the faces from the video sequences more accurately than faces from the static sequences, in *both* the negative and inverted conditions.

Lander et al.s (1999) second experiment was a partial replication of their first experiment, extending the basic premise of the findings to faces presented as thresholded images. Subjects were asked to identify 40 faces (30 famous, 10 unknown) that were presented in four image formats: moving-upright, moving-inverted, static-upright, and static-inverted. Again, Lander and colleagues reported a significant advantage for naming famous faces viewed in motion in both the inverted and thresholded conditions.

In a third experiment, Lander et al. (1999) directly addressed whether the recognition advantage found with moving faces was attributable to the additional static information provided by the moving images. To test this hypothesis, the experimenters "matched" the number of views presented in the static and moving sequences. Specifically, each subject viewed famous faces presented in three formats: a 9-frame array of static images presented in sequential order; a 9-frame array of static images presented in a "jumbled" order; and a 9-frame dynamic video sequence. With this method, participants saw the same number of views, whether the faces were presented as static or moving sequences. In addition, half of the participants viewed faces that were thresholded (but displayed upright) and half of the participants viewed faces that were inverted (but shown as positive images). Lander et al.s results revealed that for both the thresholded and inverted faces, the dynamic sequences were recognized more accurately than either the sequential-static or jumbled-static formats. This specific result is important because it indicates that the motion of the faces per se was driving the improvement in participants recognition accuracy–not simply the additional views of the faces provided by the motion.

Finally, in a fourth experiment Lander et al (1999) altered the temporal characteristics of moving faces, by either slowing down or changing the rhythm of the observed motion of thresholded faces. This manipulation was accomplished by altering the relative length of each frame in the moving sequences. Lander and colleagues results showed that the naturally moving faces were better recognized than faces in the unevenly paced sequences. This finding demonstrates that the "naturalness" or fluidity of facial motion is an important component in the recognition advantage found for moving faces, especially when the static information alone is hard to recognize. This naturalness finding was replicated and extended by Lander and Bruce (2000) in a set of two experiments. In the first experiment, participants viewed two kinds of video sequences: 9-frame "ordered" sequences and 9-frame "jumbled" sequences. All the video clips showed thresholded images. Lander and Bruce found that ordered moving sequences were recognized more accurately than jumbled moving sequences, indicating that natural motionnot simply a sense of animationis an important component in the recognition advantage found with moving faces.

In a second experiment, Lander and Bruce (2000) again used thresholded famous faces as stimuli, but compared recognition between accelerated video clips, reversedmotion clips, naturally moving (forward-motion) clips, and single freeze-frame images. All of the moving sequences contained the same number of frames, but the accelerated sequences were created by increasing the frame rate by a factor of two, and thus were presented twice to participants. Lander and Bruce reported that subjects recognized more accurately the identity of the faces from the three moving sequences than from the static freeze-frame images. More importantly, the video clips in which the faces moved at normal speed and in a normal direction elicited better recognition than the speeded-up or motion-reversed sequences. The combined findings from both experiments indicate that the pattern and tempo of facial motion is critical to the recognition benefit associated with dynamic faces. Thus, recognition is most enhanced when the observed facial motion "looks natural."

If viewing a face in motion helps recognition when the images are degraded, then does the amount of image degradation affect the usefulness of motion as a cue? Lander, Bruce and Hill (2001) addressed this question using pixelated images. Participants were asked to identify 40 faces (30 famous, 10 unknown), half of which were viewed as video clips and half of which were viewed as static images extracted from the video clip. The faces were pixelated at two different levels (10-

and 20-pixels per face). In addition, two different viewing distances were used (1.5 meters or 3 meters). Lander and her colleagues found that across all conditions, the moving faces were more accurately identified than the static images. However, motion provided no incremental benefit over the two worse viewing conditions (i.e., 10-pixels per face compared to the 20-pixels per face; or for the 3-meter viewing distance compared to the 1.5-mether viewing distance).

In a second experiment, Lander et al. (2001) investigated whether motion would improve recognition of blurred faces. Subjects were asked to name 12 faces (9 famous, 3 unknown), presented at one of three different levels of blur. Half of the participants viewed the faces as moving sequences and half of the participants viewed single static images. Lander et al. found that at relatively higher blur levels, participants identified the moving images more accurately than the static images. No motion advantage was found for the least blurred moving images they tested. This finding is noteworthy because it suggests that the beneficial effect of motion is more pronounced when an image is severely degraded and the availability of structural cues to the identity of the face is minimal.

The role of motion in memory for faces was cast into a more naturalistic context in a recent study by Burton, Wilson, et al. (1999). They asked participants to pick out familiar faces from poor quality videos similar to those used in low-cost security systems. Burton et al. used actual video footage captured from university surveillance cameras. The video was edited into 20 individual video clips showing university professors entering a building. High-quality digital photographs of each of the 20 professors were also collected. In their first experiment, Burton et al. formed three groups of participants: students who were familiar with the professors; students who were unfamiliar with the professors; and a group of trained police officers who were unfamiliar with the professors. All participants first viewed 10 of the pre-edited surveillance video clips. Later, participants were shown all 20 photographs and were asked to identify the professors they remembered seeing in the video clips. As expected, the "familiar" participant group performed more accurately than either of the two "unfamiliar" participant groups.

Given that the videos contained full body, gait, and facial information about each subject, Burton, Wilson et al. (1999) designed a second experiment to determine which aspect(s) of the stimulus contributed to the familiarity advantage. This time, the experimenters included only participants who were familiar with the targets. Each participant was instructed to watch one of four different versions of the same surveillance video footage and to identify the professors they recognized. One version of the video was edited so that the professors gait was obscured by a black rectangle, leaving only the face visible. In a second version, a solid black rectangle was shown in place of the face, leaving only the gait visible. In a third version, the tape was edited so that only a series of single static frames extracted from the original video was shown, eliminating any traces of dynamic information. As a control, an unedited fourth version, which contained no changes, was used. Burton et al. found that the "face obscured" version of the tape resulted in the worst recognition performance. And so, the authors concluded that face information plays the key role in identifying someone familiar.

The study of Burton, Wilson et al. (1999) is notable for the use of personally familiar faces, rather than famous faces and also because it is one the only studies in the literature that includes participants who were both familiar and unfamiliar with the target stimuli. The last study we consider in this section is a priming study for static and moving faces. "Priming" is defined as the facilitation demonstrated at test when the to-berecognized item has been encountered previously (Lander & Bruce, 2001). Usually, priming designs are carried out by very briefly presenting (i.e., on the order of milliseconds) subjects with a stimulus before a perceptual task. It is important to note that priming designs do not assess participants recognition memory for faces, but rather, priming is used to infer the dimensions of the face that were encoded by the subjects (Bruce et al., 1994; Kolacsai & Biederman, 1997).

In a priming study using familiar faces, Lander and Bruces (2001) participants first viewed a series of moving and static "primes," each presented for 2500 ms, and were asked to name or provide semantic information about the person displayed. At test, participants viewed static images of faces and were asked make speeded familiarity judgments (i.e., "Is this face famous or non-famous?"). Lander and Bruce found that participants correctly classified the face as famous more quickly when the face had been primed as a moving, rather than a static image. In a second experiment, Lander and Bruce found that the priming advantage for moving versus static faces persisted, even when the same static image was shown in the prime and test phases. In their third experiment, Lander and Bruce reported that a moving face acts equally effectively as a prime regardless of whether the faces shown at test time are moving or static images. Finally,

in a fourth experiment, Lander and Bruce demonstrated that priming was most effective when naturally moving primes, as opposed to slow-motion primes, were used. This specific finding is especially relevant and suggests that "dynamic information is somehow central to the stored face representations" (p. 32) of well-known faces. (See Table 1 for a summary of the psychological result)].

Studies of Unfamiliar Face Recognition

To date, the evidence for motion benefiting unfamiliar face recognition tasks is inconclusive. In most of these studies, previously unfamiliar faces/people are viewed from dynamic or static displays during a learning phase and are presented at test with either dynamic or static displays. Both eyewitness-style (i.e., whole bodies and events are viewed) and more controlled facial recognition studies have been carried out. It is important to note that for unfamiliar face recognition, motion manipulations are meaningful both at learning and at test. This is in contrast to familiar face recognition studies, for which the learning conditions are not under the control of the experimenter. We will discuss this issue in more detail shortly.

For the eyewitness experiments, two early studies found conflicting results, but neither found motion benefits for learning dynamic stimuli. Schiff, Banka, and de Bordes-Galdi (1986) found a motion advantage when participants viewed a staged robbery and were tested with "dynamic mug shots," which showed faces rotating through 180-degrees from left to right profile. Yet no corresponding benefit was found for learning with the dynamic images over viewing the robbery video from two freeze frames of the robbery. Using a similar eyewitness approach, Shepherd, Ellis, and Davies (1982) found that videotaped presentations offered no benefit over single photographs at either learning or test, although "live" presentations of suspects resulted in the most accurate recognition performance.

Both Schiff et al. (1986) and Shepherd et al. (1982) incorporated an event into the learning phase of their study. As such, their focus, and indeed the strategies of the participants, may likewise have been aimed more at witnessing the composite event, with only a secondary priority placed on remembering the faces of the individuals involved. The remainder of the studies reviewed in this section focus specifically on memory for faces.

The first of these studies provides the strongest evidence to date that recognition memory for unfamiliar faces can benefit from learning dynamic stimuli. In a series of four experiments, Pike, Kemp, Towell, and Phillips (1997) compared recognition performance for faces presented as either single static images, 5- or 10-view multiple static images, or as dynamic video clips. Pike et al. created the face stimuli by videotaping subjects seated in a rotating chair. During taping, the chair rotated through a full 360 degrees. As such, the facial motion included in this experiment was purely rigid motion.

In the first two experiments, participants viewed faces in one of the three learning conditions. In the dynamic condition, participants saw a 10-second video clip of each face. In the multiple static image condition, participants viewed time-matched presentations of 5 images extracted from the video at equally spaced angles. In the single static condition, participants viewed the frontal image for 10 seconds. At test, all participants viewed static frontal images of target and distracter faces and were asked to indicate the faces they had seen previously. In order to eliminate face-matching strategies, the test faces were duplicate images,

Table 1
Key findings from the psychological studies with moving faces.

Study	Type of Face	Motion Advantage?	Major Findings
Shepherd, Ellis, & Davies, 1982	Unfamiliar	No	Videotaped presentation of suspects offered no benefit over single photographs.
Schiff, Banka & deBordes-Galdi, 1986	Unfamiliar	Yes	"Dynamic mugshots" of suspects elicted better recognition than freeze frame image.
Bruce & Valentine, 1988	Familiar	Yes	Subjects were able to recognize friends from moving point light displays.
Knight & Johnston, 1997	Familiar	Yes	Motion helped subjects recognize negative images of faces.
Pike, Kemp, Towel,l & Phillips., 1997	Unfamiliar	Yes	Subjects recognized faces learned as dynamicsequences versus multiple-static or single-static images.
Christie & Bruce, 1998	Unfamiliar	No	Dynamic sequences offered no benefit over multiple static images at test; learning faces as dynamic sequences conferred *disadvantage.*
Bruce, Henderson, Greenwood, Hancock, Burton, & Miller, 1999	Unfamiliar	No	Dynamic sequences offered no benefit over "best" single static image extracted from video.
Burton, Wilson, Cowan, & Bruce 1999	Unfamiliar and Familiar	Yes	Subjects who were familiar with the targets more accurately identified the faces from surveillance video than unfamiliar subjects.
Lander, Christie, & Bruce, 1999	Familiar	Yes	Dynamic sequences of degraded images were recognized more accurately than static images, even when number of views was matched.
Lander & Bruce, 2001	Familiar	Yes	Moving primes faciliatated recognition more than static primes.
Lander, Bruce, & Hill, 2001	Familiar	Yes	Degraded images of moving faces recognized more accurately than single static images.
Bruce, Henderson, Newman, & Burton, 2001	Unfamiliar	Yes/No	Moving faces recognized more accurately when subject pairs talked about the faces during learning trials; longer exposure times to moving versus static faces offered no benefit.
Hill & Johnston, 2001	Unfamiliar	Yes	Participants were able to identify pre-learned facial motion sequences that were projected onto synthetic 3D heads.
Knappmeyer, Thornton, & Buelthoff 2001	Unfamiliar	Yes	Pre-learned sequences of facial motion biased subjects' subsequent identity judgments of intermediate morphs
Thornton & Kourtzi, 2002	Unfamiliar	Yes	Dynamic prime images facilitated matching responses relative to a single static prime.

collected on separate days from the face images used for the learning trials. The only difference between the first and second experiments was the type of frontal image presented at test. In the first experiment, a novel slide of the subjects taken under different illumination conditions, was presented. In the second experiment, an image extracted from the learning videotape was presented. In both cases, Pike et al. found that faces learned in the dynamic condition were recognized more accurately than faces learned in either of the static conditions.

