THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 2015
Vol. 68, No. 9, 1832–1843, http://dx.doi.org/10.1080/17470218.2014.993664

Routledge
Taylor & Francis Group

# Categorizing identity from facial motion

**Christine Girges, Janine Spencer, and Justin O'Brien**
Department of Psychology, Brunel University, Uxbridge, UK

*(Received 29 July 2014; accepted 22 November 2014; first published online 17 February 2015)*

Advances in marker-less motion capture technology now allow the accurate replication of facial motion and deformation in computer-generated imagery (CGI). A forced-choice discrimination paradigm using such CGI facial animations showed that human observers can categorize identity solely from facial motion cues. Animations were generated from motion captures acquired during natural speech, thus eliciting both rigid (head rotations and translations) and nonrigid (expressional changes) motion. To limit interferences from individual differences in facial form, all animations shared the same appearance. Observers were required to discriminate between different videos of facial motion and between the facial motions of different people. Performance was compared to the control condition of orientation-inverted facial motion. The results show that observers are able to make accurate discriminations of identity in the absence of all cues except facial motion. A clear inversion effect in both tasks provided consistency with previous studies, supporting the configural view of human face perception. The accuracy of this motion capture technology thus allowed stimuli to be generated that closely resembled real moving faces. Future studies may wish to implement such methodology when studying human face perception.

*Keywords*: Biological motion; Facial motion; Perception; Vision; Identity.

The mechanisms involved in facial identity recognition have been widely studied in both psychology (Bindemann, Attard, Leach, & Johnston, 2013; Burton, Wilson, Cowan, & Bruce, 1999) and neuropsychology (Pitcher, Walsh, Yovel, & Duchaine, 2007; Rhodes, Michie, Hughes, & Byatt, 2009; Solomon-Harris, Mullin, & Steeves, 2013). While these investigations have been highly informative, many of them utilize static stimuli such as photographs, line drawings, or CCTV images. Human faces, however, are intrinsically *dynamic* (Calder, Rhodes, Johnson, & Haxby, 2011). For example, verbal communication and emotional expressions occur via spatially distorting specific facial muscles. It is this continuous series of facial movement that provides an abundance of information necessary for all aspects of social cognition (Knappmeyer, Thornton, & Bülthoff, 2003; Krumhuber, Kappas, & Manstead, 2013; Stoesz & Jakobson, 2013).

Knight and Johnston (1997) were among the first to consider how movement influences the identity recognition of contrasted-reversed famous faces. They found accuracy to improve only when faces were displayed dynamically relative to a single static image. Later studies implementing other types of impoverished stimuli (threshold processed videos, blurred/pixelated clips, or limited frame sequences) also report a beneficial effect of motion during familiar face recognition (Lander

---

Correspondence should be addressed to Christine Girges, Department of Psychology, Brunel University, Uxbridge, UB8 3PH, UK. E-mail: Christine.girges@brunel.ac.uk

We thank the individuals who participated in the current study and David Bell for his valuable advice on animating and rendering. No funding sources declared.

& Bruce, 2000; Lander, Bruce, & Hill, 2001; Lander, Christie, & Bruce, 1999).

It suggests that motion provides three-dimensional (3D) information concerning face structure, but also prompts the recognition of idiosyncratic movements during suboptimal viewing conditions (O'Toole, Roark, & Abdi, 2002). Others argue that this does not necessarily reflect a true dynamic effect though, and recognition might actually improve because the number of static frames contained within a moving sequence increases (Lander & Chuang, 2005). Lander et al. (1999) have, however, shown that when the same frames were displayed as either a static array or an animated sequence, identity recognition was still significantly higher for the moving sequence.

