PERGAMON

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Vision Research 43 (2003) 1921–1936

Vision Research

www.elsevier.com/locate/visres

The use of facial motion and facial form during the processing of identity

Barbara Knappmeyer *, Ian M. Thornton, Heinrich H. Bülthoff

Max Planck Institute for Biological Cybernetics, Spemannstr. 38, Tübingen 72076, Germany

Received 2 December 2002; received in revised form 15 March 2003

Abstract

Previous research has shown that facial motion can carry information about age, gender, emotion and, at least to some extent, identity. By combining recent computer animation techniques with psychophysical methods, we show that during the computation of identity the human face recognition system integrates both types of information: individual non-rigid facial motion and individual facial form. This has important implications for cognitive and neural models of face perception, which currently emphasize a separation between the processing of invariant aspects (facial form) and changeable aspects (facial motion) of faces.

Keywords: Face recognition; Facial form; Facial motion; Cue integration; Computer animation

1. Introduction

Traditionally, researchers have used static stimuli, such as line drawings (e.g., Davies, Ellis, & Shepherd, 1978; Tanaka & Farah, 1993), photographs (e.g., Hancock, Bruce, & Burton, 2000; Perrett et al., 1998) or laserscans of human heads (e.g., Leopold, O'Toole, Vetter, & Blanz, 2001; Troje & Bülthoff, 1996, 1998) to explore the representation and processing of faces. However, human faces are *dynamic* rather than *static* objects. As we talk, as we raise our eyebrows, as we laugh, or as we nod our heads to signal agreement, our faces move and change in subtle though significant ways, varying along both, spatial and temporal dimensions. Although artists and impersonators have long been making use of such facial motion to mimic famous people, researchers have only recently begun to employ dynamic stimuli in studies on face processing (for review see: O'Toole, Roark, & Abdi, 2002). Probably the most obvious and intuitive functions of facial motion are the expression of emotion (e.g., Bassili, 1978; Kamachi et al., 2001) and the facilitation of communication (Campbell, de Gelder, & de Haan, 1996). But does facial motion also contribute to other aspects of face processing? Previous research has shown that facial motion can convey information about gender (Berry, 1991; Hill & Johnston, 2001), age (Berry, 1990), and, at least to some extent, identity (e.g., Bruce & Valentine, 1988; Hill & Johnston, 2001; Knight & Johnston, 1997; Lander, Bruce, & Hill, 2001; Lander, Christie, & Bruce, 1999; Rosenblum et al., 2002; Thornton & Kourtzi, 2002). It is this latter function—the role of facial motion during the processing of identity—that will concern us here.

Currently there are two main hypotheses as to how facial motion could, in principle, influence the processing of identity (O'Toole et al., 2002; Lander & Bruce, 2000). The "representation enhancement hypothesis" suggests that seeing faces in motion could indirectly facilitate face recognition by providing a better structural representation of a face. For example seeing a face moving rigidly might help to build up a 3D representation of a face. The "supplemental information hypothesis" suggests that facial motion could serve as a dynamic idiosyncratic signature independent of other sources of information. A twisted smile for example or a characteristic head tilt might be represented in addition to other identity specific information, such as the shape of the face. Clearly these two hypotheses do not need to be mutually exclusive, and as yet, there has been little experimental evidence which can choose between them.

---

* Corresponding author. Tel.: +49-7071-601-604; fax: +49-7071-601-616.

E-mail address: barbara.knappmeyer@tuebingen.mpg.de (B. Knappmeyer).

0042-6989/03/$ - see front matter © 2003 Elsevier Science Ltd. All rights reserved.
doi:10.1016/S0042-6989(03)00236-0

Regardless of how the influence of facial motion might be exerted, the ability to measure its effect in the laboratory seems to depend on a number of factors. Principle among these are the type of facial motion, the degree of familiarity with the faces and the viewing conditions.

In terms of type of motion, an important distinction is that between rigid and non-rigid movements. Rigid facial motion includes translations and rotations of the whole head whereas non-rigid facial motion refers to deformations of the face, for example while talking or displaying facial expressions of emotion. To date, advantageous effects of rigid motion have been demonstrated when unfamiliar faces were rotated (Pike, Kemp, Towell, & Phillips, 1997; Schiff, Banka, & Galdi, 1986) or when rigid head motion accompanying speech was tested (Hill & Johnston, 2001). However, Christie and Bruce (1998) failed to find beneficial effects for rigid head motion. With respect to non-rigid facial motion, the findings are up to now non-conclusive. While studies with very familiar or famous faces have consistently shown facilitating effects of (mainly) non-rigid facial motion (e.g., Bruce & Valentine, 1988; Lander & Bruce, 2000; Lander et al., 1999, 2001), it is less clear whether there are any beneficial effects of non-rigid facial motion for unfamiliar faces. Thornton and Kourtzi (2002) for example found beneficial effects of motion in a sequential matching paradigm, whereas Christie and Bruce (1998), using old-new recognition tasks, did not. Using an animated average head Hill and Johnston (2001) found that rigid motion consistently conveyed information about identity, however effects for purely non-rigid facial motion were only weak. It is probably safe to conclude from these studies that there is evidence for facial motion (rigid and non-rigid) carrying information about identity, but the effects are small and differ with familiarity and viewing conditions. For example, familiarity seems to be an important factor for beneficial effects of *non-rigid* facial motion to occur. This importance of familiarity may be related to the fact that it takes time and experience to pick up which facial movements are characteristic. Once the characteristic movements have been extracted they may be used as additional information, a suggestion which is clearly in line with the supplemental information hypothesis mentioned above.

With respect to viewing conditions, studies with full quality images have often failed to show beneficial effects of facial motion (e.g., Bruce et al., 1999; Christie & Bruce, 1998; Knight & Johnston, 1997; Lander et al., 1999). In contrast, many studies which have removed or degraded facial form cues in some way, have consistently shown advantages for moving over static presentation. For example, some studies have used point-light displays (Johansson, 1973) in which facial form cues only consist of a few high-contrast dots (Bassili, 1978; Berry, 1991; Bruce & Valentine, 1988; Rosenblum et al., 2002). Other studies have involved video images, which were degraded, either by photographic negation (Knight & Johnston, 1997; Lander et al., 1999), by thresholding (Lander & Bruce, 2000; Lander et al., 1999) or by pixelating and blurring the displays (Lander et al., 2001). Recently, an animated average head was used to explore the effects of individual facial motion in isolation by replacing individual facial form with that of an average face (Hill & Johnston, 2001). While such attempts to maximize the impact of motion have been successful and are clearly well motivated—that is, the ability to independently assess form and motion is very appealing—the resulting stimuli are nonetheless quite unnatural. That is, except in the laboratory, we will rarely be given the problem of recognizing a person purely, or even mostly, from motion (e.g., from just a few high contrast dots as in the point-light displays).

The purpose of the current work is to bring together a combination of tasks and techniques that would allow us to shed new light on the role of facial motion during the processing of identity, particularly with regard to the factors just outlined. Specifically, we made use of recent advances in computer animation and motion capture techniques to completely isolate non-rigid from rigid facial motion in an attempt to better understand the formers' contribution to identity judgments. To address the issue of familiarity, we used an incidental learning task in which exposure to both the facial form and the facial motion of a target individual was equated and controlled.

The term "facial form" is used here in the sense of "unchangeable aspects" of a face. This includes the individual shape (geometry) of a face, for example the thickness of the lips or the length, width and curvature of the nose in a neutral expression, as well as the skin texture (e.g., color) of a face. The term "facial motion" refers to "deformations over time". Such deformations over time contain both purely dynamic information and motion-induced spatial information. Purely dynamic information, for example, might be the speed with which a person reaches the peak of a smile or the duration the person persists in the full-smile expression. Motion-induced spatial information might be, for example, an asymmetric mouth position when a person displays a twisted smile or the position of the lip-corners at the peak of a smile.