The third and fourth experiments were designed to eliminate the alternative hypothesis that the improvements seen with motion were due to presenting more views rather than to motion per se. By adding additional views to the multiple view conditions in the third experiment, Pike et al. (1997)

demonstrated that the advantage of the dynamic condition was not due simply to providing extra views. By presenting the multiple views in random order, Pike et al. eliminated the potential for apparent motion cues to play a role in the multiple static conditions. The randomly ordered multiple view conditions again yielded worse performance than the dynamic condition, and proved no different than the ordered multiple view conditions.

Pike et al. (1997) suggest that the dynamic information available from a moving face helps to create a mental representation that is more accurate and more robust than static images, even when the number of views is equated between static and dynamic conditions. Moreover, the results of this study strongly support the notion that the motion per se, rather than the extra view information it provides, has a beneficial role for recognition.

Christie and Bruce (1998), however, using generally similar methods did not replicate Pike et al.s (1997) finding that motion benefits recognition. Christie and Bruce compared recognition between learning conditions using either a moving sequence of 5 frames (dynamic presentation) or a series of the same 5 frames presented in random order (static presentation). In one set of experiments, the dynamic presentation consisted of expressive motion (i.e., non-rigid motion consisting of faces moving from smile to sad). In a second set of experiments, the dynamic motion consisted of rigid head nodding and shaking. Unlike Schiff et al (1986) and Pike et al., Christie and Bruce included a direct examination of transfer effects to different test stimulus types (e.g., static-static, staticdynamic, dynamic-static, dynamic-dynamic). Viewpoint conditions (e.g., full face or three-quarter) and motion type (e.g., expressive changes vs. facial speech; or head nodding vs. head shaking) were also either matched or mismatched between learning and testing conditions.

In only one case, across the four experiments, did Christie and Bruce (1998) find a difference between the moving and static presentation sequences in the learning phase. Moreover, this case favored learning from the static stimulus. The disadvantage found for the dynamic presentation of faces is in direct conflict with the results of Pike et al. (1997), but is consistent with the results of Schiff et al. (1986). In none of the four experiments, did the effect of test condition reach significance. In a combined analysis collapsing across the four experiments, however, Christie and Bruce found a marginally significant advantage for testing with dynamic sequences.

Comparing their results with Pike et al. (1997), Christie and Bruce (1998) discuss a number of possible minor differences between the studies that might have contributed to the differences in results. We will discuss some of these differences in the conclusion of this section. Notwithstanding, the discrepancy between the findings remains puzzling, as none of these factors challenges the basic validity of the methods or stimuli used in either study.

In a subsequent related experiment, Bruce et al. (2001) found no difference between moving and still images for the matching of unfamiliar faces in conditions similar to those used by Christie and Bruce (1998).

Bruce, Henderson, Greenwood, Hancock, Burton and Miller (1999) also did not find a benefit for recognizing moving faces. They tested subjects ability to match a target face from a 10-face array of photos. Using unfamiliar faces as stimuli, participants viewed 40 faces in one of three conditions. The stimulus for the dynamic condition was a 5-second video clip condition showing a subject: a.) rotating his/her head 20-degrees; b.) looking up and down; c.) speaking a number; or d.) smiling. The single static image was a "best" still image extracted from the video and presented for 5 seconds. Finally, in the "free-viewing" condition, participants could play, pause, or rewind the video as often as they wanted, before choosing the matching face from the array. Bruce et al.s results showed that there was no difference between the two "limited viewing" conditions, (i.e., the single static image and moving video image conditions). The unlimited video condition yielded better accuracy and confidence, although the error rate was still high.

Taking a different approach, Bruce, Henderson, Newman and Burton (2001) studied the effects of familiarizing participants with initially unfamiliar faces during the course of the experiment. The authors asked subjects to view a series of video clips that showed faces rotating 360-degrees, looking up and down, and then smiling. In the first experiment, participants viewed the motion sequences for either 30-seconds or 1-minute. Bruce and colleagues found no differences in subjects subsequent ability to pick out the target faces from high-quality 8-face arrays that contained similar-looking distracters. However, in a second experiment, Bruce et al. divided participants into three viewing groups. One of the groups consisted of individual subjects, and the other two groups consisted of subject pairs. The "social exposure group pairs" were instructed to chat about the faces during viewing, while the other group of subject pairs was instructed not to discuss the faces. All participants were tested individually. Bruce et al.s results showed that the participants who were instructed to talk about the faces during the learning trials recognized the faces better than the other participant groups. Bruce et al.s results demonstrate that the nature of the experience people have with a moving face seems more important than simply the amount of experience people have with a moving face.

As we did for the section on familiar faces, we end this section with a priming study. Using unfamiliar faces as stimuli, Thornton and Kourtzi (2002) demonstrated a recognition advantage for moving primes. On each trial, observers were shown two faces in quick succession and were asked to make a speeded matching response. On half of the trials, participants first viewed a briefly presented dynamic "prime;" on the other half of the trials, participants viewed a static prime. Immediately after each prime, subjects were quickly shown a static image of a face. The participants task was to decide if the two imagesthe prime and the targetbelonged to the same person. Half of the trials contained faces that were expression matches to the prime (e.g., smile-smile or frownfrown) and half of trials contained expression mismatches to the prime (e.g., smile-frown or frown-smile). As a result, there were four blocks of test trials: same identity/different expression; same identity/same expression; different identity/same expression; and different identity/different expression. Thornton and Kourtzi reported that for the trials in which the prime and the target were "matches" (i.e., the same person), the moving primes elicited faster reaction times only for the same identity/different expression condition. However, in a second experiment in which the target faces (and not the primes) were inverted (i.e., upside-down), the moving primes elicited faster reaction times for both same identity/different expression trials and same identity/same expression trials.

Thornton and Kourtzi (2002) conclude that dynamic information from a moving face is most helpful in situations where the processing demands are higher, such as when participants must make identity judgments across changes in expression (Experiment 1) or view (Experiment 2). [See Table 1 for a summary of the psychological findings].

Conclusions from Psychological Studies

To date, the psychological studies indicate a fundamental difference in the effects of motion on memory for familiar and unfamiliar faces. Specifically, there is evidence that dynamic information can be useful for recognizing familiar faces (Knight & Johnston, 1997, Lander et al., 1999, 2001; Lander & Bruce, 2000; Burton et al., 1999). This seems to be measurable only when the quality of the spatial information has been degraded. These findings are important because they suggest that motion can serve as a back-up cue for recognizing people we know well. This is especially true under viewing conditions that make recognition based on pictorial information alone challenging. In other words, motion may make the most noticeable contribution to recognition of familiar faces when it is done from a distance, in poor illumination, or under non-optimal viewing conditions.

For unfamiliar faces, it is not clear that motion can benefit recognition and there are some hints that it may even hinder recognition (Pike et al., 1997, Christie & Bruce, 1998, Bruce et al., 1999). The primary goal of the published studies has been to examine the effects of motion, broadly defined, on memory for faces. The divergence of results indicates that a closer look at the specifics of these studies is needed to reconcile the findings. We think that three factors emerge as potentially important for explaining the differences we see in the results of recognition studies with unfamiliar faces in motion.

The most salient factor is the *type of motion* tested. Different facial motions may have different effects on memory for faces. Lets take the divergence in findings of the Pike et al. (1997) and Christie and Bruce (1998) studies as an example. Both studies used rigid rotations of the head. Pike et al. used rigid head rotations of 360 degrees, which were generated by rotating the subject in a chair. Christie and Bruce employed less passive head motions, in which the subjects nodded and shook their heads. Though both of these motions are "rigid," the motions differ in the perceptual and social information they provide to an observer. The head nodding and head shaking motions employed by Christie and Bruce can be interpreted in a social context. These motions are also active motions of the subject, rather than the camera. By contrast, the motion used by Pike et al. involved a passive subject in a rotating chair. This approximates head motions that are experienced by a moving viewer, rather than a moving subject. This latter motion may enrich the perceptual information available to a participant via structure from motion, for example, but is unlikely to be interpreted as a social signal. Moreover, much more of the head was presented to observers in Pike et al.s study than in Christie and Bruces study. Given these differences, it is perhaps not surprising that Pike et al. and Christie and Bruce found different results. The exact reason for these differences, however, is not clear. To date, there has been no systematic attention paid to the types of motion employed in these studies. Moreover, motion types have not been considered specifically in terms of the quality and quantity of the social or perceptual information they provide.

A second relevant factor has to do with the possibility of motion effects to occur both at learning and at test. This opens up the possibility for motions introduced at learning and/or test to interact. Again, this factor applies only to experiments using unfamiliar faces for which experimenters have control over the learning and test conditions. For familiar face recognition experiments, the effects of facial motion can be assessed only at test. By contrast, most (but not all) unfamiliar face studies introduce motion as a learning variable and subsequently test subjects recognition of a faces as static images (e.g., Pike et al., 1997; Bruce et al, 1999; but, see Christie & Bruce, 1998 and Schiff et al., 1986). In any case, the ability to introduce motion at either learning or testor bothadds complexity to the experimental designs using unfamiliar faces and makes it difficult to directly compare results from unfamiliar face recognition experiments with results from familiar face recognition experiments.

A third variable is the type of recognition task employed. Standard old/new recognition tasks are used most commonly (e.g., Pike et al. 1997; Christie & Bruce, 1998). In these tasks, subjects indicate whether a face is one they remember seeing during the learning session ("old") or whether the face was not presented previously ("new"). For face matching tasks (e.g., Bruce et al., 1999), participants first view a face and at test time are asked to pick out the same face from an array of similar-looking distractor faces. These two tasks make rather different demands of participants. Matching tasks involve discriminating between a single target face and several distracter faces on each test trial, whereas the old/new paradigms do not require participants to "rule-out" several distracters at once. The latter task also requires information to be retained longer in memory. Priming paradigms (e.g., Thornton & Kourtzi, 2002) provide information about how quickly a participant can identify a previously presented face, but can provide only indirect information about the effects of motion on memory. These paradigms have been used to gather evidence about the nature of the face representations based on the assumption that representations can be pre-activated by appropriately structured stimuli. Clearly, each of these designs can provide useful information about how facial motion affects recognition, but the variety of methods used complicates the task of comparing and contrasting the various results.

In summary, familiar face studies employ generally similar methods, making it relatively easy to conclude that facial motion helps recognition. The experimental paradigms and stimuli used in unfamiliar face experiments are more diverse. This complicates the task of drawing firm conclusions from the data.

Integrating the Psychological Data into the Hypotheses

Given that facial motion can affect recognition accuracy under some circumstances, the next question is: *How* might this occur? As previously discussed, three rather different hypotheses suggest that dynamic information may have several diverse roles to play in memory for faces. Again, we note that none of these hypotheses excludes any of the others. Rather, the different mechanisms may affect performance simultaneously, trading off advantages and disadvantages during a recognition task. The major challenge then, is to design experiments that enable a measure of the contribution of each hypothesized mechanism to recognition performance. We present the hypotheses formally here and evaluate the nature of the evidence that supports them. What do the psychological studies tell us about the validity of each of the hypotheses? Unfortunately, none of the experiments was conducted with the specific purpose of assessing the relative contributions of these various mechanisms to memory for moving faces. The experiments nevertheless provide useful data for considering the more general principles that might apply. Further, it is important to sort through the logic of implementing empirical tests of the various hypotheses for designing future studies. This will be needed to advance from the present empirical focus of past studies to a more theoretically grounded approach in future work.

Supplemental information hypothesis

The supplemental information hypothesis posits that, in addition to encoding the invariant structure of a face, we also encode identity-specific facial motions, in the form of dynamic facial signatures. This hypothesis has been proposed explicitly or implicitly in much of the empirical work on recognition of familiar faces in motion (Knight & Johnston, 1997; Lander et al., 1999, 2001; Lander & Bruce, 2000, 2001; Burton, Wilson, et al., 1999). These characteristic movements may form a part of our representation of the identity of a person. Notably, this hypothesis suggests the existence of an inherently dynamic visual representation of faces in the brain: a "mental video clip" of sorts.

We can make two logical assumptions about the nature of dynamic identity signatures that support their relevance in familiar face recognition studies. First, dynamic identity signatures require more time and experience to learn than the static structure of a face. For example, an observer can learn a particularly distinctive feature (e.g., deep-set eyes) from a single encounter with a face. However, a single encounter with a moving face is not sufficient to determine whether a particular sequence of facial movements is characteristic of a person. Repeated encounters with the person would seem to be necessary to establish which motions are "characteristic" of a person, and which are generated randomly.

Second, dynamic identity signatures are inherently less reliable for identification than static feature information. This is an obvious consequence of the fact that characteristic gestures are generated only intermittently. An observer cannot count on a particular facial motion being present when an identification decision is needed. Indeed, as behavioral studies using degraded images of famous faces demonstrate, people may use dynamic information to recognize someone, only if the more reliable pictorial information is unavailable or difficult to access (e.g., Knight & Johnston, 1997; Lander, Bruce, & Hill, 1999).