Several investigations have sought to examine which features of facial movement drive this increase in perception. Faces move in rigid (transient changes in head orientation) and nonrigid (expressional changes) manners. Both these categories appear to improve identity recognition (Bruce & Valentine, 1988). Pike, Kemp, Towell, and Phillips (1997) required participants to learn unfamiliar faces from static pictures (single and multiple sequences) or dynamic clips exhibiting rigid movement. At test, a single image was shown, and the task was to decide whether the face was present in the previous learning phase. The authors found that identity recognition was more accurate for faces initially presented as rigid motion sequences. Similarly, Thornton and Kourtzi (2002) observed a matching advantage for prime images of nonrigid motion (short video sequences) relative to a single static image. The benefit of nonrigid motion appears to exist regardless of task type (sequential matching versus visual search) or viewpoint (Pilz, Thornton, & Bülthoff, 2006).

Others have failed to observe advantages for faces viewed in motion over static pictures (e.g., Bonner, Burton, & Bruce, 2003; Lander & Bruce, 2003; Lee, Habak, & Wilson, 2010). Christie and Bruce (1998) found no improvement in the recognition of unfamiliar faces exhibiting rigid motion (shaking and nodding) compared to multiple static views. Lander and Chuang (2005) later replicated this finding using degraded movies of familiar and famous faces moving rigidly. Discrepant data could reflect an experimental bias caused by testing different age groups. For example, younger adults performed better than older adults when matching a learned dynamic (rigid or nonrigid) face to a static test image (Maguinness & Newell, 2014). Otsuka et al. (2009) suggests that adults benefit less from motion as their perceptual abilities are already optimal. In younger participants, however, face processing systems are less developed and need the additional data that facial motion provides.

Alternatively, the type of stimuli implemented across studies could contribute to inconsistent findings. Many use unnatural stimuli such as edited videos of image sequences (e.g., Lander & Bruce, 2003) or synthetic faces depicting computer-stimulated motion (e.g., Lee et al., 2010). These representations of facial movement may not fully capture the mechanisms underlying its perception (Schultz & Pilz, 2009). Those who do utilize naturalistic videos often do not control for irrelevant nonmotion data or residual spatial cues. For example, Lander and Bruce (2000) displayed video clips of people sometimes shown from the waist upwards. The addition of such information could confound perception.

To address this, Hill and Johnston (2001) first described a method to explore motion-based information independently of other cues. Facial animations were generated by applying the motion captured from 12 actors to the same 3D computer-generated imagery (CGI) face. The technique also allowed the authors to separate rigid and nonrigid motion, thus evaluating their contributions to the categorization of identity and gender, respectively. These stimuli have since been successfully implemented in other studies examining the discrimination of individual faces (O'Brien, Spencer, Girges, Johnston, & Hill, 2014; Spencer, O'Brien, Johnston, & Hill, 2006), viewpoint dependence (Watson, Johnston, Hill, & Troje, 2005), and neural correlates of facial motion perception within the visual cortex (Girges, Wright, Spencer, & O'Brien, 2014).

The current study builds upon the work of Hill and Johnston (2001) who used animations derived from marker-based motion capture. Here, recent developments in marker-less technology were exploited to generate accurate and realistic models of facial movement. Simultaneous sequences of rigid and nonrigid motion were recorded from real actors and were applied to a CGI face. To assess these stimuli, participants completed two tasks. In a video discrimination paradigm, a target video had to be matched to two subsequently presented animations. One was of the same target, the other a completely different foil animation. In the identity discrimination condition, the same experimental format was used. However, the two options were either of the original actor reciting another poem or of a completely different actor. The task was to choose the same actor. As the appearances of animations were identical to each other, judgements were based solely on differences in motion patterns. If facial motion is indeed a cue for identity, these tasks should be completed with minimum error rates.

Facial motion captures were also presented upside down. Inversion paradigms affect static face recognition by disrupting configural processing and early structural encoding of facial features (Itier & Taylor, 2002; Valentine, 1998). A similar effect has been found for moving faces, in which inversion impairs the ability to accurately discriminate gender and identity (Hill & Johnston, 2001). While such data suggest that facial motion is perceived configurally, others argue that it utilizes part-based processing and therefore bypasses the inversion effect (e.g., Knappmeyer et al., 2003; Xiao, Quinn, Ge, & Lee, 2012). Due to such mixed findings, we aimed to investigate this further within the current study.