The main focus of the current work was to investigate the integration of facial motion and facial form rather than to explore effects of facial motion in isolation. To do this we developed a testing method for presenting dynamic stimuli in which the *relevance* of form cues, rather than the image *quality* of form cues, was systematically varied. Form cues were manipulated by applying a 3D morphing technique (Blanz & Vetter, 1999)

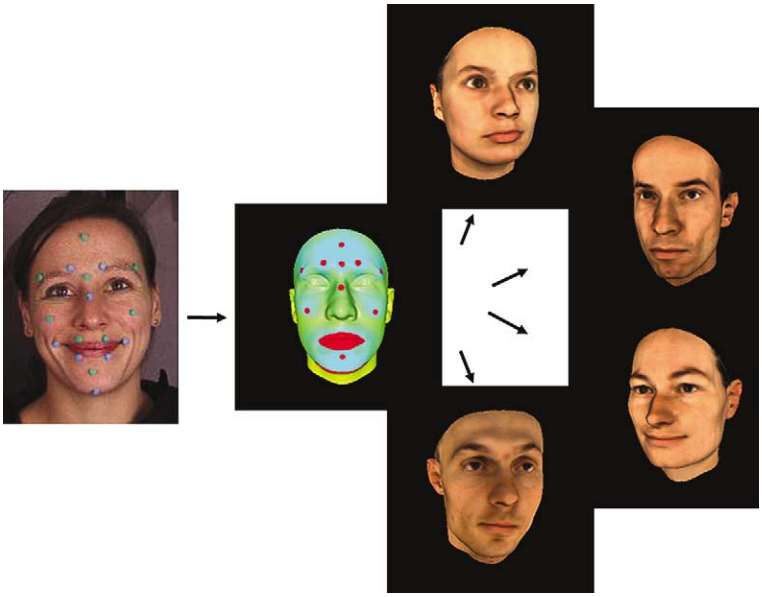

Fig. 1. Animation technique. The faces were animated using motion patterns captured from real human actors. They were filmed with a digital camera while performing a sequence of facial actions (smile, frown, surprise, chew etc.). The movement of 17 markers, which had been attached to the actor's faces, was tracked and extracted from the video using tracking software by famous3D Pty. Ltd. The facial animation software from the same company was used to apply these motion patterns to any given 3D model of a human face. To do this, marker positions, their "hot spots" (red) and their "regions of influence" (blue, green, yellow) were manually defined at first on an average face model, which was created from 200 3D Cyberware™ laserscans of human faces from the MPI face database and then automatically transferred to different faces. The motion of a marker drives its corresponding hot spot directly and animates the region of influence according to a quadratic fall-off function. The colored regions on the average face depicted here refer to the weights that result from overlapping regions of influence (decreasing weights from blue to yellow). A spline was used for the animation of the mouth. This clustering was then automatically transferred to the faces used in the current experiments by exploiting the point-to-point correspondence between all faces in the database. Thus the resulting animated faces differed either in their form (different laserscans) or in the motion pattern (different actors) that drove the animation, but never in the way in which the motion was applied to the faces (clustering). This animation technique allowed us to dissociate and independently vary facial motion and facial form. The snapshots from the animated faces (to the right of the figure) illustrate that the same motion pattern can be applied to different facial forms.

to high-quality laserscanned heads. A commercially available animation system for faces (3Dfamous Pty. Ltd.) was used to animate these heads using facial motion patterns captured from real human actors. The power of this technique is that it enabled us to animate any face with any motion (Fig. 1) to create situations where the two cues—form and motion—were either working in concert or conflict during the processing of identity. Thus, rather than trying to isolate form and motion, we wanted to explore how these two cues might be used at the same time.

In the experiments reported below we first familiarized observers with animated heads each performing the same basic sequence of non-rigid facial actions (e.g., smiling, frowning, chewing etc.), but with the slight idiosyncratic differences in the facial movements natural to different human actors. After familiarization, observers were asked to judge the identity of target faces, which were produced by morphing between the forms of the individual learned faces. The motion applied to these faces was always one of the motion patterns with which the observers were familiarized during learning (Fig. 2a). We hypothesized that observer's ability to determine the identity of the morphed target faces would be biased by the way the faces moved.

2. General methods

2.1. Participants

Seventy-five observers (age 17–40 years) from the Tübingen community were paid for their participation in these experiments. They were naïve as to the purpose of the research and had normal or corrected to normal vision. Twenty-nine observers (16 males/13 females) participated in Experiment 1, twenty-seven (12 males/15 females) in Experiment 2, thirteen (5 males/8 females) in Experiment 3 and sixteen (7 males/9 females) took part in the family resemblance task of Experiment 4. None of the observers participated in more than one of the experiments described below.

2.2. Laser scanned heads and morphing technique

All stimuli used in following experiments were created from 3D Cyberware™ laserscans of real human heads taken from the MPI database. [1] All manipulations of the heads, such as 3D morphing, anti-caricaturing,

---

[1] http://faces.kyb.tuebingen.mpg.de/

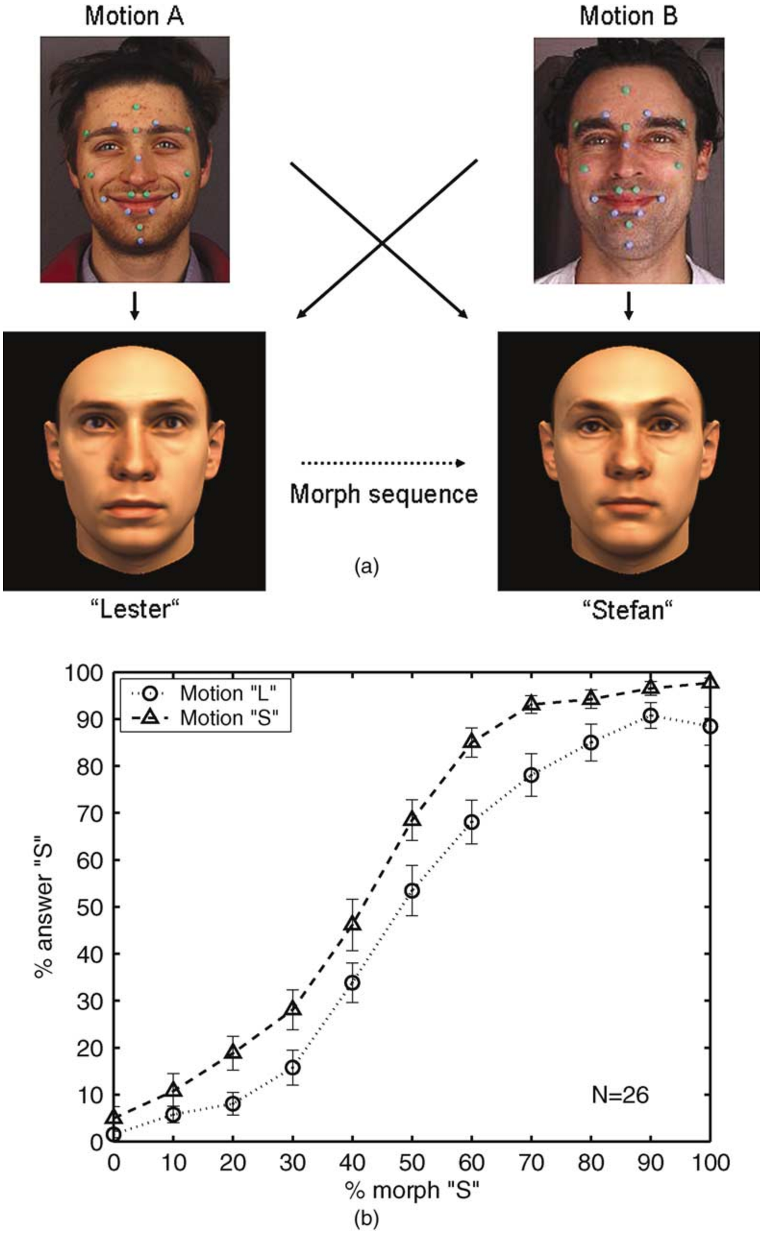

Fig. 2. Experiment 1. (a) Procedure. During a training phase observers were familiarized with two moving faces (e.g. labeled "Stefan" and "Lester"), one always animated with Motion A and the other one always animated with Motion B. The motion patterns consisted of the same sequence of facial expressions performed by different human actors. At test, each face of a morph sequence between "Stefan's" and "Lester's" facial form was combined with each of the two motion patterns, e.g. "Stefan's" face was animated with "Lester's" motion and "Lester's" face was animated with "Stefan's" motion. Observers had to decide whether these moving target faces looked more like "Stefan" or more like "Lester". (b) Results. Mean distribution (collapsed across observers and two different face pairs) of "Stefan" responses as a function of morph level. The psychometric functions reveal a biasing effect of facial motion for most morph levels. That is, when faces move with "Stefan's" motion, observers are more likely to respond "Stefan" than when exactly the same faces move with "Lester's" motion suggesting that observers based their identity judgments not solely on cues to individual facial form, but also on cues to individual facial motion. Table 1 summarizes the PSE analysis which was applied to assess the magnitude of this motion bias.

calculating an average head, applying a generic facial outline to the faces and replacement of individual skin texture were done using software developed by Blanz and Vetter (1999). An average facial outline served as a uniform aperture for all faces to prevent observers from using the cutting line from hair removal as a feature.