The most intriguing and provocative aspect of this hypothesis is that it suggests the direct encoding of inherently dynamic, identity-specific information about faces. This represents a fundamental departure from, or addition to, the classic assumption that the identity of faces (and perhaps objects) is encoded via feature sets that capture the invariant structure and configuration of a face. We know of no direct evidence that the visual system retains memory traces of moving faces (though we consider this topic in more detail in the neural framework section). Notwithstanding, proof of the human ability to identify faces based only on their characteristic or idiosyncratic motions was provided recently in two studies. The studies used animated synthetic three-dimensional head models in which the identity of faces from the pictorial cues was either unspecified (Hill and Johnston, 2001) or altered systematically and "played off against" the identity information specified through motion (Knappmeyer, Thornton, & Blthoff, 2001).

Hill and Johnston (2001) projected facial animations generated by human actors onto a computer-generated average head (Blanz & Vetter, 1999). Participants learned to discriminate among four individuals based solely on the facial motion information that was projected onto the structurally unchanged head model (see Figure 7). In another study, Knappmeyer et al. (2001) trained participants to discriminate two synthetic faces that were animated with different characteristic facial motions. When later viewing morphs, created by combining the two head models, the participants identity judgments about the intermediate morphed heads were biased by the animated motion information participants originally learned to associate with the faces.

Given that supplemental motion cues are available and accessible to human perception, to what extent have these motions contributed to behavioral findings with moving faces? Logically, supplemental motion information could account for someor even allof the advantages seen in the familiar face studies we reviewed. The stimuli used in these experiments, usually taped interviews and TV appearances, almost certainly contain idiosyncratic motion cues to the identity of famous faces. Indeed, our ability to appreciate celebrity impersonators indicates that we are familiar with the characteristic gestures of famous people. Impersonators, who have faces quite different from the celebrities and politicians they mimic, use dynamic identity signatures very effectively to convey an impression of the person they portray.

Yet a problem remains. Despite evidence consistent with the supplemental information hypothesis, familiar face recognition experiments cannot exclude a role for the representation enhancement hypothesis. Structure-from-motion analyses, a key mechanism of the representation enhancement hypothesis (see section below), may also be at work in experiments with familiar faces and may be especially useful in the case of spatially degraded stimuli

Representation enhancement hypothesis

The representation enhancement hypothesis posits that facial motion can improve recognition by enriching the quality of the three-dimensional structural information accessible from a human face (Pike et al., 1997; Christie & Bruce, 1998; Lander & Bruce, 2001; Lander et al., 1999, 2001; Bruce et al., 1999, 2001; Thornton & Kourtzi, 2002). The mechanism proposed by this hypothesis differs substantially from the supplemental information hypothesis, because it does not involve a direct representation or encoding of inherently dynamic visual information about faces. Rather, facial motion information benefits recognition by bootstrapping the encoding of the invariant structure and features of the face. Due to the inherently perceptual nature of these motion benefits, the representation enhancement mechanism should be equally beneficial for both familiar and unfamiliar faces.

This hypothesis draws implicitly on structure-from-motion phenomena, which have been studied extensively in psychology and in computer vision (Johansson, 1973; Ullman, 1979). It is well known that the visual system can extract the structure of an object by the cohesion of a relatively small number of object sample points in motion. The kinetic depth effect (Wallach & OConnell, 1953) is a compelling example of how the three-dimensional structure of an object or scene can be revealed by motion. Viewed in a stationary display, the sample points give no hint of the shape or structure of an object. Once the points are set in motion, however, a complex and three-dimensional object structure is revealed. Kinetic depth effects provide a proof that motion, by itself, is sufficient to specify relatively complex aspects of three-dimensional object structure.

The classic point-light displays of biological motion created by Johansson in the early seventies are a convincing demonstration of the power of motion information to specify complex biological forms (Johansson, 1973). These displays provide an important control in studying the role of motion in form perception because they isolate motion information from the pictorial cues, (e.g., shape-from-shading), usually available. To cre-

Figure 7. Hill and Johnston (2000) projected facial motion sequences from real human actors onto synthetic heads. Participants learned to discriminate among four individuals based solely on the facial motion information that was projected onto the structurally unchanged head model. (Reprinted from *Current Biology*, **11**, Hill, H. and Johnston, Memory for moving faces).

ate a pointlight display, small patches of lights or light reflectors are attached to the body joints of a person wearing dark clothing, so that in low levels of illumination, only dots of light are visible. When the person is stationary, we see a formless cloud of dots. As soon as the person begins to walk, dance, jump, or otherwise move in a natural fashion, we immediately perceive a human form in motion. The human perceptual system can extract, from these simple displays, characteristics of both the model (e.g., gender, Cutting & Kowalski, 1977; identity, Cutting & Kowalski, 1977; Stevenage, Nixon, & Vince, 1999; Hill & Pollick, 2000) and the motion (e.g., walking, jumping, Dittrich, 1993).

More direct applications to face recognition can be found in several studies using point-light faces. In these displays, reflective dots are scattered on the surface of a moving face (Bassilli, 1978, 1979; Bruce & Valentine, 1988). When the brightness of these displays is reduced so that only the patches of light are visible, people can accurately judge the gender and age of a face (Berry, 1990). Viewing point light faces while speaking can also improve speech comprehension (Rosenblum & Saldaa, 1996; Rosenblum, Johnson, & Saldaa, 1996).

The quality of *identity* information that can be derived from point light displays of faces, however, is less certain. Although Bruce and Valentine (1988) found that subjects were able to recognize the faces of their friends from moving point light displays, accuracy was generally quite poor. The interpretation of point light studies of faces in the context of the supplemental information and representation enhancement hypotheses is limited by the fact that dynamic identity signature information and structure-from-motion information are intermixed in these displays. Historically, biological motion studies have stressed the importance of structure-from-motion information rather than the dynamic identity information. In cases where the participants know the faces (e.g., Bruce & Valentine, 1988), it is impossible to determine the extent to which point light displays facilitate recognition based on structural information versus supplemental characteristic motion information.

A look at the perception of gender from point light walkers is informative in understanding the difference between "motion pattern" and "structure" information that might be specified by these displays. Gender perception from point light displays is partially dependent on structural cues that are revealed by the animation of the walker (e.g., center of gravity and shoulder-to-hip ratio), and partially dependent on characteristically male and female walking styles (e.g., hip swing, see Stevenage et al., 1999, for a recent overview). Likewise, the motions of a point-light face might reveal some components of the structure (e.g., configuration of the features), but might also capture characteristic motions of the face. At present, there is no direct evidence for sorting through the structure versus dynamic signature cues in the perception of facial identity from point light displays.

This returns us to the question of whether it is possible to get a relatively pure measure of the viability of the representation enhancement hypothesis in unfamiliar face recognition studies. Logically, by controlling the learning and test conditions, it is possible to ensure that supplemental information (e.g., person-specific patterns of facial motion) is not available as a cue to the identity of a face. The use of unfamiliar faces can be informative in this regard, because short-term recognition paradigms probably do not allow sufficient contact with the faces to acquire a working knowledge of characteristic motions. Further, even if it were possible to learn these motions from a single video clip, the contribution of supplemental information can be eliminated by testing with a static stimulus. By testing with a static image, the study of Pike et al. (1997), and some conditions of Christie and Bruces (1998) study, meet the criteria for eliminating any possible contribution of supplemental motion to recognition. Yet, despite their similarity, the results of these two studies do not converge. Remember that Pike et al. (1997) found a recognition advantage for faces learned as moving images–a result which offers clear support for the representation enhancement hypothesis. However, recall also that Christie and Bruce (1998) reported no recognition advantage, either at learning or at test, for moving faces–a result which offers no support for the representation enhancement hypothesis. Differences in the social content of the movements may be helpful in understanding the difference in results (see below).

Just as it is difficult to tease apart the contribution of the supplemental information and representation enhancement hypotheses in familiar face recognition experiments, it is also problematic to obtain a pure measure of the representation enhancement hypothesis in unfamiliar face recognition experiments. Specifically, one cannot exclude a role for the social signal hypothesis in unfamiliar face recognition studies. This predicament is most relevant when the motion is introduced at learning. As we argued in the first section of this paper, most facial motions can be interpreted as social signals. The primary tenant of the social signal hypothesis is that the social content of facial motions affects the attention paid to the facial identity, and thereby, affects recognition accuracy. Although different kinds of socially-laden face motions may affect identity processing in different ways, most of the unfamiliar face recognition studies we cite allow for a role for the social content of the motions to impact identity processing.

Interestingly, the motion used by Pike et al. (1997) might be a notable exception to the rule that facial motions are "social." Recall that subjects in this experiment were filmed by a stationary video camera, while sitting in a motorized rotating chair. The perception of motion experienced by participants in this study was probably more akin to that experienced from the motion of a moving observer than from the motion of a selfanimated face. This is the only study that defines facial motion in this passive way. We assume that the passivity of this motion may make it less likely to be interpreted in a social context. Notably, Pike et al. (1997) is the only unfamiliar face recognition study that reports consistent motion advantages for recognition. This brings us to our next hypothesis.

Motion as a social signal hypothesis

Social communication information is embedded in the movements of the face. In natural settings, face processing involves simultaneous attention to both the identity and social information in a face. Who is this person, and what is she trying to communicate to me? The social signal hypothesis posits that the effects of motion on memory for faces are mediated by the constraints imposed by the social information processing of facial motions. Although most behavioral studies have sought to identify situations in which facial motion can benefit recognition, it is clear that this hypothesis predicts that, in some cases, motion could hinder recognition. Specifically, the simultaneous processing of identity and social communication information from a face may be a divided attention task, where accurate processing of one type of information disturbs the processing of the other type of information.

Under other circumstances, the social signal hypothesis predicts a recognition advantage. For example, in some situations (when we are busy or pre-occupied), we may pay only limited attention to the identity of the people we encounter. At other times, however, facial movements may engage and potentially focus our attention on a person, encouraging identity processing that would otherwise not have occurred, or may have occurred in a more cursory fashion. These effects are most likely to occur when we are learning a new face. (i.e., it seems unlikely that processing social signals from a face would impede or facilitate the identification of someone we already know). Thus, the social information processing of facial motion is most likely to affect memory for unfamiliar faces. To the best of our knowledge, this hypothesis has not been suggested in the context of previous studies.

The social signals conveyed by facial motions may affect recognition accuracy in a variety of ways. Although the face studies we review do not explicitly consider how the social aspects of facial motion affect identity processing, there are some hints about the potential interaction between social and identity processing in other studies, not designed to test face recognition. One study that comes to mind provides a real-world demonstration of the potential for identity processing to be overwhelmed by social processing. In Simons and Levins (1998) study, an experimenter, posing as pedestrian on a college campus, asks an unsuspecting participant for directions. Midway through their conversation, the face-to-face interaction is interrupted briefly by two construction workers carrying a door between the experimenter and the participant. During this brief separation, the experimenter asking for directions changes places with another similarlooking experimenter. Remarkably, Simons and Levin found that 60 percent of participants failed to notice the “person change” and continued to give directions as if nothing had happened. In subsequent experiments, Simons and Levin discovered that the likelihood of detecting a “person change” varied as a function of social status differences between the experimenter and the subject. For example, person changes were detected more frequently when the subject and experimenter were close in age or matched in social cohort. In this latter case, when the participant was a university student, the person change was less likely to be detected if the experimenter posed as a construction worker, wearing a hard hat and workmans clothes. This was true even when the subject and experimenter were similar in age.

Simons and Levins (1998) findings are reminiscent of the mechanisms suggested in the social signal hypothesis. Namely, we may fail to notice or attend to identity information in situations where the processing demands necessary to actively participate in a social interaction are high, and possibly also, when ones interest in the identity of a person is low. To give directions, a participant must conjure up a mental map of campus, locate the present and desired locations of the experimenter, map out a route between the two locations, and then translate a spatially based map representation into a verbally expressed set of instructions, replete with useful landmarks. Further, the attention of the participant to the experimenters identity is also sapped by the fact that the identity of the requester is generally irrelevant in a case like this. It is unlikely that we will need to remember the face of this stranger, because we do not anticipate encountering them in the future (unless we have given them really bad directions).

Simons and Levins (1998) experiment showed that identity processing can be overwhelmed by social-processing demands that are part of our everyday interactions with others. To what extent might this kind of explanation help us to understand the results of the studies on moving faces?

There is no direct evidence to date that social signals from a moving face distract the observer away from processing face identity. There is likewise no direct evidence that social signals from a moving face might focus more attention on identity than would have otherwise occurred. However, the study of Bruce et al. (2001) provides an intriguing hint that focusing attention on a person, in a more social context, improves recognition accuracy. Recall that Bruce et al. (2001) found that participants who talked about the faces performed better than participants who did not. These effects are highly reminiscent of the classic study of Bower and Karlin (1974) on depth of processing with faces. The authors found that faces judged for the "deep characteristics" of personality traits were recognized more accurately than faces judged for "surface characteristic" of sex.

As we have found with the previous two hypotheses, the task of isolating the contribution of this hypothesis to successful/unsuccessful recognition of a moving face is difficult. The problem lies in directly manipulating the variables that affect attention to the identity of a face. This is a challenging, but not impossible task. In fact, standard attention paradigms have been applied previously to the task of assessing the independence of identity and facial expression processing in faces. (We will argue in more detail in the neural framework section of this paper why we think that the processing of facial expressioneven from static images of facesmay be related to processing motion information from faces). For present purposes, these attention paradigms provide a useful illustration of how one might tease apart interference and facilitation effects involved in the simultaneous processing of both identity and social/emotional content from faces.