## EXPERIMENTAL STUDY

### Method

#### *Participants*

Ethical approval was obtained from Brunel University. Twenty individuals (6 male, 14 female, mean age = 33.45 years, range = 23–58 years) with normal or corrected-to-normal vision took part. Eighteen participants were right-handed, and two participants were left-handed (or ambidextrous). None of the sample had any history of neurological or psychological disorders. Participants were given a description of the study, and written informed consent was obtained. They were debriefed after the experiment was terminated. No reimbursement was given.

#### *Stimuli creation*

Fifteen nonprofessional human actors (7 male, 8 female) recited extracts from six short novelty poems (total of 90 different performances). Each poem contained a similar number of words and took approximately the same time to read. The extracts ranged in emotional content, therefore eliciting a variety of different natural facial expressions (nonrigid motion), mannerisms, head movements (rigid motion), and speech. Before recording, actors had a practice trial run to ensure they were familiar with the content and spoke clearly, fluently, and at an even pace.

A Kinect for Windows v2 sensor and Software Development Kit (SDK) captured the facial motion in 3D without the use of facial markers. The device featured an RGB camera (8-bit VGA resolution, 640 × 480 pixels) with a Bayer colour filter and both infrared and monochrome CMOS depth sensors (11-bit depth VGA resolution, 640 × 480 pixels, 2048 levels of sensitivity). As the sensor captured the 3D motion, images were reconstructed (via *Light Coding* scanner systems) and directly live streamed into a motion tracking software (FaceShift Studio 1.1, www.faceshift.com) at 30 frames per second (fps; Figure 1). Motion was tracked in real time, ensuring high accuracy.

In FaceShift, actors were first asked to elicit 23 training facial expressions prior to real motion recordings (neutral, open mouth, smile, brows down, brows up, sneer, jaw left, jaw right, jaw front, mouth left, mouth right, dimple, chin raise, pout, funnel, frown, m phoneme, grin, cheek puff, chew, lip down, eye blink left, and eye blink right). Scanning these sets of expressions enabled the program to calibrate each actor's motion and create a personalized avatar used for accurate motion tracking. Forty-eight blendshape parameters were tracked in total, meaning that