2.3. Motion capture and animation

The faces were animated using motion patterns captured from real human actors. Six non-professional human actors (4 males/2 females) were trained to perform the following sequence of posed facial expressions within a fixed time interval (8–10 s): neutral, smile, frown, surprise, chew three times, disgust, smile, neutral. A total of seventeen blue and green foam markers were placed on each actor's face, with markers positioned on or near the eyebrows, forehead, brow furrow, mouth, chin, nose and cheeks. Actors were filmed using a standard digital video camera. Head position was fixed to reduce rigid head movements and the actors were able to watch their faces in a monitor as they performed the facial actions.

The motion of the markers was tracked from the 25 f/s video clips using commercial tracking software by famous3D Pty. Ltd. The marker on the nose was used as a reference point to remove slight remaining head translations in the image plane. Thus the resulting motion patterns did not contain any rigid head motion, had the same overall duration and differed only in the subtle idiosyncratic way in which the actors naturally moved their faces.

These motion patterns were then applied to 3D models of human faces. Specifically, we animated 3D Cyberware™ laserscans of human heads using commercial facial animation software by famous3D Pty. Ltd. To do this, we manually defined corresponding marker positions on an average face model, which was calculated from 200 laserscans from the MPI face database (Blanz & Vetter, 1999). The motion of a marker drives its corresponding "hot spot" directly and animates a "region of influence" according to a quadratic fall-off function. The colored regions on the average face depicted in Fig. 1 refer to the weights that result from overlapping regions of influence. Each red spot corresponds to a "hot spot" for a given marker. The blue, green and yellow regions correspond to the overlapping "regions of influence" with blue standing for larger and yellow for smaller weights. This map of weights is referred to as "clustering". The clustering is somewhat arbitrary and was optimized to produce natural looking animations. Most importantly, the clustering was exactly the same for each face used in the current experiments. This was achieved by automatically transferring the clustering from the average face to every other face exploiting the dense point-to-point correspondence between all faces in the database (Blanz & Vetter, 1999). Thus the resulting animated faces differed either in their form (different laserscans) or in the motion pattern (different actors) that drove the animation, but never in the way in which the motion was applied to the faces (clustering). This animation technique allowed us to dissociate and independently vary facial motion and facial form. The animated faces were rendered into AVI format. During the experiments video clips were displayed with 12 f/s on a CRT monitor using IRIS Mediaplayer (SGI O2). Faces were presented in frontal view and covered approximately 3.6°×4.6° of visual angle.

2.4. Training procedure

While a single static picture can be enough to communicate the characteristic structural information of a face, significantly more exposure appears to be required in order to convey the characteristics of complex facial movements (Christie & Bruce, 1998). In the current studies, an incidental learning procedure was used to familiarize observers with individual faces moving in idiosyncratic ways. Observers were repeatedly shown two animated faces one after the other in a looped display along with a corresponding name label. Each face was presented for 30 s. Half of the observers were familiarized with face A animated with actor A's motion and face B animated with actor B's motion and the other half learned face A animated with actor B's motion and face B animated with actor A's motion. This was done to counterbalance for potential differences in the distinctiveness of the motion patterns.

While watching these animations observers were asked to fill out a questionnaire assessing personality traits of the faces. The questions were for example "Who looks overall happier to you?", "Who appears more dominant?". Observers spent approximately 30 min answering these questions and they were not aware that there would be a further categorization task. After this familiarization phase observers were able to accurately (100%) label the learned faces. Our intention with this training procedure was to familiarize observers with the particular faces and facial motions without them trying to explicitly memorize any aspect of the display.

2.5. Testing procedure

In Experiments 1–3, observers were shown spatial morphs that represented a gradual transition between the form of the learned faces and they were asked to identify these morphs as one of the two previously learned faces (2AFC). In Experiment 4 observers were presented with new faces that were morphed (50% morphs) with the form of the learned faces. They were asked to classify these morphs into two families. To test whether the incidentally learned motion patterns influenced observers' decisions in both of these tasks, the target faces were presented a number of times, half of the time animated with the motion pattern from one learned face and half of the time animated with the motion pattern from the other. The strength of the motion cue was held constant, i.e. there was no motion morphing involved.

3. Experiment 1

The purpose of Experiment 1 was to establish whether incidentally learned facial motion patterns would bias observers' perception of facial identity even when relevant form cues were available. Observers were first familiarized with two animated faces using the procedure described above. The faces differed in their form as well as in the way they were moving. During testing, the form cue in the target faces was systematically varied but the motion cue was held constant. This allowed us to measure the direct trade-off between facial form and facial motion. If characteristic motion influences the processing of identity, we would expect more "face A" responses for a face that moves like "face A" than for the same face moving like "face B". We assumed that such biasing effects might be particularly evident when form information was ambiguous (i.e. around the 50% morph).

3.1. Stimuli and procedure

For the training procedure two pairs of head models (2 female heads, 2 male heads) were chosen from the MPI head database. Since discriminating between just two faces is a very easy task, there was the risk that ceiling effects would leave little room for any motion-induced biases. To minimize this risk, and to account for the fact that our recording techniques capture facial form in more detail than facial motion, we decided to slightly weaken the form cue in the training faces for this initial experiment. This was achieved by morphing the faces 20% towards the average head (anti-caricaturing; Blanz & Vetter, 1999) and by applying an average skin texture to the faces. After this transformation the two faces looked a little more similar (Fig. 3), but were still easily distinguishable from each other after familiarization.

The two female faces were animated with facial motion captured from two female actors and the two male faces were animated with motion recorded from two male actors. Fourteen observers (9 females/5 males) were familiarized with the two male faces labeled "Stefan" and "Lester". The remaining fifteen observers (6 females/9 males) were familiarized with the two female faces labeled "Susi" and "Lara". At test, observers were asked to categorize single moving faces as either "Stefan" ("Susi") or "Lester" ("Lara"). The animated target faces were taken from a spatial morph sequence representing a gradual transition between "Stefan's" ("Susi's") and "Lester's" ("Lara's") facial forms. Eleven morphs covering the whole range between the form of "Stefan's" ("Susi's") face and the form of "Lester's" ("Lara's") face in 10% steps were used as target faces. For example, the 50% morph contained equal form information from "Stefan's" ("Susi's") and "Lester's" ("Lara's") face. To examine whether the incidentally learned motion patterns would nevertheless influence the perception of identity each target face was presented 20 times, 10 times animated with "Stefan's" ("Susi's") facial motion and 10 times animated with "Lester's" ("Lara's") facial motion. Observers were instructed that they would see faces, whose facial form might sometimes have been modified. Thus observers were if anything cued to pay attention to the form rather than to the motion, which is conservative with respect to our hypothesis. They were asked to indicate (via key press "S" or "L") after each target video (10 s), whether the face looked more like "Stefan" ("Susi") or more like "Lester" ("Lara"). Presentation order was randomized for each observer.

3.2. Results

Fig. 2b shows mean proportion of "S" responses for each morph and each motion pattern, collapsed across

Fig. 3. Stimuli. From left to right the columns depict: the female face pair used in Experiment 1 (top: "Lara", bottom: "Susi"), the male face pair used in Experiment 1 (top: "Lester", bottom: "Stefan"), both face pairs were slightly anti-caricatured, had an average facial outline and an average skin texture; the same male face pair from Experiment 1, but without anti-caricaturing (used in Experiments 2a, 3 and 4); and finally to the right: again the same face pair, but now without anti-caricaturing and with individual skin texture (used in Experiment 2b).

25 observers, the male and female face pairs and two form/motion combinations. Data from 3 out of 29 observers (2 trained with the male faces, 1 trained with the female faces) were excluded from the analysis because neither of the two psychometrical functions (data for each motion pattern) crossed the 50% level. This exclusion criterion was applied to avoid an overestimation of the biasing effect caused by single observers who categorized the faces solely based on the motion pattern. In terms of our hypothesis this is a conservative treatment of the data.