A long-standing tenant in face perception is the functional independence between facial expression and identity processing (Bruce & Young, 1986). The initial evidence for this claim came from neuropsychological cases of double dissociations of expression and identity. For the case of expression and identity, double dissociations occur when the processing of identity and expression can both be impaired selectively. Indeed, some prosopagnosiacs, who are unable to recognize faces, nonetheless retain the ability to perceive facial expressions accurately (Kurucz & Feldmar, 1979; Kurucz et al., 1979; for a review see Damasio, Damasio, & Van Hoesen, 1982). Conversely, some patients with an impaired ability to process facial expression retain the ability to recognize faces (Bruyer, Laterre, Seron, Feyereisen, Strypstein, Pierrard, & Rectem, 1993; Shuttleworth, Syring, & Allen, 1982).

More recently, psychological studies using the Garner speeded classification task have partially supported a dissociation of processes for expression and identity. The Garner task was developed to test the (in)dependence of stimulus dimensions. By this method, two stimulus dimensions are considered independent if people can attend selectively to either dimension, while ignoring irrelevant variations in the other dimension. For expression and identity, early evidence indicated mutual independence of processing (Etcoff, 1984).

More recent studies using Garner tasks have suggested asymmetric dependence rather than independence between expression and identity (Schweinberger, Burton, & Kelley, 1999; Schweinberger & Soukoup, 1998). Specifically, these studies found that identity judgments are independent of variations in expression, but that expression judgments are influenced by irrelevant variations in identity. A similar result was found for the processing of identity and facial speech (Schweinberger & Soukoup, 1998). The basic findings for expression and identity indicate at least some ability for identity information to leak into the processes responsible for the analysis of expression and facial speech. We will return to this issue in the section concerning the neural processing of moving faces.

To conclude, an assessment of the contribution of the motion as social signal hypothesis will require the manipulation of the parameters that mediate attention to the identity of a face. This is, at present, an unsolved empirical challenge.

Summary

The supplemental information, representation enhancement, and social signal hypotheses suggest three diverse ways in which facial motions can affect face recognition. The supplemental information hypothesis is most relevant for the processing of familiar faces and may boost recognition performance when viewing conditions are sub-optimal.

There is supportive but not conclusive evidence for this hypothesis in the familiar face recognition studies that find motion benefits for recognition. The representation enhancement hypothesis, however, cannot be excluded from a possible role in these findings.

There is much less direct support for a pure form of the representation enhancement hypothesis. Though widely cited as a mechanism by which motion could help recognition, studies using unfamiliar faces provide the only reasonable test bed for this hypothesis, and so far, these studies do not consistently show motion benefits. Moreover, if such benefits are found, the role of the social signal hypothesis cannot be excluded. As we have pointed out, paying attention to the social information from a moving face may actually counteract any potential recognition benefit brought about by representational enhancement. Indeed, if representation enhancement is a useful tool for recognition, but facial motions distract or divert attention away from the identity of the face, then the two results could balance out to produce the null effects we see in several studies (e.g., Christie & Bruce, 1998; Bruce et al., 1999).

Finally, the studies to date offer some hints, although no direct evidence, for a possible role of the social signal hypothesis. Again, the relevant factors are difficult to manipulate explicitly. One possible related factor to consider in future studies is the type of facial movement employed. The extent to which different facial movements engage attention to social information rather than to identity information might prove an interesting avenue to explore.

Developmental Insights

Developmental psychologists have long appreciated the importance of movement for infants attention to and processing of naturalistic stimuli. An undercurrent of much of the developmental research is the social importance of moving faces to young infants. In this section, we will emphasize the findings that relate to the question of how infants extract identity from moving faces. Although connections between the adult and developmental literatures are not straightforward, we believe that useful and insightful links between these literatures can and should be made. The primary hurdle to overcome in making these links comes from the broad differences in the kinds of questions that typify research in these two domains and obvious differences in methodologies. If we can look beyond these differences, there is a store chest of valuable insights to be gained from infant research.

The studies we review in this section support the idea that motion helps infants acquire information about faces. In fact, motion may impact infant face perception and memory in ways that are similar to the mechanisms we have outlined for adults. Namely, data regarding infants ability to abstract structure that is revealed through movement are consistent with the representation enhancement hypothesis. In addition, data regarding the adjustments adults make to maintain the infant's attention during face-to-face interactions map onto, and even suggest potential refinements of, the motion as social signal hypothesis. However, the supplemental motion hypothesis is probably less relevant for infants than the other two hypotheses, although findings in the infant imitation literature suggest a tenuous connection (see Meltzoff & Moore, 1999). Importantly, the wealth of data in the developmental literature enables a relatively fine-grained analysis of how motion affects attention, and concomitantly, how attention in turn may affect the acquisition of information from faces.

With this perspective in mind, we begin our review with some basic perceptual considerations that underscore the salience of motion information for the infants visual world. Next, we look at evidence that young infants are sensitive to and can discriminate the basic categories of facial motion (i.e., rigid and non-rigid) that we discussed in the context of the adult studies. Finally, we will argue that findings in the developmental literature suggest two additional hypotheses, both of which invoke the mechanism of attention to explain how motion might affect face recognition.

Perceptual Considerations

Marr (1982) proposed that the purpose of vision is to recover the three-dimensional structure of objects from the two-dimensional images available on the retinae. Shading, occlusion, shadows, perspective, and a host of other image-based features provide clues to the three-dimensional structure of objects. The structure of objects can also be determined through motion (Marr, 1982; Ullman, 1979). Studies in the adult face recognition literature suggest that the relative importance of motion for face recognition increases when the more reliable pictorial information is unavailable or difficult to access. It is interesting to note that "non-optimal viewing conditions," which hinder the acquisition of pictorial information for adults, may approximate the normal state of perceptual affairs for young infants.

The visual system develops considerably in the first year of life with the most salient changes occurring in spatial resolution (Atkinson, 1998; Banks & Crowell, 1993; Danemiller, 2001; Mauer & Lewis, 2001) and in the infants ability to make use of image-based cues to object structure (e.g., Granrud, Yonas, & Opland, 1985; Yonas & Arterberry, 1994). Although the role of movement in object perception is complex (see Burnham, 1987 for a review), neonates look significantly longer at a moving stimulus when paired with a stationary one (Slater, Morison, Town, & Rose, 1985). Additionally, and perhaps surprisingly, differential tracking of moving faces has been observed in the first hours of life (Goren, Sarty, & Wu, 1975; Johnson, Dziurawiec, Ellis & Morton, 1991). These findings, combined with the fact that faces are one of the most interesting objects in the infants perceptual and emotional world (Bushnell, 1998; Nelson, 2001), suggest that motion may play an important role in how infants acquire information from faces. Motion may thus help to "tune-up" the less functional pictorial analysis tools that will come to characterize the adult visual system.

Attentional Considerations

The stimulus characteristics of human faces attract the attention of young infants. For example, young infants are attracted to areas of high contrast (Bronson, 1990), curvilinearity (Fantz & Miranda, 1975), and symmetry (Bornstein, Ferdinandsen, & Gross, 1981). Additionally, the faces infants see during interactions with caregivers are characterized by distinctive facial movements. Caregivers use a distinctive interaction style with young infants in which they produce speech and coincident facial movements with properties that engage and maintain infants attention (Fernald & Simon, 1984; Papousek, 1992). The facial expressions that accompany infant-directed speech tend to be exaggerated, repeated, slowed in tempo, and held for a comparatively long time (Stern, 1974, 1977; Werker & McLeod, 1989). For example, consider the "mock surprise" expression with its open mouth, wide eyes and raised eyebrows. Babies find such an expression captivating.

The temporal patterning of the facial movements during mother-infant interaction is highly distinctive (Stern, 1974). At times, mothers actions are fast and exaggerated; at other times, they seem to "be playing with" the speed as the pace changes unpredictably during an interaction. By varying the temporal properties of the movement, mothers maintain the infants interest. That is, the mothers facial movements help to prevent the baby from habituating to her face. In addition, two-month-old infants are especially attentive to faces that are imitative (Field, 1977), suggesting that the temporal parameters of such movements may be optimally suited to the infants' processing skills. By three months of age, infants have developed expectations about facial movements experienced during social interactions. When a face suddenly becomes still or non-responsive, infants show gaze aversion, decreased smiling, and increased distress (Trevarthen, 1983; Tronick, Als, Adamson, Wise, & Brazelton, 1978).

Although newborns may be less tuned to specific social properties of moving faces, it is clear that moving faces captivate very young infants. The fact that infants only a few days old will imitate facial movements (Meltzoff & Moore, 1977; 1983) underscores their attention to and processing of moving faces. Preference studies also suggest a strong early interest in moving faces. For example, 1-month-olds prefer animated to static faces (Sherrod, 1979), and 2-month-olds will attempt to communicate with animated, but not static faces (Trevarthen, 1977).

Different Roles for Different Movements?

Developmental studies provide a much firmer link than adult studies do between motion types and the perceptual and attentive factors that mediate the acquisition of information from a moving face. Namely, rigid motions and non-rigid motions make decidedly different contributions to infant face perception. Rigid motion helps infants find and track faces and aids structure-from-motion analyses, whereas non-rigid motion seems geared more toward helping infants participate in social interactions.

At the level of basic perceptual ability, it seems that infants can distinguish between rigid and non-rigid motion. For example, in a study using moving objects, Gibson, Owsley, and Johnson (1978) found that infants as young as 3 months of age treat rigid and non-rigid movements as two distinct categories of motion. As we shall see, several studies have shown that infants respond both to rigid motion and to the three kinds of non-rigid motion we discussed in the context of the adult literatureexpressions, facial speech and gaze.

Rigid motions.

There is evidence that infants between 3 and 6 months of age process rigid motion. This type of motion seems especially effective for focusing the babys attention on the parents face. For example, infants will follow adult head turns with their eyes when adults redirect their attention from the infant to a puppet (dEntremont, Hains, & Muir, 1997). Also, rigid facial motion may be especially important for establishing and finetuning face-tracking skills. For example, from birth to approximately 6 weeks of age, infants track moving faces longer than other equally complex stimuli (Goren, Sarty, & Wu, 1975; Johnson, Dziurawiec, Ellis, & Morton, 1991). Moreover, it seems that certain properties of the face may be important for engaging the tracking system at this early age, although how specifically tuned these processes are to faces remains unclear. According to Johnson and Morton (Morton & Johnson, 1991; Johnson & Morton, 1991), CONSPEC, a process that depends heavily on subcortical structures, tracks face-like stimuli that appear in the periphery. Whether or not CONSPEC continues to mediate performance later than 6 weeks of age is less certain. After 6 weeks, schematic faces are preferred when they are presented centrally, but schematic faces moving in the periphery no longer attract special attention.

Non-rigid motion.

The sub-types of non-rigid facial motionsgaze, facial speech and expressionare difficult to separate in the infant's world. Indeed these movements are continuously combined and “fed” to the infant during most social interactions. Researchers have nonetheless successfully demonstrated infants sensitivity to each of these three subtypes.

Gaze. Gaze changes form an important component of face-to-face interactions. When mothers avert their gaze, thereby breaking eye contact with their infants, 3- to 7-montholds pay less attention to their mothers face and produce fewer smiles (Hains & Muir, 1996). By 5 months of age, infants can detect shifts in eye gaze of a mere 5 degrees horizontally (Symons, Hains, & Muir, 1998). Detection of gaze shifting becomes increasingly differentiated and becomes useful for infants in a number of different contexts from 5 to 12 months of age (see Rochat, 1999).

Facial Speech. Intermodal matching studies suggest that infants are sensitive to the non-rigid speech production movements of a face. For example, 5-month-olds can perceive that sounds such as /a/ and /i/ are produced by different mouth movements and will preferentially attend to phoneme-synchronized displays (Kuhl & Meltzoff, 1982; Patterson & Werker, 1999). More generally, 4-month-old infants prefer facial speech that is synchronized with a sound track (Dodd, 1979; Spelke & Cortelyou, 1981).

Expression. By 7 months, infants can match dynamic facial emotions (e.g., happy, angry, sad, neutral) with a soundtrack of the corresponding vocal emotion when presented faces and voices of unfamiliar females (for a review, see Walker-Andrews, 1997). To do this, infants need to attend to moving configurations of internal features. We will revisit this issue of internal face features shortly.

Identity-specific Movements. A particularly remarkable demonstration of infants sensitivity to non-rigid motion is the finding that infants can recognize the movement patterns that characterize a face, even in the absence of pictorial cues. Using stimuli analogous to point-light displays, Stucki, Kaufmann-Hayoz, and Kaufmann (1987) showed that 3-month-olds can discriminate between the motion patterns that specify a human face and those created by deforming a rubber mask. Infants viewed either a pointlight display of a womans face as she pretended to interact with a baby or a point-light display of a hand-animated rubber mask. Stucki et al. found that 3-month-olds could discriminate the upright but not the inverted stimuli, indicating that pure motion information is sufficient to specify the familiar animated structure of a face.