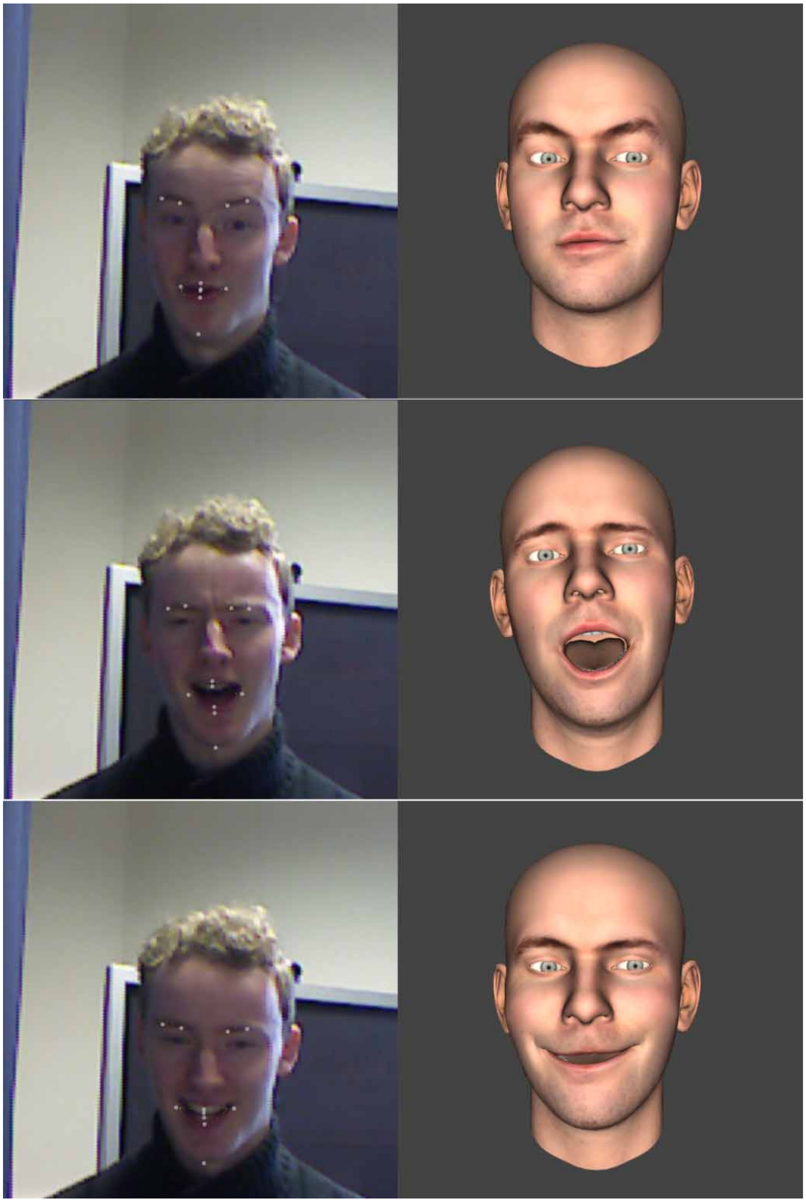

**Figure 1**. *Example of how the motion was tracked using the Kinect Sensor and FaceShift studio. The left panel of screenshots shows the real actor communicating. The right panel shows how the real motion is mapped onto an avatar in FaceShift. Note that this avatar was not the final model used in the experiment. To view this figure in colour, please visit the online version of this Journal.*

emotions of all magnitude, eye gaze (including eye blinks), and head pose were captured. Optimal recordings were best achieved by actors being seated 60 cm away from the sensor (sensors angular field of view = 57° horizontally and 43° vertically). Actors were allowed to adjust their seating position and move in their chair during recordings.

Offline, each complete performance was imported into a 3D CGI rendering and animation application (Roosendaal, 1995) as a.bvh file. These files contained 35 motion data points representing major facial regions (eyes, nose, chin, mouth, forehead, cheeks, and ears). These points were all connected to a common reference point (neck bone), which controlled any rigid motion present in the motion sequence. The reference point also preserved the relative spatial structure between each point so that they all moved correctly in relation to one another. The motion data was then "parented" to a greyscale computer-generated 3D face model[1] to begin the rigging process (Figure 2). Before this could happen, each individual motion point had to be readjusted to fit the computer-generated face. This was done by visual realignment and using a technique called "snapping", which placed each point on the surface of the model's skin. Once attached, the points essentially pulled and distorted the face into the specified motion pattern originally recorded from the Kinect Sensor (Figure 3). Any performances that did not map correctly onto the CGI faces or contained many artefacts were discarded from the database (although this was not the case here).

The greyscale face model was used for all 90 performances, allowing motion-based information to be measured independently from spatial cues. The appearances of all motion capture faces were identical to each other and only differed in the way they expressed motion. Each animation was encoded in h.264 format as an MP4 file. An orientation-inverted version of each animation was produced by rotating the stimuli along a 180-degree axis. None of the stimuli contained audio information.

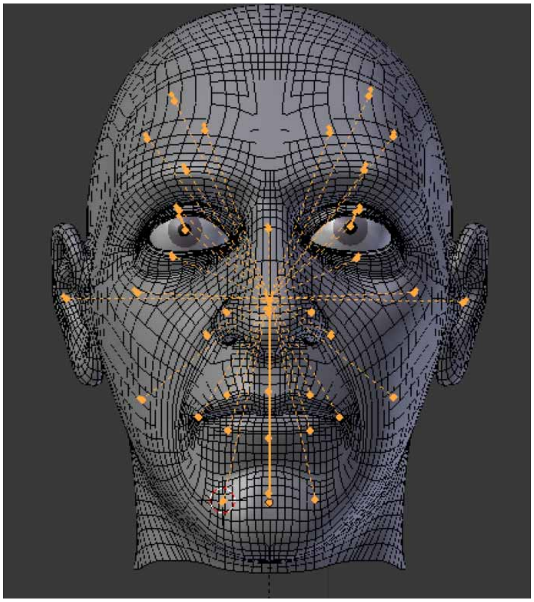

**Figure 2**. *Computer-generated face model with the motion data points attached to the major facial landmarks. To view this figure in colour, please visit the online version of this Journal.*