Across almost the whole range of the morph sequence, observers were more likely to respond "S" when the morphs moved with motion "S" than when exactly the same morphs moved with motion "L". The response differences varied between 3.5% and 16.9% with the largest difference for the 60% morph and the smallest for the 0% morph, which is identical to the learned face "Lester" ("Lara"). To more carefully assess the magnitude of the shift between the two psychometrical functions, a standard PSE (points of subjective equality) analysis was performed. The point of subjective equality (PSE), i.e. when observers' responses are at chance (50%), denotes the morph which is *perceived* as most ambiguous. This does not necessarily coincide with the physically most ambiguous morph (50%), for instance because of differences in the initial distinctiveness of the two faces or because of individual observer bias.

Cumulative Gaussian functions were fitted to individual observer data for each motion pattern separately. [2] The PSE values were extracted from the fitted data and were submitted to a 2 (face pair at training) × 2 (form-motion combination at training) × 2 (motion pattern at test) ANOVA. The ANOVA revealed a main effect of motion ($F(1,22) = 10.3$, $p = 0.004$) showing that the morph to which observers responded equally often with "S" and "L" needed to contain less form information from face "S" (38.8%, SE 2.3%) when it moved with the facial motion from face "S" than when it moved with the facial motion from face "L" (53.9%, SE 3.3%). No other main effects or interactions were found. Since there was no effect of face pair or form-motion combination, the data is presented collapsed across these conditions (Fig. 2b). However, for completeness we also present the PSE values for the two face pairs separately in Table 1. The fact that the PSEs are not symmetrical around the 50% morph may reflect variations in the distinctiveness of the faces. If, for example, the face "Stefan" was more distinctive than "Lester", the psychometric functions would be shifted to the left.

In addition to the PSE values, P25 and P75 values were calculated from the fitted data and one-tailed t-tests were applied to evaluate whether the differences (motion "L"–motion "S") were significantly larger than zero. The P75 (P25) value demarks the amount of form from face "S" in the morph required to obtain 75% (25%) "S" responses. Table 1 summarizes the results from this analysis. At all levels, less form information from face "S" was needed when the faces were moving with motion "S" than when they were moving with motion "L". These differences were significant, except for the P25 difference for the female faces, which reached only marginal significance.

While the motion bias for the male faces seems to be slightly larger than that obtained with the female faces, one-tailed t-tests comparing these differences (motion "L"–motion "S") between the two experiments (male face pair–female face pair) revealed that this apparent trend was not significant (P25: $t(24) = -0.0135$, $p = 0.5053$; PSE: $t(24) = 0.3648$, $p = 0.3592$; P75: $t(24) = 0.8424$, $p = 0.2039$). This is consistent with the lack of a main effect for face pair in the ANOVA on the PSE values reported above. Taken together (collapsed across face pairs) the magnitude of the bias varied between 14% and 16% and was present at all three levels of performance (PSE, P25, P75) suggesting that the motion biased observers' decisions not only when form information was completely ambiguous (at the point of subjective equality) but also when observers were able to reliably identify the faces (P25, P75).

3.3. Discussion

Using two pairs of animated faces we have shown in this first experiment that observers were using both facial form and facial motion while making their identity decision. The psychometric functions clearly show that observers were sensitive to form information. That is, the proportion of "Stefan" ("Susi") responses was close to 100%, when the face looked exactly like "Stefan" ("Susi") and close to 0%, when the face looked exactly like "Lester" ("Lara"). More interestingly, the shift between the two psychometrical functions suggests that the characteristic motion associated with an individual face during learning biased observers' identity judgments. The magnitude of this motion bias was equivalent to a 14%–16% change in relative form information. Furthermore, the bias was not only present when form information was completely ambiguous (at the point of subjective equality, PSE), but also when relevant form cues were available (across almost the whole range of the morph sequence). Thus, rather than exclusively relying on either facial form or on facial motion, observers seem to integrate both sources of information. However,

---

[2] The PSE analysis reported here was carried out using the MATLAB statistics toolbox. Re-analyses of the data using software by Wichmann and Hill (2001) specifically developed for the fitting of psychometrical functions yielded the same pattern of results.

Table 1
Experiment 1: Mean PSE, P25, P75 values (% form from face "S" in the morph) collapsed across observers for male and female face pairs

Face pair		Motion "L"	Motion "S"	Difference	*t*	*df*	*p*
Male	P25	34.9 (SE 3.2)	20.9 (SE 5.2)	14.1	1.9	11	0.0409
	PSE	53.2 (SE 5.1)	36.2 (SE 3.3)	17.0	2.3	11	0.0225
	P75	71.5 (SE 7.6)	51.5 (SE 2.8)	20.0	2.5	11	0.0140
Female	P25	39.6 (SE 3.4)	25.4 (SE 7.0)	14.2	1.5	13	0.0745
	PSE	54.5 (SE 4.3)	41.1 (SE 3.0)	13.4	2.0	13	0.0310
	P75	69.4 (SE 5.9)	56.8 (SE 2.7)	12.6	2.8	13	0.0082
Collapsed (male and female)	P25	37.4 (SE 2.3)	23.3 (SE 4.4)	14.2	2.4	25	0.0123
	PSE	53.9 (SE 3.3)	38.8 (SE 2.3)	15.1	3.1	25	0.0024
	P75	70.3 (SE 4.7)	54.4 (SE 2.0)	16.0	3.7	25	0.0006

To assess the magnitude of the biasing effect in Experiment 1, points of subjective equality (PSEs), P25 and P75 were calculated for each observer and each motion pattern by fitting cumulative gauss functions. The values denote how much form information from "Stefan's" face was required in the morph to elicit 25%, 50% or 75% "S" responses. One-tailed *t*-tests were applied to assess the magnitude of the differences between the two psychometric functions. The data show that at each level of performance less form information of "S" is contained in the morph when it is moving with motion from face "S" than when it is moving with motion from face "L".

the faces used in this initial experiment were manipulated to look quite similar (Fig. 3). In the following experiment we will investigate whether these manipulations influenced our findings.

4. Experiment 2

The faces used for training in the previous experiment looked very similar, since we had weakened the form cue to ensure that the task was not trivial. This manipulation might have encouraged observers to pay more attention to the facial motions than under more "natural" conditions. The purpose of Experiment 2 was to systematically investigate whether the motion bias observed in Experiment 1 crucially depended on these form manipulations, which included 20% anti-caricaturing, i.e. morphing towards the average face, and substitution of the individual skin texture with an average skin texture. Thus we replicated Experiment 1, but now the training faces were not anti-caricatured, but they retained their original inner features (Experiment 2a) and individual skin texture was applied to the training faces (Experiment 2b). We assumed that increasing the strength of the form cue (shape and texture) at training would weaken the motion bias.

5. Experiment 2a

5.1. Stimuli and procedure

The same male face pair from Experiment 1 was used, but now the faces were not anti-caricatured (Fig. 3). That is, the inner features of the faces differed in their natural way. However the skin texture was still taken from the average face. Since there was no effect of face pair in the previous experiment, all 14 observers (6 females/8 males) were now familiarized with the same pair of male heads. The faces were animated with the same motion sequences as before. Unlike in the first experiment, observers watched each morph only five times animated with Stefan's motion and five times animated with Lester's motion. Otherwise the procedure and the stimuli were identical to those in Experiment 1.