Infant Face Recognition.

If both rigid and non-rigid motions can effectively modulate an infants attention to a face, then how does this heightened attention affect infants face recognition ability? We should point out that by "infant face recognition" we mean the ability to differentiate among individual faces. Young infants do not require, nor do they exhibit, the kind of high capacity within-class discrimination abilities that characterize adult face recognition. However, infants perform remarkably well at the subset of face recognition skills that are most important for their survival (Sherrod, 1990). Specifically, infants can discriminate their mothers face from the face of a stranger and can discriminate between a relatively small number of family members and other commonly encountered caretakers.

Despite numerous studies on the perception and recognition of static faces by infants and children (for reviews see Carey, 1996; de Haan, 2001; Nelson & Ludemann, 1989; Sherrod, 1990; Slater & Pascalis, 2001), only a handful of studies have used moving faces to explore the role that motion has in infants ability to discriminate faces. For example, infants as young as 6 weeks of age can learn to identify strangers based on the facial imitation games they play (Meltzoff & Moore, 1992). Consistent with these findings, Spelke and Cortelyou (1981, p. 63) briefly mention findings from their lab suggesting that young infants can abstract identity from moving faces: "Infants learn rapidy to recognize the face of an unfamiliar person by 4 months of age, if the person moves expressively (Spelke, 1975)." Thus, even after relatively brief encounters with a moving face, infants can ably encode motion-based cues to facilitate recognition.

Recognition of external vs. internal facial features.

The fact that non-rigid motions inherently involve the internal features of the face (e.g., eyes, nose, mouth) brings us to an important point regarding infant face recognition. That is, non-rigid movements seem especially useful for drawing infants attention away from the external features of a face and toward the internal features. As we will argue shortly, there is evidence that this shift in attention to internal face features may ultimately benefit infants ability to recognize faces.

In the first month of life, when infants are presented with a novel face, they spend more time scanning the external regions than the internal contours (Haith, Bergman, & Moore, 1977; Maurer & Salapatek, 1976). This scanning bias is seen even when the face moves. However, it should be noted that 1-month-olds are not efficient information seekers. Thus, even when a face is the only object in sight, a considerable proportion of 1-montholds fixations are off the outer contour of the face (i.e., chin and hairline). But, by two months of age, infants tend to fixate the internal regions of the face, especially the eyes, and facial movement affects fixation patterns. For example, Haith et al. compared fixations across still faces, talking faces, and faces that moved from side to side (rigid motion). Haith and colleagues found that 2- and 3-month-olds spent more time scanning the eyes of the face when it was talking than they did when it was still or displayed rigid movements. A more concentrated focus on the eyes could also lead to heightened interest in the internal regions of the face. Consistent with these results, Sherrod (1980) posits that the animated social meaning of faces "pushes" babies to switch from encoding strategies that are largely based on external contours to new strategies that encompass a more global scanning style.

Concomitant with these early scanning strategies, early face recognition is biased toward external features followed by a shift toward processing of internal features at about 2 months of age. Three-day-old infants look longer at their mothers face than the face of an unfamiliar female (Bushnell, Sai, & Mullin, 1989; Field, Cohen, Garcia, & Greenberg, 1984; Pascalis, de Schonen, Morton, Deruelle, & Fabre-Grenet, 1995; Walton, Bower, & Bower, 1992). However, when the hair/face separation line and the outer contour of her head are masked by a scarf, 3-day-old infants no longer show a preference for their mother's face (Pascalis et al., 1995). Three-week-olds also fail to show a preference for their mother when she is wearing a scarf although 5- and 7-weekold infants do show a preference under these conditions (Bartrip, Morton, & de Schonen, 2001). Other evidence for a shift in attention from external to internal facial features was provided by demonstrations that 19-week-olds recognize their mothers face when mothers and strangers were wearing identical wigs, but 5- and 12-week-olds fail to discriminate the faces (Bushnell, 1982).

Although infants can distinguish their mothers internal facial features by 4 months of age, it may be worth noting that a recognition bias for outer features persists through childhood (Campbell, Walker, & Baron-Cohen, 1995) and is even seen in adults when unfamiliar faces are processed. Indeed in a classic study with adults, Ellis, Shepherd, and Davies (1979) showed that unfamiliar face recognition relies more on external features, whereas familiar face recognition relies more on internal features.

Combined, the data on processing of internal versus external features leads us to propose the heightened internal features hypothesis. This hypothesis posits that non-rigid facial motions help infants (and possibly adults as well) shift their focus of attention from the external to the internal features of the face. As we described in the previous section, infants eye scanning patterns suggest an interesting shift in attention from external to internal features as the child develops. This heightened internal features hypothesis is grounded in the notion that non-rigid motions may benefit recognition through an attention-based mechanism. Because the attention-getting signals conveyed by a nonrigidly moving face are essentially social in nature, this hypothesis relates strongly to the motion as social signal hypothesis. The extent to which a particular type of movement enhances face recognition may be a function of its attentional consequences, and more specifically, the extent to which it engages the social interactional system. In other words, we might expect socially engaging facial motions, such as emotional expressions, to draw the interest of the baby (or adult) to the internal features of the face. This in turn, may benefit recognition.

Recognition in the context of the social interaction.

A primary characteristic of face-to-face interactions with infants is that they are interactive. Non-rigid motions of the face help sustain the infants attention during faceto- face interactions and thereby enhance the likelihood that the infant will get a clear view of the face. When mothers are expressive, infants are less likely to turn away (Fogel, 1977). Parents talk slowly and use their voice and subtle facial movements to maintain the baby's interest. These actions are frequently tuned to the infant's developmental level and fluctuating attentional state (Field, 1981).

Infants quickly become more competent social partners. Two- and three-month-old infants and their mothers engage in "protoconversations," which are face-to-face interactions characterized by patterned facial, vocal, and gestural expressions (Bateson, 1979; Brazelton, Koslowski, & Main, 1974; Trevarthen, 1977). During face-to-face interactions, parents pause between their own facial gestures and words as if to allow their preverbal infant to have a turn in the conversation. The parents actions resume after this pause, sometimes with a hiccup or gurgle counting as the infants turn in the conversation. Hence, infants view faces that move and pause and then begin moving again. Moreover, the pauses that occur during these interactions afford clear static views of the face at a time when the infant is actively attending to the face. These clear static views may, in turn, facilitate the encoding of the face.

Infants are not passive perceivers, but rather, through their expressions and changes in gaze, play a part in controlling the cadence of the interaction. With a smile, the baby communicates to the parent that arousal is at an optimal level (Sroufe & Waters, 1976). Gaze is also used to signal continued interest as well as a readiness to receive or produce social information. As Sroufe (1995) points out, during the first year of life, infants become increasingly adept at using gaze to modulate arousal. Infants make and break eye contact in order to control the pace of an interaction and to regulate the emotional intensity of the exchange (e.g., Brazelton, Koslowski, & Main, 1974; Stern, 1974; Field, 1977; Rochat & Striano, 1999). When the infant looks away, the parent maintains his or her gaze (Fogel, 1977). Parents seemingly assume the role of an attentive listener whose face is available when the infant chooses to re-establish visual contact. The changing tempo of the caregivers facial movements serves to maintain the

infants interest in the caregivers face. As the babys social repertoire widens, parent-infant interactions become increasingly playful, and exaggerated facial expressions, accompanied by words or other vocalizations, are often repeated in the babys presence. Importantly, these facial movements may help the baby learn the identity of the face. As Stern (1974) noted nearly three decades ago,

> The slowing of the tempo, along with the full extent of the exaggeration, may enable the infant to maintain the identity of the mother's face across its various physical transformations and thus facilitate the acquisition of a stable face schema. "Normal" adult facial expressions flash very rapidly and conceivably could present the infant with a discontinuous sequence of faces. (p.192)

Moreover, Nelson and Horowitz (1983)posit a reciprocal relationship between the pauses and motions that characterize face-to-face interactions with an infant. They presented infants with holographic stereograms of a woman who either winked and smiled or blew a kiss. Infants viewed either a static or dynamic version of the display. Nelson and Horowitz found that infants who were habituated to a static image looked longer when the stimulus moved, regardless of whether the woman posed the old or new expression. Infants who were habituated to a dynamic display showed a novelty preference for the new expression during the test phase. Specifically, the infants dishabituated to a static image of the woman posing with a new expression, but failed to dishabituate to a static image of the old expression. In light of these findings, Nelson and Horowitz suggest that subtracting motion facilitates face recognition whereas reanimating a face serves to recapture the infants attention.

The data on the interactive nature of the face-to-face interactions leads us to propose the enhanced views hypothesis. In brief, expressive face-to-face interactions consist of pauses intermixed with motion. According to the enhanced views hypothesis, movement helps maintain the viewers interest in the face thereby increasing the likelihood that attention will be focused on the face during these conversational pauses. Hence, attention to non-rigid motions of the internal facial features could provide infants (and adults) with more high quality static views of the face than would be experienced from a face that does not similarly capture the attention of a viewer. The enhanced views hypothesis suggests two potentially overlapping interpretations. First, by prolonging interest in the face, socially engaging movements could result in the viewer encoding more distinct views of the face. Second, heightened attention to the face could lead to enhanced processing and a better representation of each view. These "prime views" may be optimal for building a representation of the static features of a face.

The results of Nelson and Horowitz (1983) and others are relevant for adult studies because they posit a role for the quantity and duration of static pauses afforded the viewer in natural, non-rigidly moving facial stimuli. Interestingly, the duration of these pauses may vary systematically with different types of facial motions. This could, in part, account for the conflicting set of results found in the adult face recognition literature, in which the type of facial motion varies considerably across experiments. An additional relevant factor to consider in the context of face learning and recognition may be the number of pauses a given social encounter affords the viewer. More pauses may help the viewer extract static-based information that is useful for recognition.

Revisiting the Supplemental Information Hypothesis. Although the supplemental information hypothesis may be less relevant for infants than the other two hypotheses, data from the infant imitation literature provide a tenuous, but interesting, connection. Meltzoff and Moore (1999) have suggested the intriguing possibility that facial gestures produced in the context of imitation games may play a role in how infants come to recognize the identity of other people. For example, Meltzoff and Moore (1994) had 6- week-olds play an imitation game with an adult partner. The next day, when infants reencountered their adult partner posing with a neutral expression, the infants looked intently at the adults while making the gesture from the game they had played on the previous day. Meltzoff and Moore suggested that the infant was using this gesture to probe the persons identity as if to ask, “Aren’t you the one who does this?”

Interestingly, in an earlier study, infants often did not realize that the partner in the game had switched. In Meltzoff and Moores (1992) design, mothers first made a particular gesture to her 6-week-old and, after she left, a stranger made a different gesture. Meltzoff and Moore found that infants who visually tracked the individuals coming and going were able to learn and repeat the different gestures associated with each individual. However, the infants who did not visually track the individuals merely repeated the same actions they performed with the person from the previous exchange. Meltzoff and Moore (1999) view this error as evidence that very young infants do not use featural information, but rather spatiotemporal information (i.e., places and trajectories), to determine identity. When infants, whose spatial resolution is quite poor, lose track of a targeted individual, they tend to use facial movements, interactional styles, and other movement-based information, not the face per se, to clarify identity. Thus, the infants ability to use motion as a supplemental source of information for establishing identity may provide a foundation later in life for the use of person-specific facial movement patterns (i.e., dynamic facial signatures) as a recognition cue.

In summary, results from the developmental literature highlight the importance of facial motion, both socially and perceptually. These studies make it clear that processing facial motion is a high priority for infants. Facial motion not only helps capture the infants attention, but ultimately seems to facilitate the perception and encoding of the facial structure. This in turn, helps the infant discriminate between faces. As we will argue in the next section, data from the neural literature on face processing are suggestive of an organization that is divided along the lines of socially-relevant motion information and identity-relevant feature information.

Integrating the Psychological Data into A Neural Systems Model

In the past two decades, enormous progress has been made in mapping the brain regions involved in processing faces. Relevant data have come from the full complement of methods available for associating brain tissue and visual/cognitive function. These include functional neuroimaging, electrophysiology, evoked responses studies, and neuropsychological case studies. A remarkable aspect of these findings is the number and diversity of cortical and sub-cortical brain areas that seem to respond to faces. This diversity has made the interpretation of the complete neural system for face processing challenging.

In this final section, we present the data on memory for moving faces in the context of a particular neural systems model of face processing (Haxby et al., 2000; Haxby, Hoffman, & Gobbini, 2002). We concentrate on this model, because it distinguishes explicitly between the neural processing of moving and static faces. This model is aimed at understanding the anatomical and functional differences that may underlie face processing in these two cases. What follows is by no means an exhaustive look at the full range of data on the neural systems responsive to faces. Rather, it is a selective look at how the presently available data on memory for moving faces might fit into the simplified sub-network proposed by Haxby et al. We think this model provides a useful neural framework for understanding the challenges humans face in recognizing moving faces.