To ensure that the stimuli still represented the actual motion recorded from the original actors, we performed a small preliminary experiment with a different set of participants ($N = 15$, 7 male; age: $M = 33.20$ years, $SD = 12.04$). Participants observed a real video recording, followed by two facial motion animations presented side by side. Using a two-alternative forced-choice procedure, the task was to indicate which animation represented the real video. This carried on for 20 trials. On average, participants scored 18.40 out of 20 possible correct answers ($SD = 1.30$, range = 16–20, percentile score = 92%). A one-sampled $t$-test confirmed an above-chance performance (50%), $t(14) = 25.06$, $p < .001$, Cohen's $d = 13.36$.

### *Procedure*

The dynamic stimuli were presented using an LCD display with a resolution of 1024 × 768 pixels and a 60-Hz refresh rate. Viewing distance was 60 cm, at which the distance of the 30 × 22.5-cm display subtended an angle of 26.6° × 20.6°. The height of the average face was approximately 10.5°, and the frame rate of the animation was 25 fps.

Participants engaged in two experimental conditions, each with two manipulations (upright versus orientation-inverted facial motion).

---

[1] The computer-generated model was created by Kent Trammel and was available online to download from http://cgcookie.com/blender/author/theluthier/. The model was edited to achieve a more unisex appearance.

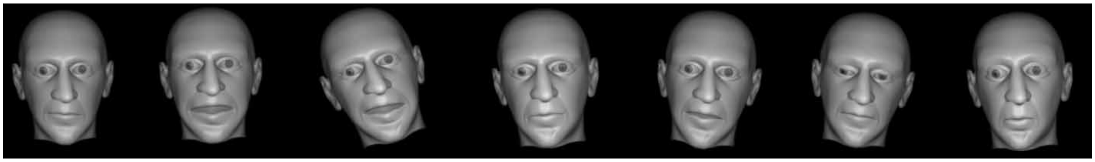

**Figure 3**. *Screenshots of final stimuli.*

Conditions comprised 21 experimental trials, plus four attention-control trials. Videos were edited so that only the first 5 s were shown, to ensure equal viewing durations. The same experimenter always sat behind the participants to manually record their verbal responses. No feedback was given. The average duration of the experiment was approximately 25 min.

The first condition consisted of discriminating between different *videos* of facial motion. A single facial animation was displayed in the centre of a black screen. Immediately after, the same animation was presented again plus a completely different animation (shown side by side). The foil animation was chosen at random and could have been from the same actor reciting a different poem. Using a two-alternative forced-choice procedure, participants had to indicate which stimuli (left or right) were present in both trials. A similar format was used for the second condition, in which participants were required to discriminate between different *identities* of facial motion. A single facial animation was selected at random, and its presentation was followed by another two animations. One was of the original actor reciting a different poem (target), and the other was of a completely different actor reciting any poem (foil). Using characteristic mannerisms and individuality of movements, participants had to discriminate which animation (left or right) represented the same individual from the first presentation.

All observers viewed the same combination of videos across trials. Male and female performances were not intermixed within the same trial to avoid indirect judgements based on gender. Each task (video or identity discrimination) was carried out using upright and orientation-inverted stimuli. The order of conditions was counterbalanced across participants to avoid familiarity effects, boredom, or fatigue influencing the behavioural data.