5.2. Results

Data from 1 out of the 14 observers met the exclusion criterion described in Experiment 1 and was thus not included in the analysis. Fig. 4a shows the mean responses (collapsed across observers) for each morph and each motion pattern. The proportion of "Stefan" responses was very low for the 0% morph ("Lester's" face) and very high for the 100% morph ("Stefan's" face) suggesting that form influenced observer's decision. Furthermore, across a large portion of the morph sequence observers responded more often "Stefan" when the morphs were moving with "Stefan's" motion than when they were moving with "Lester's" motion. This response difference was largest for the 60% morph (32.3%), smallest for the 0% morph (3.1%)

The PSE analysis revealed significant differences at all three levels of performance (PSE, P25 and P75) (see Table 2). More specifically, the corresponding morphs contained significantly less form information from "Stefan" when it was moving with "Stefan's" motion than when it was moving with "Lester's" motion at all three levels of performance. The magnitude of the bias ranged from 22.7% to 25.1%. The motion bias seems to be larger than in the first experiment. However, one-tailed *t*-tests (comparing data from the male faces in Experiment 1 with data from Experiment 2a) revealed that this difference was not significant (P25:

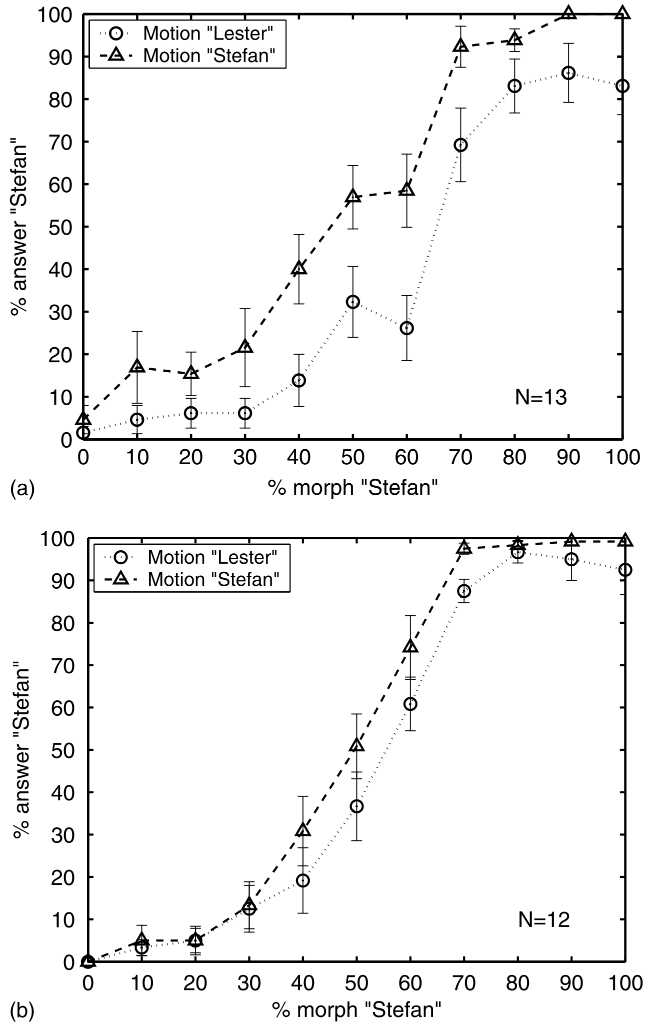

Fig. 4. (a) Experiment 2a. Mean distribution (collapsed across observers) of “Stefan” responses as a function of morph level. In contrast to Experiment 1, the training faces were not morphed towards the average face. Thus they were slightly more distinct. The psychometrical functions still reveal a motion bias. (b) Experiment 2b. Mean distribution (collapsed across observers) of “Stefan” responses as a function of morph level. In contrast to the previous experiment, individual skin texture was applied to the training faces. Thus they were even more distinct. The psychometrical functions still reveal a motion bias, which was smaller than in the previous experiments. After the task, observers were asked to discriminate the learned motion patterns applied to an average head. They performed 87% ($N = 11$, SE 4.6%) correct suggesting that they were able to distinguish between the motion patterns even though the bias was smaller.

$t(23) = -0.6914, p = 0.2481$; PSE: $t(23) = -0.5467, p = 0.2949$; P75: $t(23) = -0.3958, p = 0.3480$).

5.3. Discussion

As in Experiment 1 the data show a clear trade-off between facial form and facial motion. Since the motion bias was still present we conclude that the anti-caricaturing (Experiment 1) was not crucial to obtain the effect.

6. Experiment 2b

6.1. Stimuli and procedure

The procedure and the stimuli were the same as in Experiment 2a except for one further manipulation: individual skin texture was applied to the training faces. In addition, after having completed the whole task (training and testing phase) observers were presented with an average face which was animated with the learned motion patterns. They were asked to decide which of the two motion patterns was used for the animation. This task consisted of 20 trials: 10 presentations of each motion pattern, randomly presented.

6.2. Results

The data from 1 out of 13 observers met the exclusion criterion and was thus not included in the analysis. The mean proportions of “Stefan” responses are shown in Fig. 4b. Again across the whole range of the morph sequences there was a trend to respond “Stefan” more often when the morphs were moving with “Stefan’s” motion than when they were moving with “Lester’s” motion. The difference varied from 0% (for the 0 and the 20% morph) to 14.2% (for the 50% morph).

The PSE analysis revealed significant differences at the P25 and the PSE (Table 3), but only a marginally significant trend at the P75 level. The bias was apparently smaller than in the previous experiments (3.9%–11.4% “Stefan” in morph). One-tailed t-tests revealed that the bias was significantly smaller than in Experiment 2a at the P25 level ($t(23) = 1.7968$, $p = 0.0428$), marginally smaller at the PSE level ($t(23) = 1.4706$, $p = 0.0775$) and not significantly smaller at the P75 level ($t(23) = 1.1036$, $p = 0.1406$).

When observers were asked to discriminate the two motion patterns applied to an average head, they

Table 2
Experiment 2a: Mean PSE, P25, P75 values (% form from face “S” in the morph) collapsed across observers

	Motion “Lester”	Motion “Stefan”	Difference	$t(12)$	p
P25	51.3 (SE 5.6)	28.6 (SE 7.2)	22.7	2.3	0.020
PSE	66.8 (SE 5.9)	42.9 (SE 5.0)	23.9	2.4	0.016
P75	82.3 (SE 7.6)	57.2 (SE 3.8)	25.1	2.5	0.015

In contrast to the previous experiment the training faces were not morphed towards the average face in this experiment. See Table 1 for a more detailed description of the data format.

Table 3
Experiment 2b: Mean PSE, P25, P75 values (% form from face "S" in the morph) collapsed across observers

	Motion "Lester"	Motion "Stefan"	Difference	$t(11)$	p
P25	42.7 (SE 3.4)	38.8 (SE 3.9)	3.9	1.95	0.038
PSE	55.3 (SE 3.6)	47.7 (SE 2.9)	7.7	1.88	0.044
P75	68.0 (SE 6.3)	56.5 (SE 2.2)	11.4	1.71	0.058

In contrast to Experiment 2a the training faces also differed in their skin texture. See Table 1 for a more detailed description of the data format.

performed very accurately, with performance averaging around 87% ($N = 11$, SE 4.6%).

6.3. Discussion

The data from Experiment 2b still reflect a trade-off between facial form and facial motion. However the motion bias is smaller, which is in line with our prediction concerning increased form cues. The fact that motion has any impact in this experiment is impressive given that adding individual skin texture considerably increases the useful form information in the animations (Fig. 3). The fact that observers were still able to accurately (87%) distinguish between the two different motion patterns when these were applied to an average face, suggests that the increased form information did not block the extraction of motion during learning, but rather provided a much more robust cue during testing.

7. Experiment 3

The previous experiments provide evidence that both facial form and facial motion seem to be used during the processing of identity. While it has been well established that processing of facial form is tuned to upright faces, the so-called inversion effect (e.g., Thompson, 1980; Yin, 1969), it is less clear whether this is also true for the processing of facial motion. For example, using an animated average face Hill and Johnston (2001) found that even when the animated face was turned upside down, observers were still able to identify one out of three facial motions taken from a different human actor. However, the performance was worse than for the upright presentation. Similarly, Lander et al. (1999) reported an advantage for moving compared with multiple static displays even when faces were presented upside down. In contrast, Knight and Johnston (1997) did not find such an advantage for inverted faces. The purpose of the following experiment was to test whether the motion bias we observed in the previous experiments would be robust against rotation in the image plane.

7.1. Stimuli and procedure

The stimuli at training were exactly the same as in Experiment 2a. Observers were familiarized with the faces presented in upright orientation, but now at test the faces were presented upside down, i.e. rotated 180° in the image plane.

7.2. Results

Data from 1 out of the 13 (5 males/8 females) observers were discarded from the analysis according to the exclusion criterion described above. The results are summarized in Fig. 5. The performance at the endpoints of the morph sequence was worse than in the previous experiments, e.g. 10–25.8% "Stefan" responses for the 0% morph (= "Lester's" facial form) and 83.3–93.3% for the 100% morph (= "Stefan's facial form). The difference in "Stefan" responses depending on the motion pattern varied between 9.2% (for the 90% morph) and 25.8% (for the 80% morph).