We begin with a brief overview of Haxby et al.s (2000) model and then discuss the psychological hypotheses about the role of face motion in memory in the context of this model. We will suggest that there are plausible neural analogues for the various psychological mechanisms we proposed for motion information to affect memory for faces. Understanding these possible connections might provide guidance to psychologists in designing future experiments.

Haxby et al. (2000) integrated findings from across the diverse lines of research on the neural basis of face processing and proposed a "distributed" neural system for human face perception[2]. Of particular interest for the present review are the anatomical and functional distinctions made in this model between the processing of the invariant and changeable aspects of a face. Haxby et al. propose that the analysis of the invariant aspects of faces supports the function of face identification. This is not surprising, in that we would expect the non-changeable aspects of faces to be the most useful for identifying people over time. By contrast, the analysis of the changeable aspects of faces, (e.g., gaze, facial expression, and speech production movements), are proposed to serve a social communication function. This makes sense, also, when we consider the taxonomy of facial motions presented previously and the argument that nearly all facial motions can be interpreted as social signals.

Anatomically, Haxby et al. (2000) propose an analogous split in the brain areas responsible for processing the invariant and changeable features of faces. The model includes three core brain areas as well as four "extender" areas, which subserve facerelated tasks. We concentrate on two of the three core areasthe lateral fusiform gyrus and the posterior superior temporal sulcusthose posited to be most relevant for processing the invariant and changeable aspects of faces.

The first core area is the lateral fusiform gyrus, reported to be active in many neuroimaging studies of face perception. This area is commonly known as the "fusiform face area" (FFA). Haxby et al. (2000) propose that the FFA responds to the invariant or non-changing information in faces that is most useful for specifying identity. In primate studies, the homologous area is referred to generally as the inferotemporal cortex (IT). Though bilateral activation of this region is frequently found in human neuroimaging studies, the most consistent findings indicate activations that are lateralized in the right hemisphere.

The second core region is the posterior superior temporal sulcus (STS) which is the proposed site for processing the changeable aspects of faces. Studies of single unit physiology in nonhuman primates and neuroimaging studies in humans indicate that this area is important for detecting gaze information, head orientation, and facial expression (Narumoto, Okda, Sadato, Fukui, & Yonekura, 2001; see Allison, Puce, & McCarthy, 2000 for a review). More generally, Haxby et al. (2000) note that many kinds of biological motion, including motion of the whole body, the hand, and the eyes and mouth activate the STS and that the STS is the primary brain region tied to the processing of facial motion.

The third core area in Haxby et al.s model, the inferior occipital gyri, is posited to be involved in the early perception of facial features. This region is proposed as a "feeder" system to both the lateral fusiform and superior temporal sulcal regions.

In addition to the core regions of the neural systems model, Haxby et al. (2000) propose four brain regions to carry out auxiliary tasks on faces. These "extender" areas channel information from faces to other brain areas implicated in face-related tasks. We describe each of these in turn. As we will see, the functions associated with regions that extend the motion-based STS core system, are tied closely to processing the three types of non-rigid motions we defined previously, (i.e., gaze, speech production movements, and facial expression) and to rigid motions (e.g., head orientation). In all cases, the extender regions are concerned with applying the information extracted from faces to concrete tasks, such as using gaze information to redirect ones own attention, using lip movements to

[2] For readers inclined to interpret "distributed" in terms of computational or artificial neural networks, this term is a bit misleading. By distributed, Haxby et al. (2000) mean that the brain areas responsive to faces are not localized to a single area of the brain.

help in speech comprehension, and using facial expression in the many ways required for adaptive social interaction.

The first auxiliary region is the intraparietal sulcus, which Haxby et al. (2000) propose as an extender region for the STS system to process spatially directed attention from information about eye gaze and head direction. As noted previously, an important social function of facial motion is to provide information about the object(s) of another persons attention. For example, eye gaze changes are especially effective for helping us to gauge the focus of another persons attention. Spatial attention cues are provided also by large head movements, such as head tilting. Such motions fall into the category of motions presumably processed by the intraparietal sulcus.

The second auxiliary region is the auditory cortex, which Haxby et al. (2000) propose as an extender region for STS system to aid in prelexical speech perception. The third extender system consists of limbic structures, including the amygdala and insula, which are activated by facial expressions.

Finally, Haxby et al. (2000) propose the anterior temporal area as an extender region for the fusiform face area. This extender region is thought to be involved in the retrieval of personal identity, name, and biographical information and may have access to the identity details of the person whose face we are viewing (Haxby et al., 2000).

The Neural Systems Model and Visual Streaming

The functional division of these brain areas for processing the changeable versus invariant information about faces in the Haxby et al. (2000) model is consistent with what is known about the early neural processing of visual information. Beginning at the retina, visual information is channeled into the: a) high resolution, color sensitive, *parvocellular* stream, and the b) lower resolution, motion sensitive, *magnocellular* stream (Merigan, 1991). These visual inputs map onto the "what" (ventral) and "where" (dorsal) visual streams, whose putative functions are to guide object recognition and spatial orientation, respectively (Ungerlieder & Mishkin, 1982). The ventral system projects to inferotemporal areas including the fusiform gyrus. The dorsal stream, on the other hand, projects to motion-processing areas including medial temporal cortex (MT), STS, and the parietal lobe.

In summary, Haxby et al. (2000) propose that socially relevant facial motions are processed in the STS and identity-relevant feature information is processed in the fusiform gyrus. The STS is extended further to regions that make use of the motion information for specific face processing tasks that are related to social communication. The fusiform system is extended further to retrieve the identity information that may be linked to the processing of the individual as a person. This two-route system for faces is a natural extension, both theoretically and functionally, of the parallel processing hypothesis for visual perception. The IT/FFA system is responsible for invariant, high-resolution facial feature analysis, from parvocellular, ventral stream input. The STS region is responsible for the analysis of facial motions from magnocellular, dorsal stream input.

Recognition of moving faces and the interaction of two systems

If the processing of feature-based facial information is dominant in the ventral stream, and the processing of face motion is dominant in the dorsal stream, how do the two systems interact when we recognize a moving face? In other words, how does the brain integrate the dynamic, static, and social information cues that we encounter when viewing a face in motion? These questions return us to the hypotheses we proposed originally about how motion might affect memory for faces. We will propose two speculative modifications to the model of Haxby et al. (2000) that might offer a clearer view of the various processing routes for facial information when a face is in motion (See figure8). We discuss each hypothesis in turn.

We have argued that a primary way in which motion could benefit recognition is through the use of supplemental information. This presupposes a direct encoding of the motion-based, identity-specific attributes of a face, such that the charac-

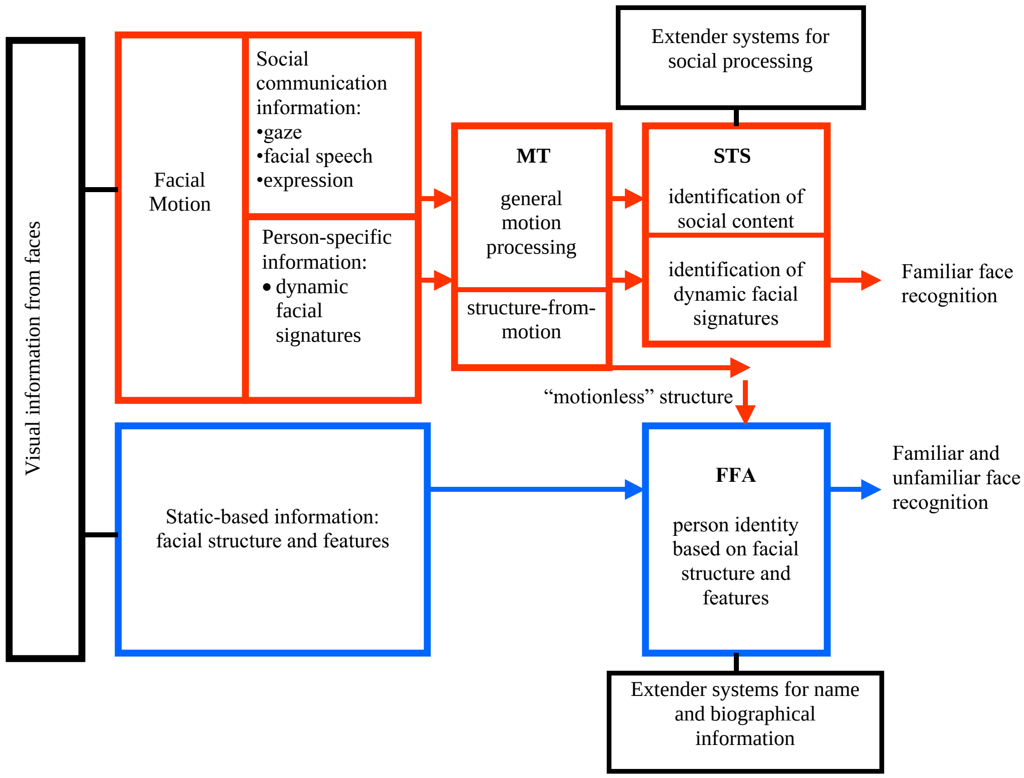

Figure 8. The psychological findings regarding the effects of facial motion on recognition can be mapped onto separate sets of brain structures, based on Haxby et al.s (2000) distributed neural system for face perception. The ventrally-based stream (dashed line) processes the static features and structure of a face, and the dorsally-based stream (solid line) processes facial motion. Facial motion contains two different types of information: social communication signals (gaze, expression and facial speech) and person-specific dynamic facial signatures. Both sets of information are forwarded first to MT for general motion processing and then to the STS, where two types of output are possible. The social communication information is forwarded from STS to extender systems responsible for additional social processing. STS processing of dynamic facial signatures provides identification-based output, beneficial for familiar face recognition (supplemental information hypothesis). A second route from MT to IT (defined more specifically in humans as FFA) may underlie structure-from-motion processing. "Motionless" structure information becomes accessible to IT, enriching static-based representations of faces in the ventral stream. This added input can contribute to recognition of unfamiliar faces (representation enhancement hypothesis), but may also help recognition for familiar faces. (*Figure from O'Toole, Roak, & Abdi; 2002.*)

teristic facial motions of an individual are recognized in their dynamic form. Following this line of thought, the first modification we propose to the neural systems model is a conceptual one. Namely, "identity-specific information," in the form of dynamic facial signatures, is processed by the STS system. Dynamic facial signatures are entwined and embedded in facial expressions, facial speech, and orienting head/face movements. We speculate that these identity-laden facial motions are processed in the dorsal stream, transiting through the general visual areas that support motion processing (e.g., MT) to the STS.

Thus, in addition to the primary face recognition system that processes static information along the ventral stream to IT, we hypothesize that the

STS system may act as a secondary system for face recognition based on solely on the supplemental information provided by facial motion. However, some caveats apply. First, because this system is based on the recognition of characteristic gestures and facial motions, it is likely to be useful only for people we know reasonably well (i.e., familiar faces). Second, given that under most circumstances, static feature-based information is more reliable for identification than dynamic facial signatures, this system may contribute more to recognition when the viewing conditions are non-optimal. Aspects of dorsal-stream processing, like its ability to operate in poor illumination and with low-resolution stimuli, make it ideal as a plausible secondary route to recognition for stimuli that are not "ventral system quality." It is logical to consider, therefore, that dorsal stream processes might support the kind of robust face recognition we exhibit for people we know well, even when viewing conditions are at their poorest.

To summarize, the first modification of Haxby et al.s (2000) neural systems model amends the role of STS stream to include *face identification*, via processing dynamic identity signatures along with other kinds of social communication information from moving faces. This modification maps onto the supplemental information hypothesis because it ties recognition benefits to the inherently dynamic cues embedded in facial motion, and puts this analysis in a visual stream primed for motion processing–the dorsal stream. See Figure 8 for a diagram that maps out the proposed processing of dynamic facial signatures.

Can the ventral stream access structure from motion processing?

Our second proposed modification of Haxby et als (2000) model is more speculative and concerns structure-from-motion analyses for face recognition. Recall that the representation enhancement hypothesis suggests that the benefits of motion in face recognition come from using motion to extract a more accurate representation of the static or invariant structure of a face. Thus, motion bootstraps the encoding of face structure, but is not useful itself. Structure-from-motion analysis is a classically dorsal stream process, with evidence for the involvement of MT (Bradley, Chang, & Anderson, 1998). We propose an addition to the neural systems model of Haxby et al. that allows for the output of structure-from-motion analyses in MT to project back to IT–but as *static form* information (i.e., "motionless form").

There is no direct proof for this rather speculative suggestion, though there is some neurophysiological evidence that puts the visual perception basics on firmer ground. In particular, Sary, Vogels, & Orban (1993) demonstrated that neurons in IT, selective for particular forms, continued to respond to the form even when it was specified by pure motion-induced contrasts. This suggests that form information specified by motion is available, at least in principle, to the ventrally based IT system.

In a complementary study, Britten, Newsome, and Saunders (1992) found that although IT-lesioned monkeys were unable to learn shape discriminations based on form-from-luminance cues, they could nonetheless discriminate shapes based on form-from-motion information. Britten et al. suggest the dorsal pathway–including MT, MST, and other parieto-occipital areas–as a "logical candidate" for the site of form-from-motion processing when IT is impaired.