To ensure maximal attention to the stimuli throughout the testing period, attention-control trials were also included. This provided a conservative criterion for rejecting any data where there was a possibility of nonperceptual factors (fatigue, intermittent confusion) influencing the responses (Spencer & O'Brien, 2006). Attention controls occurred on every eighth trial in all conditions and were presented in a similar format to that for experimental trials. A single facial motion was displayed. Immediately after, the same video was shown again plus an orientation-inverted animation. Participants had to state which video (left or right) was present in both trials. As the orientation of one test stimuli was inverted, it could be excluded as a correct answer. Participants were aware that these manipulations would indicate the correct answer. The responses to these trials were not included in any subsequent analysis. All participants completed these trials without error; therefore no data were discarded.

### *Statistical analysis*

A one-sampled *t*-test (test value = 10.5) was applied to the data to ensure that all participants performed above chance level. A repeated measures analysis of variance (ANOVA) was then used to observe any main effects of *task type* (video, identity) and *orientation* (upright, orientation-inverted). Post hoc (paired-samples *t*-test) tests were applied where appropriate.

## Results

Table 1 presents the mean data (and standard deviations) from each task. Participants significantly performed above chance level on all tasks: *upright*

**Table 1.** *Mean correct scores for each task*

| *Task* | *Orientation* | *Mean* | SD | *% score* | *Range* |
|---|---|---|---|---|---|
| Video discrimination | Upright | 20.65 | 0.67 | 98.33 | 19–21 |
| | Orientation-inverted | 17.05 | 2.21 | 81.19 | 14–21 |
| Identity discrimination | Upright | 19.25 | 2.17 | 91.67 | 15–21 |
| | Orientation-inverted | 13.10 | 2.29 | 62.38 | 9–17 |

*Note:* Scores out of 21.

*video discrimination*, $t(19) = 67.67$, $p < .001$, Cohen's $d = 31.05$; *orientation-inverted video discrimination*, $t(19) = 13.24$, $p < .001$, Cohen's $d = 6.07$; *upright identity discrimination*, $t(19) = 18.01$, $p < .001$, Cohen's $d = 8.26$; and *orientation-inverted identity discrimination*, $t(19) = 5.07$, $p < .001$, Cohen's $d = 2.33$.

There was a significant main effect of task type, $F(1, 19) = 48.01$, $p < .001$, $\eta_p^2 = .72$, with participants scoring higher on video discrimination tasks. Orientation also produced a significant main effect, $F(1, 19) = 194.475$, $p < .001$, $\eta_p^2 = .91$. *T*-tests revealed a significant difference between the upright and orientation-inverted video scores, $t(19) = 7.71$, $p < .001$, Cohen's $d = 1.72$, and between the upright and orientation-inverted identity scores, $t(19) = 12.46$, $p < .001$, Cohen's $d = 2.79$. Such results indicate an inversion effect present in our sample. There was also a significant interaction between task type and orientation, $F(1, 19) = 32.51$, $p = .001$, $\eta_p^2 = .44$. A larger inversion effect occurred for identity discrimination tasks than for video discrimination tasks.

## Discussion

Building upon the pioneering work of Hill and Johnston (2001), recent developments in markerless technology were exploited to generated models of facial movement. Simultaneous sequences of rigid and nonrigid motion (including eye gaze and blinks) were recorded from real people and were applied to a CGI display. These animations were used to evaluate motion-specific contributions in the categorization of identity.

### *Recognition from facial motion*

Consistent with our hypotheses, participants were able to distinguish between different facial motion videos and to discriminate the faces of unfamiliar individuals. Other studies of face learning and recognition from dynamic stimuli report parallel findings (e.g., Hill & Johnston, 2001; Knight & Johnston, 1997; Pilz et al., 2006). Similar results have also been documented in infant populations (e.g., Otsuka et al., 2009). Spencer et al. (2006) reported that infants between 4 and 8 months could discriminate sequences of facial motion and the identity of a speaker. Layton and Rochat (2007) observed an effect of motion cues at 8 months of age when infants viewed familiar faces (their mother's face). Bulf and Turati (2010) extended these findings, demonstrating that newborns were able to recognize the profile pose of unfamiliar faces moving rigidly. Evidently, the ability to perceive and utilize facial motion is acquired very early on in visual development (see Xiao et al., 2014, for a review).