The PSE analysis revealed a significant motion bias at all three levels of performance (Table 4). The magnitude of the bias varied between 18.7% and 25.7%. Two-tailed t-test revealed that the magnitude of this bias was neither different from that in Experiment 2a (P25: $t(23) = 0.2920$, $p = 0.7729$; PSE: $t(23) = 0.1346$, $p = 0.8941$;

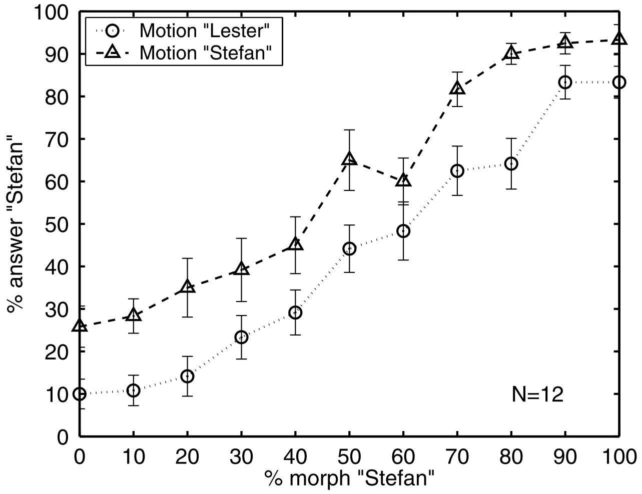

Fig. 5. Experiment 3. Mean distribution (collapsed across observers) of "Stefan" responses as a function of morph level. Observers were trained with upright faces and tested with faces that were turned upside down. The overall performance was worse than in the previous experiments (inversion effect), but the motion bias was still present suggesting that some aspects of the spatio-temporal pattern seem to be invariant to rotations in the image plane.

Table 4
Experiment 3: Mean PSE, P25, P75 values (% form from face "S" in the morph) collapsed across observers

	Motion "Lester"	Motion "Stefan"	Difference	$t(11)$	p
P25	29.3 (SE 7.8)	10.6 (SE 4.4)	18.7	2.1	0.033
PSE	59.3 (SE 4.2)	37.1 (SE 4.7)	22.2	2.9	0.007
P75	89.3 (SE 5.8)	63.7 (SE 6.2)	25.7	3.2	0.004

Observers were trained with the same faces as in Experiment 2a. While trained with upright faces, at test observers were presented with faces that were turned upside down. See Table 1 for a more detailed description of the data format.

P75: $t(23) = -0.0411$, $p = 0.9676$) nor from that obtained for the male faces in Experiment 1 (P25: $t(22) = -0.3980$, $p = 0.6945$; PSE: $t(22) = -0.4834$, $p = 0.6336$; P75: $t(22) = -0.5074$, $p = 0.6169$).

7.3. Discussion

The data show that even though observers were familiarized with upright moving faces the facial motion still influenced their identity decision when the target faces were presented upside down. This is quite impressive given the subtlety of the differences in the motion patterns. Some aspects of the motion patterns seem to be rather invariant across rotations in the image plane. This is consistent with Lander et al. (1999) and Hill and Johnston (2001) who also found that some useful aspects of facial motion seem to be preserved in inverted displays. However a direct comparison with these studies has to be handled with care due to differences in task and stimuli. The magnitude of the motion bias in the current experiment is comparable with the equivalent upright condition (Experiment 2a). However, it is larger at the endpoints, probably because turning the faces upside down is a non-optimal viewing condition for extracting the facial form. This might lead to additional cues (such as facial motion) becoming more important. The fact that the overall performance is worse is consistent with the well-known inversion effect for pictures of faces (e.g., Thompson, 1980; Yin, 1969). There may be several reasons why motion still biases observers' responses under these conditions. For example, purely dynamic information (e.g., the rhythm and timing of a sequence) is unaffected by rotation in the image plane. Possibly it is this aspect of the motion pattern that is responsible for the continued motion bias. This would be consistent with research showing that observers are sensitive to the exact rhythm of facial motion (Lander & Bruce, 2000). Alternatively, the motion patterns used in the current experiments might create some particularly distinctive spatio-temporal feature (e.g., a skewed smile) which might also be easily perceived in the rotated display. Clearly, further research is needed to more fully understand the effect of inversion on the processing of moving faces.

8. Experiment 4

While the previous experiments provide convergent evidence that facial motion influenced observer's perception of identity even when relevant form cues were present, the particular judgment task we used might have encouraged them to adopt strategies quite different from the way they would usually process facial identity. That is, since observers were required to make very fine-grained distinctions between highly similar faces within the morph sequence, they might have focused on very subtle features in the animated faces. To reduce the likelihood of a feature based strategy, we designed a new "family resemblance" task, which we hoped would encourage observers to rely more on their overall impression of the faces. The target faces were now created by spatially morphing 20 new individual faces with the learned facial forms (50% morphs). Observers were instructed that they would see novel faces of people who are related to one of the two learned faces and they were asked to categorize them with respect to their "family membership". Each novel face was presented with each facial motion of the learned faces. Based on our previous findings, we assumed that observers' responses would reflect an integration of cues from both sources of information, facial form and facial motion.

8.1. Stimuli and procedure

The same male face pair as in Experiment 2a was used for the current experiment. The faces were animated with facial motions from two new male actors. The sequence of facial expressions remained the same as in the previous experiments, but the overall duration was shorter (8s). Observers were familiarized with these animated faces labeled "Stefan" and "Lester" in the same way as described above.

At test, observers were now presented with 40 novel faces created by spatially morphing a novel face from the database (20 different faces per "family": 10 males/ 10 females) with either "Stefan" or "Lester" (50% morphs). Thus, faces within a "family" shared some common geometry but were nonetheless considerably more distinct from each other than the morphed faces used in the previous experiments. Examples of these

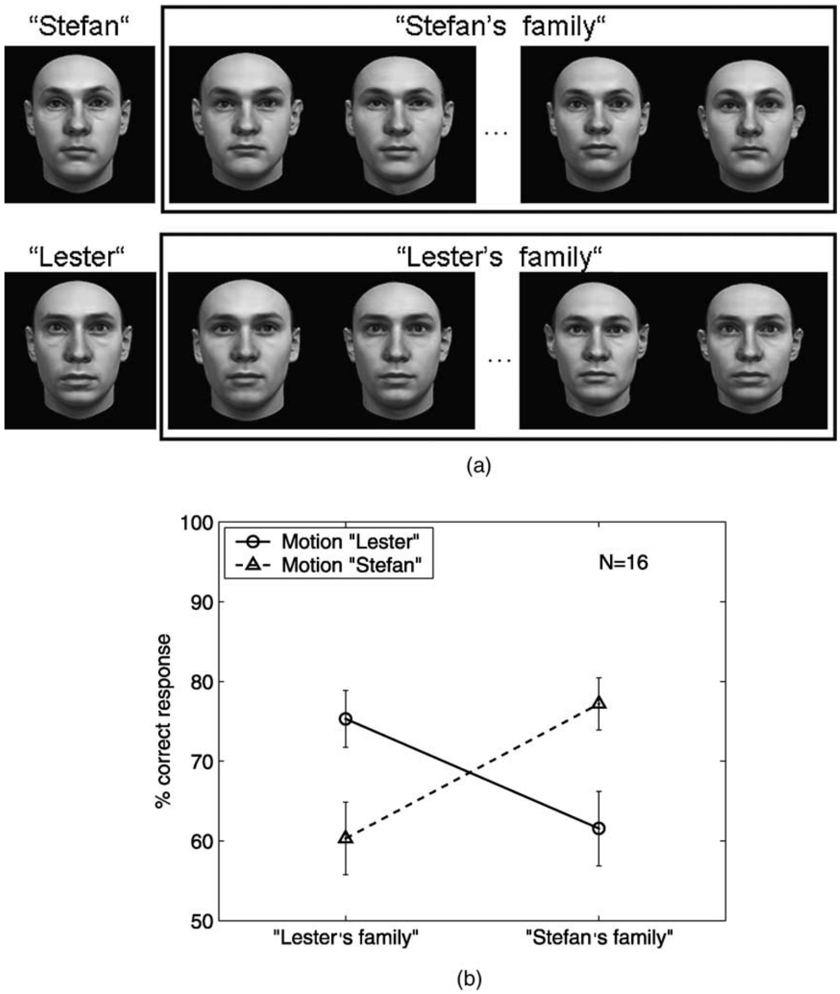

Fig. 6. Experiment 4. (a) Procedure. At training, observers were again familiarized with two animated faces (labeled "Stefan" and "Lester"). At test they were now shown 40 new moving faces and were asked to decide whether these faces belong to members of either "Stefan's" or "Lester's family". Each "family" consisted of 20 novel faces (10 males/10 females) morphed halfway towards "Stefan" or "Lester" (50% morphs). Thus faces within one "family" resembled each other with respect to their form. Each face was presented twice: once animated with "Lester's" motion and once animated with "Stefan's" motion. (b) Results. Mean percent correct (collapsed across observers) defined on the basis of the form cue, e.g. response "Stefan's family" counts as correct for a face that was morphed with "Stefan". Error bar represent standard errors. Performance was above chance in all conditions, suggesting that observers used the facial form cue in this task. However, observers also used the facial motion cue to make their decision, as revealed by a strong interaction between form and motion. That is, when the motion cue was consistent with the form cue, performance was considerably more accurate than when it was inconsistent.