Both Sary et al. (1993) and Britten et al. (1992) suggest known connections from MT to IT via V4 (Maunsell & Van Essen, 1983; Ungerleider & Desimone, 1986) as a plausible basis of their findings. At present, however, there are no direct demonstrations of the usefulness of this route for face recognition. Further studies are needed to flesh out the details of the relationship between IT neurons and motion cues for face recognition.

To summarize, we propose amending Haxby et al.s neural systems model to allow for the possibility that the structure-from-motion analysis of moving faces can bootstrap the encoding of the static face features and that this information is ultimately made available to the IT/fusiform system. This kind of processing could improve the perceptual quality of the feature information we can extract from a moving face, thereby aiding recognition. Finally, because of the inherently perceptual nature of the information posited in the representation enhancement hypothesis, this benefit to recognition should be available to both for both familiar and unfamiliar face recognition. See Figure 8 for a schematic diagram of the processing involved.

Concluding Remarks

Until recently, face recognition research has been dominated by studies using static images of faces as stimuli. Because the faces we learn and recognize in most daily situations are in motion, determining the factors that affect memory for moving faces can provide a foundation for understanding how we recognize people in natural settings. Recent studies have demonstrated that the perceptual and social complexities introduced by facial movements affect face recognition performance in non-trivial ways.

Currently, what we know about the effects of facial motion on memory is based largely on a set of loosely connected empirical findings. The main purpose of this review has been to place the empirical results into a theoretical framework, from which future studies can proceed in a more systematic fashion. The hypotheses we have outlined offer a point of departure for this endeavor. The primary challenge for future studies will be to devise experimental designs that are powerful enough to isolate the contributions of the specific perceptual and social factors that impact memory for moving faces. Experimental efforts in this direction may offer new insights into understanding face identification in naturalistic contexts. Eyewitness identifications and computer-based face recognition systems are prime examples of recognition outcomes that are difficult to predict from data based on static face recognition experiments. The issues involved in recognizing faces as they move, gesture, and communicate are complex, but they are fundamental to our understanding of human face recognition in real world contexts.

References

Adolphs, R. (2002). Recognizing emotion from facial expressions: Psychological and Neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews, 1*, 21–60.

Allison, T., Puce, M.A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences, 4*, 267–278.

Atkinson, J. (1998). The where and what or who and how of visual development. In F. Simion & G. Butterworth (Eds.), *The development of sensory, motor and cognitive capacities in early infancy: From perception to cognition*, (pp. 3 24). Hove, UK: Psychology Press.

Bahrick, H.P., Bahrick, P.O., & Wittlinger, R.P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General, 104*, 54–75.

Banks, M.S. & Crowell, J.A. (1993). Front-end limitations to infant spatial vision: Examination of two analyses. In K. Simons (Ed.), *Early visual development: Normal and abnormal* (pp. 91–116). New York: Oxford University Press.

Bartrip, J., Morton, J. & DeSchonen, S. (2001). Responses to mothers face in 3-week to 5-month-old infants. *British Journal of Developmental Psychology, 19*, 219–232.

Bassilli, J.N. (1978). Facial motion in the perception of faces and emotional expression. *Journal of Experimental Psychology, 4*, 373–379.

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

Bateson, M.C. (1979). The epigenesis of conversational interaction: a personal account of research development. In D. Aaronson & R.W. Reiber (Eds.), *Developmental Psycholinguistics and Communication Disorders; Annals of the New York Academy of Sciences, vol. CCLXIII* (pp. 101–113). New York: New York Academy of Sciences.

Bernstein, L. E., Demorest, M. E. & Tucker, P. E. (2000). Speech perception without hearing. *Perception and Psychophysics, 6*, 233–252.

Berry, D.S. (1990). What can a moving face tell us? *Journal of Personality and Social Psychology, 58*, 1004–1014.

Blanz, V. & Vetter, T. (1999). A morphable model for the synthesis of 3D faces. *SIGGRAPH 99 Proceedings*, Los Angeles, CA.

Bornstein, M.H., Ferdinandsen, K. & Gross, C.G. (1981). Perception of symmetry in infancy. *Developmental Psychology, 17*, 82–86.

Bower, G.H. & Karlin, M.B. (1974). Depth of processing of pictures of faces and recognition memory. *Journal of Experimental Psychology, 103*, 751–757.

Bradley, D.C., Chang, G.C., & Anderson, R.A. (1998). Encoding of 3D structurefrom-motion by primate area MT neurons. *Nature, 392,* 714.

Braje, W. L., Kersten, D., Tarr, M. J. & Troje, N.F. (1999). Illumination effects in face recognition. *Psychobiology, 26,* 371–380.

Brazelton, T. B., Koslowski, B. & Main, M. (1974). The origins of reciprocity: The early mother-infant interaction. In Michael Lewis & Leonard A. Rosenblum (Eds.), *The effect of the infant on its caregiver* (pp. 49–76). New York: Wiley.

Britten K.H., Newsome, W.T., & Saunders, R.C. (1992). Effects of inferotemporal lesions on form-from-motion discrimination. *Experimental Brain Research, 88,* 292–302.

Bronson, G.W. (1990). Changes in infants visual scanning across the 2-24-week age period. *Journal of Experimental Child Psychology, 49,* 101–125.

Bruce, V., Burton, A.M., Carson, D., Hanna, E., & Mason, O. (1994). Repetition priming of face recognition. *Attention and Performance, 15,* 179–201.

Bruce, V. & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77,* 305–327.

Bruce, V. & Valentine, T. (1988). When a nods as good as a wink: The role of dynamic information in facial recognition. *Practical Aspects of Memory: Current. Research and Ideas, 1*, 169–174.

Bruce, V., Henderson, Z., Greenwood, K., Hancock, P.J.B., Burton, A., & Miller, P. (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied, 5,* 339–360.

Bruce, V., Henderson, Z., Newman, C., & Burton, M.A. (2001). Matching identities of familiar and unfamiliar faces caught on CCTV images. *Journal of Experimental Psychology: Applied, 7,* 207–218.

Bruyer, R., Laterre, C., Seron, X., Feyereisen, P., Strypstein, E., Pierrard, E., & Rectem, D. (1983). A case of prosopagnosia with some preserved covert remembrance of similar faces. *Brain and Cognition, 2,* 257–284.

Burnham, D. K. (1987). The role of movement in object perception by infants. In B.E. McKenzie & R.H. Day (Eds.), *Perceptual development in infancy: Problems and issues* (pp. 143–172). Hillsdale, NJ: Erlbaum.

Burton, A.M., Wilson, S., Cowan, M., & Bruce, V. (1999). Face recognition in poor quality video. *Psychological Science, 10,* 243–248.

Burton, A.M., Bruce, V., & Hancock, P.J.B. (1999). From pixels to people: A model of familiar face recognition. *Cognitive Science, 23,* 1–31.

Bushnell, I. W. R. (1982). Discrimination of faces by young infants. *Journal of Experimental Child Psychology, 33,* 298–308.

Bushnell, I.W.R. (1998). The origins of face perception. In F. Simion & G. Butterworth (Eds.), *The development of sensory, motor and cognitive capacities in early infancy: From perception to cognition* (pp. 69–86). Hove, UK: Psychology Press.

Bushnell, I.W.R., Sai, F. & Mullin, J.T. (1989). Neonatal recognition of the mothers face. *British Journal of Developmental Psychology, 7,* 3–15.

Calvert, G.A., Brammer, M.J., & Iversen, S.D. (1998). Crossmodal identification. *Trends in Cognitive Sciences, 2,* 247–253.

Campbell, R., Dodd, B., & Burnham, D. (Eds.). (1998). *Hearing by Eye II: Advances in the Psychology of Speechreading Audiovisual Speech.* Psychology Press.

Campbell, R., Walker, J., & Baron-Cohen, S. (1995). The development of differential use of inner and outer face features in familiar face identification. *Journal of Experimental Child Psychology, 59,* 196–210.

Carey, S. (1996). Perceptual classification and expertise. In R. Gelman & T. Kit-Fong Au (Eds.), *Perceptual and cognitive development* (pp. 49–69). NY: Academic Press.

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

Cohn, J.F. & Tronick, E.Z. (1988). Mother-infant face-to-face interaction: Interaction is bi-directional and unrelated to periodic cycles in either partners behavior. *Developmental Psychology, 24,* 386–392.

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

Damasio, A.R., Damasio, H., & Van Hoesen, G.W. (1982). Prosopagnosia: Anatomic basis and behavioral mechanisms. *Neurology, 32,* 331–341.

Danemiller, J.L. (2001). Brain-behavior relationships in early visual development. In C.A. Nelson & M. Luciana (Eds.), *Handbook of Developmental Cognitive Neuroscience* (pp. 221–235). Cambridge, MA: MIT.

DeHaan, M. (2001). The neuropsychology of face processing during infancy and childhood. In C.A. Nelson & M. Luciana (Eds.), *Handbook of Developmental Cognitive Neuroscience* (pp. 381–398). Cambridge, MA: MIT.

DEntremont, B., Hains, S. M. J., & Muir, D. W. (1997). A demonstration of gaze following in 3- to 6-montholds. *Infant Behavior and Development, 20,* 569–572.

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

Dodd, B. (1979). Lip-reading in infants: Attention to speech presented in- and out-of synchrony. *Cognitive Psychology, 11,* 478–484.

Ellis, H.D., Shephard, J.W., & Davies, G.M. (1979). Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception, 8,* 431–439.

Etcoff, N. (1984). Selective identity to facial identity and facial emotion. *Neuropsychologia, 22,* 281–295.

Fantz, R.L. & Miranda, S.B. (1975). Newborn attention to form of contour. *Child Development, 46,* 224–228.

Farah, M. J., Wilson, K. D., Drain, M. & Tanaka, J. N. (1998). What is "special" about face recognition? *Psychological Review, 105,* 482–498.

Fernald, A. & Simon, T. (1984). Expanded intonation contours in mothers speech to newborns. *Developmental Psychology, 20,* 104–113.

Field, T. (1977). Effects of early separation, interactive deficits, and experimental manipulations on infant-mother face-to-face interaction. *Child Development, 48,* 763–771.

Field, T. (1981). Infant arousal attention and affect during early interactions. In L.P. Lipsitt & C.K. Rovee-Collier (Eds.), *Advances in Infancy Research, Vol. 1* (pp. 57–100). Norwood, NJ: Ablex.

Field, T. M., Cohen, D. Garcia, R. & Greenberg, R. (1984). Mother-stranger face discrimination by the newborn. *Infant Behavior and Development, 7,*19–25.

Fogel, A. (1977). Temporal organization in mother-infant face-to-face interaction. In H.R. Schaffer (Ed.), *Studies in mother-infant interaction.* New York: Academic Press.

Galper, R. E. & Hochberg, J. (1971). Recognition memory for photographs of faces. *American Journal of Psychology, 84,* 351–354.

Gibson, E.J., Owsley, C.J., & Johnston, J. (1978). Perception of invariants by fivemonth-old infants: Differentiation of two types of motion. *Developmental Psychology, 14,* 407–415.

Goren, C.C., Sarty, M. & Wu, P.Y.K. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics, 56,* 544–549.

Granrud, C. E., Yonas, A. & Opland, E. A. (1985). Infants' sensitivity to the depth cue of shading. *Perception & Psychophysics, 37,* 415–419.

Hains, S.M.J. & Muir, D.W. (1996). Infant sensitivity to adult eye direction. *Child Development, 67,* 1940–1951.

Haith, M.M., Bergman, T. & Moore, M.J. (1977). Eye contact and face scanning in early infancy. *Science, 198,* 853–855.

Hancock, P.J.B., Bruce, V., & Burton, A.M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences, 4,* 330–337.

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

Haxby, J.V., Hoffman, E.A., & Gobbini, M.I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4,* 223–233.

Henderson, Z., Bruce, V., & Burton, A.M., (2001). Matching the faces of robbers captured on video. *Applied Cognitive Psychology, 15,* 445–464.

Hill, H. & Bruce V. (1996). Effects of lighting on the perception of facial surface. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 986–1004.

Hill, H. & Pollick, F.E. (2000). Exaggerating temporal differences enhances recognition of individuals from point light displays. *Psychological Science, 3,* 223–228.

Hill, H. & Johnston, A. (2001). Categorizing sex and identity from the biological motion of faces. *Current Biology, 11,* 880–885.

Hood, Bruce, M., Willen, J. D., & Driver, J. (1998). Adults eyes trigger shifts of visual attention in human infants. *Psychological Science, 9,* 131–134.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics, 14,* 201–211.

Johnson, M.H., Dziurawiec, S., Ellis, H. & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition, 40,* 1–19.

Johnson, M.H. & Morton, J. (1991). *Biology and cognitive development: The case of face recognition.* Cambridge, MA: Blackwell.

Kalocsai. P. & Biederman, I. (1997). Difference of face and oject recognition in utilizing early visual information. In H. Wechsler, P.J. Phillipps, V. Bruce, F. Fogelman Souli, & T.S. Huang (Eds), *Face Recognition: From Theory to Applications, Nato ASI series, Series F computer and systems sciences, 163*.