There are two prominent hypotheses regarding how facial motion influences identification processes (O'Toole et al., 2002). First, the "*supplemental information hypothesis*" states that idiosyncratic facial movements aid identification. This cue may be particularly useful when recognizing already familiar faces. For example, you might identify a close friend by the way they smile, or characteristically nod their head during conversations. By contrast, the "*representation enhancement hypothesis*" suggests that facial motion provides a better structural depiction of a 3D face. Learning new faces benefits from such enhancement. The number of viewpoints increases, therefore refining mental representations of less familiar faces (O'Toole et al., 2002).

While these hypotheses describe two different ways in which facial motion contributes to identity judgements, it does not mean that they are strictly exclusive for a specific type of recognition. Rather, they are interlinked. There is some neuroimaging evidence to support this conclusion. Encoding new views of an individual has been thought to operate within the fusiform gyrus (Longmore, Liu, & Young, 2008), while identifying characteristic motion takes place within a portion of the superior temporal sulcus (Longmore & Tree, 2013). Past studies have shown both regions to be collectively active during facial motion perception (Furl et al., 2010; Schultz & Pilz, 2009). It is unclear, though, whether ventral–temporal areas are showing a true dynamic response, or are simply sensitive to the static information contained within a motion sequence (Schultz, Brockhaus, Bülthoff, & Pilz, 2013).

As with static face perception, inverting the stimuli significantly reduced participants' ability to discriminate video sequences or recognize the faces of different individuals. This was particularly true for judgements concerning facial identity, perhaps reflecting task complexity. To successfully discriminate different identities, participants had to perceive characteristic mannerisms, which would have been more difficult to do when the animations were inverted. It seems that the inversion effect is sensitive to task type and what information must be extracted. Many preexisting studies report similar inversion effects with dynamic stimuli (Longmore & Tree, 2013; Watson et al., 2005). Observers were poor on tasks requiring them to judge the gender of an inverted dynamic face (Thornton, Mullins, & Banahan, 2011). It suggests that motion information is processed configurally by a system tuned to upright faces, rather than by extraction of low-level cues (Hill & Johnston, 2001; Watson et al., 2005).

In contrast, facial motion might utilize part-based processing and bypass the disruption caused by inversion. Indeed, a less pronounced inversion effect has been observed when faces are shown dynamically (Hill & Johnston, 2001; Knappmeyer et al., 2003; Lander et al., 1999). More recent investigations using composite faces also support the featural influence hypothesis of facial motion perception. Xiao et al. (2012) found that the upper and lower portions of composite faces were processed in a part-based manner, allowing participants to identify the test faces more accurately. Xiao, Quinn, Ge, and Lee (2013) later replicated and extended these findings, reporting a smaller composite effect for elastic (nonrigid) facial motion. The current data are evidently mixed, and further clarification is needed. It may be that dynamic faces are still subjected to the inversion effect, but the addition of motion minimizes the disruption.

### *Comparison of methodology with other approaches*

A handful of face perception studies implement dynamic stimuli inspired by the Facial Action Coding System (FACS; Ekman & Friesen, 1978). FACS quantifies all possible facial muscle expressions and decomposes them into action units. Each unit is then plotted as a time course so that the spatiotemporal properties of local movements can be represented. This technique has been applied to motion-capture data to create highly controlled and meaningful facial animations (e.g., Curio et al., 2006; Dobs et al., 2014). The advantage here is that facial motion is annotated accurately and precisely with reference to underlying muscle activations. It is also easy to retarget motion onto any face model that uses the same semantic structure (Curio et al., 2006). Yet, these FACS-derived animations typically present only nonrigid motion—that is, facial expressions without changes in viewpoint. Head position and orientation, however, represent a powerful cue, especially with reference to identity recognition (e.g., Hill & Johnston, 2001). The stimuli we present may therefore be more suitable when studying face perception.