stimuli are shown in Fig. 6a. Observers were instructed that they would see novel faces of people who were related to "Stefan" or "Lester". On each trial, their task was to categorize a single novel face as either "a member of Stefan's family" or "a member of Lester's family". Each face was presented twice, once moving with "Lester's" facial motion and once moving with "Stefan's" facial motion. Response was given via key press ("S" or "L").

8.2. Results

Fig. 6b shows the mean percentage (collapsed across observers) of correct responses for each family and each motion pattern. Responses were defined as "correct" when they were consistent with the form cue in the stimulus, e.g. the response "Stefan's family" to a face that was morphed with "Stefan" counted as a correct response irrespective of the motion pattern that was used to animate the face. A 2 (form motion combination at training) × 2 (form cue) × 2 (motion cue) ANOVA revealed a significant interaction between form and motion ($F(1,15) = 7.6$, $p = 0.02$). When the faces were animated with the consistent motion, i.e. "Lester's family" with "Lester's" motion and "Stefan's family" with "Stefan's" motion, observers were considerably more accurate, than when exactly the same faces were animated with the inconsistent motion (Table 5). There

Table 5
Experiment 4: Family resemblance task

Facial form	Facial motion "Lester"				Facial motion "Stefan"			
	% Correct	SE	$t(15)$	p	% Correct	SE	$t(15)$	p
New + 50% "Lester"	75.3	3.6	7.1	$p < 0.001$	60.3	4.6	2.3	$p = 0.038$
New + 50% "Stefan"	61.6	4.7	2.5	$p = 0.026$	77.2	3.3	8.3	$p < 0.001$

Percent correct responses (i.e. response "Lester's family" when facial form was morphed with "Lester") averaged across observers ($n = 16$). The t-values were calculated to asses whether performance was above chance level (50%).

were no other main effects or interactions. Finally, t-tests ($p < 0.05$) reveal that observers were consistently above chance (50%) in all conditions.

8.3. Discussion

Using a more natural task, these data again indicate that cues to both facial form and facial motion, contribute to the processing of identity. Given the definition of "correct response" we used in this experiment, an ideal observer who relied solely on form information would perform with 100% correct responses irrespective of the facial motion pattern that was used to animate the faces. In contrast, an ideal observer who relied exclusively on the motion cue would perform with 100% correct responses, when the facial motion is consistent with the facial form, but with 0% correct response when form and motion cue were inconsistent. The data clearly does not conform to either of these two cases. Rather, observers seem to base their decision on some combination of form and motion, a pattern more similar to that predicted from an ideal observer, who integrated form and motion cues with equal weights. Such an observer would perform at 100%, when form and motion cue are consistent and at chance (50%), when form and motion cue are inconsistent The fact that performance was above chance in all conditions, even in the inconsistent condition, may reflect a slight advantage of the form over the motion cue.

9. General discussion

In the series of experiments reported here we found consistent evidence that non-rigid facial motion biased observers' perception of identity. Furthermore, by employing a variety of new tasks and new techniques, we have demonstrated that information provided by facial form and facial motion seems to be integrated during the processing of identity.

In the first three experiments we measured responses to morphed faces that represented a continuous transition between the forms of two learned faces. In Experiment 1, there was a consistent shift between the psychometrical functions measured for the different learned facial motions that were applied to these morphs, suggesting that facial motion biased observers' identity decisions. This shift was observable across almost the whole range of the morph sequence, even when relevant form information was available. While the learned faces in this initial experiment looked very similar, Experiment 2 replicated these findings with faces that were considerably more distinct. Although the motion bias was slightly weaker in this experiment (Experiment 2b), it was still present. Furthermore, observers were still able to reliably distinguish between facial motions when they were presented on an average face, suggesting that the individual motion patterns had still been extracted, but were given less weight during the integration. In Experiment 3, we found that a motion bias could still be observed even when target faces were rotated 180° in the picture plane, suggesting that some aspect of the spatio-temporal pattern was rotation-invariant. Possibly it is the purely dynamic information (e.g., the speed or a characteristic rhythm) which causes the bias in such a condition. Alternatively, the motion patterns might contain some very distinctive feature (for example a twisted smile) which can be easily observed even when the face is turned upside down. Finally, in Experiment 4, a family resemblance task was used to demonstrate that the observed motion bias generalized to tasks involving a larger variety of facial forms. Again, the results suggested that observers integrated both facial form and facial motion during the processing of identity.

The finding that facial motion biased observers' identity decisions is consistent with previous research showing that such motion patterns can carry information about identity (e.g. Bruce & Valentine, 1988; Knight & Johnston, 1997; Lander et al., 1999; Rosenblum et al., 2002; Thornton & Kourtzi, 2002). Recently, however, Hill and Johnston (2001), using a very similar technique, found only weak effects of purely non-rigid facial motion compared to robust effects of rigid head motion. The stronger effects of non-rigid motion observed in the current work may reflect subtle differences in either the task or the stimuli used in these studies. For example, we used expressive, rather than speech-related, movements and we introduced an incidental learning phase to familiarize observers with specific motion patterns. Familiarity seems to be one factor that has a strong impact on the detection of motion effects (e.g., O'Toole et al., 2002 for a review).

The presence of robust non-rigid motion effects in the current work is particularly interesting as the observers had access to both facial form and facial motion cues at learning and test. That is, in contrast to previous research (e.g., Bassili, 1978; Bruce & Valentine, 1988; Hill & Johnston, 2001; Knight & Johnston, 1997; Lander et al., 1999; Rosenblum et al., 2002), which focused on reducing or eliminating the form cue to investigate effects of motion in isolation, the current study explored the integration of facial form and facial motion. While investigating effects of isolated or enhanced motion may be very useful in order to more fully understand their potential impact on face processing, outside of the laboratory, such isolated cues may rarely be used. Although the current animation and morphing techniques may raise similar concerns regarding ecological validity, our experimental situation was more natural in the sense that the two major sources of information—form and motion—were available in stimuli with high image quality. The fact that motion still biased observers' judgments under these conditions strongly suggests that facial movements are not redundant cues to identity, as has sometimes been suggested (e.g., Bruce & Valentine, 1988; Knight & Johnston, 1997). More specifically, the results from Experiment 4 suggest that facial form and facial motion might be integrated with almost equal weights in decisions about identity, with only a slight advantage for facial form. That is, observers performed close to chance when form and motion cue were inconsistent, while they were performing well above chance when both cues were consistent. If they had mainly relied on form or mainly on motion, this would have been revealed in a different pattern of results for the inconsistent condition. However, more research is needed to determine the exact weights and functions which are applied during the integration of these two cues.

In the introduction we outlined two current hypotheses (O'Toole et al., 2002; Lander & Bruce, 2000) as to how facial motion could, theoretically, contribute to face recognition. The representation enhancement hypothesis suggests that seeing a face move provides a better structural representation of that face than static images. The supplemental information hypothesis suggests that facial motion can be used as an additional source of identity specific information. As the experiments reported here do not directly compare static versus dynamic presentation modes, our findings are uninformative with respect to the first hypothesis. However, the fact that we found response differences for identical faces that differed only in the way they moved is clearly consistent with the supplemental information hypothesis. Furthermore, as mentioned above, the overall pattern of results in all of our studies seems to suggest that both form and motion are being used during the processing of identity, a finding which again would seem to support the supplemental information hypothesis.