Knappmeyer, B., Thornton, I.M., & Blthoff, H.H. (2001). Facial motion can determine identity. *Journal of Vision, 3,* 337.

Knight, B. & Johnston, A. (1997). The role of movement in face recognition. *Visual Cognition, 4,* 265–273.

Kuhl, P.K. & Meltzoff, A.N. (1982). The bimodal perception of speech in infancy. *Science, 218,* 1138–1141.

Kurucz, J. & Feldmar, G. (1979). Prosopo-affective agnosia as a symptom of cerebral organic brain disease. *Journal of the American Geriatrics Society, 27,* 91–95.

Kurucz, J., Feldmar, G. & Werner, W. (1979). Prosopo-affective agnosia associated with chronic organic brain syndrome. *Journal of American Geriatrics Society, 27,* 225–230.

Lander, K., Christie, F., & Bruce, V. (1999). The role of movement in the recognition of famous faces. *Memory & Cognition, 27,* 974–985.

Lander, K. & Bruce, V. (2000). Recognizing famous faces: exploring the benefits of facial motion. *Ecological Psychology, 12,* 259–272.

Lander, K., Bruce, V., & Hill, H. (2001). Evaluating the effectiveness of pixelation and blurring on masking the identity of familiar faces. *Applied Cognitive Psychology, 15,* 101–116.

Lander K. & Bruce V. (2001). Repetition priming of moving faces. *Talk presented at the European Conference Of Cognitive Psychology,* (ESCOP) X, Edinburgh Scotland, Sept. 2001.

Langton, S.R.H., Watt, R.J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences, 4,* 50–59.

Marr, D. (1982). *Vision.* San Francisco: Freeman.

Maunsell, J.H.R. & Van Essen, D.C. (1983). The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *Journal of Neurophysiology, 3,* 2563–2586.

Maurer, D. & Lewis, T.L. (2001). Visual acuity and spatial contrast sensitivity: Normal development and underlying mechanisms. In C.A. Nelson & M. Luciana (Eds.), *Handbook of Developmental Cognitive Neuroscience* (p. 237–251).Cambridge, MA: MIT.

Maurer, D. & Salapatek, P. (1976). Developmental changes in the scanning of faces by young infants. *Child Development, 47,* 523–527.

McGurk, H. & MacDonald, T. (1976). Hearing lips and seeing voices. *Nature, 264,* 746–748.

Merigan, W. H. (1991). P and M pathway specialization in the macaque. In A. Valberg and B.B. Lee (Eds.). *From Pigments to Perception.* (pp. 117–125). Plenum Press.

Meltzoff, A.N. & Moore, M.K. (1977). Imitation of facial and manual gestures by human neonates. *Science, 198,* 75–78.

Meltzoff, A.N. & Moore, M.K. (1983). Newborn infants imitate adult facial gestures. *Child Development, 54,* 702–709.

Meltzoff, A.N. & Moore, M.K. (1992). Early imitation within a functional framework: The importance of person identity, movement, and development. *Infant Behavior and Development, 15,* 479–505.

Meltzoff, A.N. & Moore, M.K. (1994). Imitation, memory, and the representation of persons. *Infant Behavior and Development, 17,* 83–99.

Meltzoff, A.N. & Moore, M.K. (1998). Infant intersubjectivity: Broadening the dialogue to include imitation, identity, and intention. In S. Braten (Ed.), *Intersubjective Communication and Emotion in Early*

Ontogeny: A Sourcebook (pp. 47–62). Cambridge, UK: Cambridge University Press.

Meltzoff, A.N. & Moore, M.K. (1999). Persons and representation: Why infant imitation is important for theories of human development. In J. Nadel & G. Butterworth (Eds.), *Imitation in Infancy* (pp. 935). New York: Cambridge University Press.

Morton, J. & Johnson, M.H. (1991). CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological Review, 98,* 164–181.

Narumoto, J., Okda, T., Sadato, N., Fukui, K., & Yonekura, Y. (2001). Attention to emotion modulates fMRI activity in human right superior temporal sulcus. *Cognitive Brain Research, 12*, 225–231.

Nelson, C.A. (2001). The development and neural bases of face recognition. *Infant and Child Development, 10,* 3–18.

Nelson, C.A. & Horowitz, F.D. (1983). The perception of facial expressions and stimulus motion by two- and five-month-old infants using holographic stimuli. *Child Development, 54,* 868–877.

Nelson, C.A. & Ludemann, P.M. (1989). Past, current, and future trends in infant face perception research. *Canadian Journal of Psychology, 43,* 183–198.

OToole, A. J., Edelman, S. & Blthoff, H.H. (1998). Stimulus-specific effects in face recognition over changes in viewpoint. *Vision Research, 38,* 2351–2363.

OToole, A.J., Roark, D.A. & Abdi, H. (2002). Recognizing moving faces: A psychological and neural synthesis. *Trends in Cognitive Sciences, 6*, 261–266.

Papousek, M. (1992). Early ontogeny of vocal communication in parent-infant interactions. In H. Papousek, U. Jurgens, & M. Papousek (Eds.), *Nonverbal vocal communication: Comparative and developmental approaches* (pp. 230–261). Cambridge: Cambridge University Press.

Pascalis, O., de Schonen, S., Morton, J., Deruelle, C., & Febre-Grenet, M. (1995). *Mothers face recognition by neonates: A replication and extension. Infant Behavior and Development, 18,* 79–85.

Patterson, M. L. & Werker, J. F. (1999). Matching phonetic information in lips and voice is robust in 4.5-month-old infants. *Infant Behavior & Development, 22,* 237–247.

Pike, G., Kemp, R., Towell, N., & Phillips, K. (1997). Recognizing moving faces: The relative contribution of motion and perspective view information. *Visual Cognition, 4,* 409–437.

Perrett, D. I., Harries, M.H., Mistlin, A.J., Hietanen, J.K., Benson P.J., Bevan, R., Thomas, S., Oram, M.W., Ortega, J., and Brierley, K. (1990). Social signals analyzed at the cell level: Someone is looking at me, something touched me, something moved. *International Journal of Comparative Psychology, 4,* 25–54.

Rochat, P. (Ed.) (1999). *Early Social Cognition: Understanding others in the first months of life.* Mahwah, NJ: Erlbaum.

Rochat, P. & Striano, T. (1999). Social-cognitive development in the first year. In P. Rochat (Ed.), *Early Social Cognition.* Mahwah, NJ: Erlbaum. (p 3–34).

Rosenblum, L.D., Johnson, J. A., & Saldaa, H.M. (1996). Point-light facial displays enhance comprehension of speech in noise. *Journal of Speech and Hearing Research, 39,* 1159–1170.

Rosenblum, L. D., & Saldaa, H.M. (1996). An audiovisual test of kinematic primitives for visual speech perception. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 318–331.

Sary G., Vogels, R., & Orban, G.A. (1993). Cue invariant shape selectivity of macaque inferior temporal neurons. *Science, 260,* 995–997.

Schiff, W., Banka, L., & de Bordes-Galdi, G. (1986). Recognizing people seen in events via dynamic "mug shots." *American Journal of Psychology, 99,* 219–231.

Schweinberger, S.R., Burton, A.M., & Kelly, S.W. (1999). Asymmetric dependencies in perceiving identity and emotion: Experiments with morphed faces. *Perception and Psychophysics, 61,* 1102–1115.

Schweinberger, R.S. & Soukoup, G.R. (1998). Asymmetric relationships among perceptions of facial identity, emotion, and facial speech. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1748–1765.

Shepherd, J.W., Ellis, H.D., & Davies, G.M., (1982). *Identification Evidence: A Psychological Evaluation.* University of Aberdeen Press.

Sherrod, L.R. (1979). Social cognition in infants: Attention to the human face. *Infant Behavior and Development, 2,* 279–294.

Sherrod, L. R. (1980). Issues in cognitive-perceptual development: The special case of social stimuli. In Michael Lamb & Lonnie R. Sherrod (Eds.), *Infant social cognition* (pp. 11– 36). Hillsdale, NJ: Erlbaum.

Sherrod, L.R. (1990). How do babies know their friends and foes? *Human Nature, 1,* 331–353.

Shuttleworth, E.C., Syring, V., & Allen, N. (1982). Further Observations on the Nature of Prosopagnosia. *Brain and Cognition, 1,* 307–322.

Simons, D.J. & Levin, D. T. (1998). Failure to detect changes to people during a realworld interaction. *Psychonomic Bulletin Review, 5,* 644–649.

Slater, A., Morison, V., Town, C. & Rose, D. (1985). Movement perception and identity constancy in the new-born baby. *British Journal of Developmental Psychology, 3,* 211–220.

Slater, A. & Pascalis, O. (Eds). (2001) *Special Issue: Infant & Child Development, 10.*

Spelke, E.S. & Cortelyou, A. (1981). Perceptual aspects of social knowing: Looking and listening in infancy. In Michael Lamb & Lonnie R. Sherrod (Eds.), *Infant social cognition* (pp. 61–84). Hillsdale, NJ: Erlbaum.

Sroufe, L.A. (1995). *Emotional development*. New York: Cambridge University Press.

Sroufe, L.A. & Waters, E. (1976). The ontogenesis of smiling and laughter: A perspective on the organization of development in infancy. *Psychological Review, 83,* 173–189.

Stern, D. N. (1974). Mother and infant at play: The dyadic interaction involving facial, vocal, and gaze behaviors. In Michael Lewis & Leonard A. Rosenblum (Eds.), *The effect of the infant on its caregiver* (pp. 187–213). New York: Wiley.

Stern, D.N. (1977). *The first relationship: Mother and infant*. Cambridge, MA: Harvard University Press.

Stevenage, S., Nixon, M.S., & Vince, K. (1999). Visual analysis of gait as a cue to identity. *Applied Cognitive Psychology, 13,* 513–526.

Stucki, M., Kaufmann-Hayoz, R., & Kaufmann, F. (1987). Infants' recognition of a face revealed through motion: Contribution of internal facial movement and head movement. *Journal of Experimental Child Psychology, 44,* 80–91.

Symons, L A., Hains, S. M. J., & Muir, D. W. (1998). Look at me: Five-month-old infants sensitivity to very small deviations in eye-gaze during social interactions. *Infant Behavior and Development, 21,* 531–536.

Thornton, I.M. & Kourtzi, Z. (2002). A matching advantage for dynamic human faces. *Perception, 31,* 113–132.

Trevarthen, C. (1983). Interpersonal abilities of infants as generates for transmission of language and culture. In A. Oliverio & M. Zappella (Eds.), *The behavior of human infants* (pp. 244–274). New York, NY: Plenum Press.

Trevarthen, C. (1977). Descriptive analyses of infant communicative behavior. In H.R. Schaffer (Ed.), *Studies in mother-infant interaction* (pp. 227–270). London: Academic Press.

Tronick, E.Z. (1989). Emotions and emotional communication in infants. *American Psychologist, 44,* 112–119.

Tronick, E. Z., Als, A., Adamson, L., Wise, S., & Brazelton, T.B. (1978). The infants response to entrapment between contradictory messages in face-to-face interaction. *Journal of American Academy of Child Psychiatry, 17,* 1–13.

Troje, N. F. & Blthoff, H. H. (1996). Face recognition under varying pose: The role of texture and shape. *Vision Research, 36,* 1761–1771.

Ullman, S. (1979). The Interpretation of Visual Motion. Cambride, MA : MIT Press.

Ungerlieder, L.G. & Mishkin, M. (1982). Two cortical visual systems. In Ingle, D.J. et al., (Eds.). *Analysis of Visual Behavior.* (pp. 549–586). MIT Press.

Ungerleider, L.G. & Desimone, R. (1986). Cortical connections of visual area MT in the macaque. *Journal of Comparitive Neurology, 248,* 190–222.

Valentin, D., Abdi, H., Edelman, B., Nijdam, A. (1996). Connectionism "face"-off: Different algorithms for different tasks. *Psychologica Belgica, 36,* 65–92.

Valentine, T. (1988). Upside-down faces : A review of the effect of inversion upon face recognition. *British Journal of Psychology, 79,* 471–491.

Vecera, S. P. & Johnson, M. (1995). Gaze detection and the cortical processing of faces: Evidence from infants and adults. *Visual Cognition, 2,* 59–87.

Walker-Andrews, A.S. (1997). Infants perception of expressive behaviors: Differentiation of multimodal information. *Psychological Bulletin, 121,* 437–456.

Walton, G.E., Bower, N.J., Bower, T.G.R. (1992). Recognition of familiar faces by newborns. *Infant Behavior and Development, 15,* 265–269.

Wallach, H. & OConnell, D.N. (1953). The kinetic depth effect. *Journal of Experimental Psychology, 45,* 205–217.

Werker, J. F. & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology, 43,* 230–246.

Yin, R.K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81,* 141–145.

Yonas, A. & Arterberry, M.E. (1994). Infants perceive spatial structure specified by line junctions. *Perception, 23,* 1427–1435.