More generally, the motion capture and animation techniques employed in the current work open the door for the systematic study of the use of form and motion across a whole range of topics, which have previously only been explored with static images of faces. For example, by using dynamic morphing and caricaturing methods (Blanz & Vetter, 1999; Giese & Poggio, 2000), it is possible to investigate the influence of motion on facial caricature (e.g., Giese, Knappmeyer, & Bülthoff, 2002; Hill et al., 2002; Rhodes, Brennan, & Carey, 1987) and viewpoint effects (e.g., Troje & Bülthoff, 1996; Watson, Hill, & Johnston, 2002; Troje & Kersten, 1999). As already mentioned the techniques and tasks described in this paper also allow researchers to disentangle rigid from non-rigid facial motion and would make it feasible to systematically study the role of motion during learning, i.e. when an unfamiliar face becomes familiar. While the current paper has been exclusively concerned with the contribution of facial motion to the processing of identity, similar techniques can be applied to the study of other aspects of face processing, for example facial attractiveness (Knappmeyer, Thornton, Etcoff, & Bülthoff, 2002).

Finally, we believe the current findings have important implications for cognitive and neural models of face perception (Bruce & Young, 1986; Haxby, Hoffman, & Gobbini, 2000). Such models have typically stressed a separation of the invariant aspects (facial form) and changeable aspects (facial motion) of faces into independent processing systems and have assigned decisions about facial identity firmly with the former system. While earlier studies have shown that either of these systems can compute identity in isolation, here we have shown that when operating together, a compromise is reached with responses reflecting input from both types of information. Such a compromise, is consistent with a growing body of evidence from behavioral (Bernstein & Cooper, 1997; Lorenceau & Alais, 2001; Stone & Harper, 1999; Wallis & Bülthoff, 2001), computational (Giese & Poggio, 2003) and neural studies (Bradley, Chang, & Andersen, 1998; Decety & Grezes, 1999; Grossman & Blake, 2002; Haxby et al., 2000; Kourtzi, Bülthoff, Erb, & Grodd, 2002; Oram & Perrett, 1994) which emphasize the use of both form and motion during the recognition of many classes of objects.

Acknowledgements

The authors would like to thank Fiona Newell, Chris Christou, Markus Graf, Alice O'Toole and Zoe Kourtzi for comments on an earlier draft of the manuscript; Mario Kleiner, Volker Blanz and Curzio Basso for helping to produce the stimuli and our colleagues, who allowed us to record their facial motions. We would also like to thank two anonymous reviewers for their valuable and thoughtful comments.

References

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

Bernstein, L. J., & Cooper, L. A. (1997). Direction of motion influences perceptual identification of ambiguous figures. *Journal of Experimental Psychology: Human Perception and Performance, 23*(3), 721–737.

Berry, D. S. (1990). What can a moving face tell us? *Journal of Personality and Social Psychology, 58*(6), 1004–1014.

Berry, D. S. (1991). Child and adult sensitivity to gender information in patterns of facial motion. *Ecological Psychology, 3*(4), 348–366.

Blanz, V., & Vetter, T. (1999). A morphable model for the synthesis of 3D faces. *Computer Graphics Proceedings SIGGRAPH*, 187–194.

Bradley, D. C., Chang, G. C., & Andersen, R. A. (1998). Encoding of three-dimensional structure-from-motion by primate area MT neurons. *Nature, 392*, 714–717.

Bruce, V., Henderson, Z., Greenwood, K., Hancock, P. J. B., Burton, A. M., & Miller, P. (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology—Applied, 5*(4), 339–360.

Bruce, V., & Valentine, T. (1988). When a nod's as good as a wink: The role of dynamic information in facial recognition. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: current research and issues* (Vol. 1, pp. 169–174). Chichester: Wiley.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 7*, 305–327.

Campbell, R., de Gelder, B., & de Haan, E. (1996). The lateralization of lip-reading: A second look. *Neuropsychologia, 34*(12), 1235–1240.

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

Davies, G., Ellis, H. D., & Shepherd, J. (1978). Face recognition accuracy as a function of mode of representation. *Journal of Applied Psychology, 63*, 180–187.

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

Giese, M. A., & Poggio, T. (2000). Morphable models for the analysis and synthesis of complex motion patterns. *International Journal of Computer Vision, 38*(1), 59–73.

Giese, M. A., & Poggio, T. (2003). Neural mechanisms for the recognition of biological movements. *Nature Reviews Neuroscience, 4*, 179–192.

Giese, M. A., Knappmeyer, B., & Bülthoff, H. H. (2002). Automatic synthesis of sequences of human movements by linear combination of learned example patterns. In H. H. Bülthoff, S. W. Lee, T. Poggio, & C. Wallraven (Eds.), *Biologically motivated computer vision* (pp. 538–547).

Grossman, E. D., & Blake, R. (2002). Brain Areas active during visual perception of biological motion. *Neuron, 35*, 1167–1175.

Hancock, P. J. B., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences, 4*(9), 330–337.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*(6), 223–233.

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

Hill, H., Pollick, F. E., Kamachi, M., Troje, N. F., Watson, T., & Johnston, A. (2002). Using the principles of facial caricature to exaggerate human motion. *Perception, 31*(Suppl.), 60.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics, 1*, 201–211.

Kamachi, M., Bruce, V., Mukaida, S., Gyoba, J., Yoshikawa, S., & Akamatsu, S. (2001). Dynamic properties influence the perception of facial expressions. *Perception, 30*, 875–887.

Knappmeyer, B., Thornton, I. M., Etcoff, N., & Bülthoff, H. (2002). *Facial motion and the perception of facial attractiveness*. Vision ScienceS Society, 2nd Annual Meeting, Sarasota, Florida, 10–15 May.

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

Kourtzi, Z., Bülthoff, H. H., Erb, M., & Grodd, W. (2002). Object-selective responses in the human motion area MT/MST. *Nature Neuroscience, 5*(1), 17–18.

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 pixalation and blurring on masking the identity of familiar faces. *Applied Cognitive Psychology, 15*, 101–116.

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

Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level after-effects. *Nature Neuroscience, 4*(1), 89–94.

Lorenceau, J., & Alais, D. (2001). Form constraints in motion binding. *Nature Neuroscience, 4*(7), 745–751.

Oram, M. W., & Perrett, D. I. (1994). Responses of anterior superior temporal polysensory (STPa) neurons to "biological motion" stimuli. *Journal of Cognitive Neuroscience, 6*, 99–116.

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

Perrett, D. I., Lee, K. J., Penton-Voack, I., Rowland, D., Yoshikawa, S., Burt, D. M., Henzi, S. P., Castles, D. L., & Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature, 394*, 884–887.

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

Rhodes, G., Brennan, S., & Carey, S. (1987). Identification and ratings of caricatures: Implications for mental representation of faces. *Cognitive Psychology, 19*, 473–497.

Rosenblum, L. D., Yakel, D. A., Baseer, N., Panchal, A., Nordase, B. C., & Niehus, R. P. (2002). Visual speech information for face recognition. *Perception and Psychophysics, 64*(2), 220–229.

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

Stone, J., & Harper, N. (1999). Object recognition: View-specificity and motion-specificity. *Vision Research, 39*, 4032–4044.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology-A, 46*, 225–245.

Thompson, P. (1980). Margaret thatcher: A new illusion? *Perception, 9*, 483–484.

Thornton, I. M., & Kourtzi, Z. (2002). A matching advantage for dynamic human faces. *Perception, 31*(1), 113–132.

Troje, N. F., & Bülthoff, H. H. (1998). How is bilateral symmetry of human faces used for recognition of novel views. *Vision Research, 38*, 79–89.

Troje, N. F., & Bülthoff, H. H. (1996). Face recognition under varying poses: The role of texture and shape. *Vision Research, 36*(12), 1761–1771.

Troje, N. F., & Kersten, D. (1999). Viewpoint-dependent recognition of familiar faces. *Perception, 28*, 483–487.

Wallis, G., & Bülthoff, H. H. (2001). Effects of temporal association on recognition memory. *PNAS, 98*(8), 4800–4804.

Watson, T. L., Hill, H., & Johnston, A. (2002). View invariance in facial motion. *Perception, 31*, 119.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling and goodness-of-fit. *Perception and Psychophysics, 63*(8), 1293–1313.

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