Further, marker-based motion capture records data from a predefined set of facial points. Because of this, subtle or extremely implicit facial movements occurring in other "unmarked" areas are disregarded. The method described here minimizes such issues. Motion in all face regions was recorded, resulting in extremely detailed and naturally fluid animations. As the animations closely resembled real human facial movement in the absence of spatial cues, it is possible to generalize the current data to faces in real life. Indeed, there is evidence that synthetic and natural faces are processed by similar neural mechanisms (but see Han, Jiang, Humphreys, Zhou, & Cai, 2005). Moser et al. (2007) demonstrated that avatars elicit similar patterns of activation to human faces, particularly in the emotion-sensitive amygdala.

As a side note, viewing such motion-rich stimuli could explain why some observers performed at ceiling in all but one condition (orientation-inverted identity discrimination). Stimuli high in detail would provide much information concerning identity, which in turn would facilitate perception. On the other hand, this could reflect aspects of our task design. Stimuli were presented consecutively without delays, and participants were asked to make their decision immediately after each trial. This decrease in working memory could have evoked superior levels in face recognition (Weigelt, Koldewyn, & Kanwisher, 2013).

### *Limitations and future directions*

While the current data indicate a significant ability in categorizing identity from facial motion, it is possible that observers could do this just as easily with multiple static frames or snapshots of different head positions (Lander & Chuang, 2005). We are disinclined to accept this view, however, as others have shown that it is the dynamic quality of motion, rather than the amount of static information, that is crucial for recognition (Lander & Bruce, 2000). For example, Lander et al. (1999) report that identity recognition was better with moving sequences than with a static array even though both stimuli contained the same amount of frames.

Moreover, several papers have attempted to discover which aspect of facial motion contributes to recognition. Unfortunately they provide mixed results. Hill and Johnston (2001) suggest that it is head rotations/translations that are useful when categorizing identity. However, at least three research groups have shown no advantage for rigid motion compared to static images (e.g., Christie & Bruce, 1998; Lander & Chuang, 2005; Lee et al., 2010). As rigid and nonrigid motion cues were not separated within the current study, we cannot comment on what aspect is driving the performance here. Of course, it may be that perception is facilitated by a combination of both cues. In real life, changeable facial expressions and head movements are encountered simultaneously rather than in isolation. Supporting this assumption, Maguinness and Newell (2014) report that motion facilitates face learning across changes in both viewpoint (rigid) and expression (nonrigid).

Future studies are encouraged to extend this experiment by implementing conditions that compare performances based on rigid motion, nonrigid motion, and combined motion cues. This would facilitate our understanding of what type of facial movement is facilitating its perception.

In addition, the stimuli method could be applied to the study of emotion recognition. It has been previously shown that dynamic presentations aid recognition, yet these conclusions are derived from implied or morphed videos (Bould & Morris, 2008; Puce, Allison, Bentin, Gore, & McCarthy 1998). Implementing such facial motion captures would significantly help in fully understanding the underlying mechanisms.

## CONCLUSION

Taken together, we provide a new method to create facial motion stimuli that are free from surfaced-based visual cues yet are still realistic and accurate. While similar to those used by Hill and Johnston (2001), the current marker-less animations contain much more detail and move more naturally. From the use of such advanced stimuli, we have shown that adult observers are able to perceive facial motion and can use it to make sensible categorizations concerning unfamiliar facial identities.

As such discrimination is of a social nature, we provide further evidence that facial motion has a prominent role in social cognition. The data also support the configural view of human face perception.

## REFERENCES

Bindemann, M., Attard, J., Leach, A. M. Y., & Johnston, R. A. (2013). The effect of image pixelation on unfamiliar-face matching. *Applied Cognitive Psychology*, *27*, 707–717. doi:10.1002/acp.2970

Bonner, L., Burton, A. M., & Bruce, V. (2003). Getting to know you: How we learn new faces. *Visual Cognition*, *10*(5), 527–536. doi:10.1080/13506280244000168

Bould, E., & Morris, N. (2008). Role of motion signals in recognizing subtle facial expressions of emotion. *British Journal of Psychology*, *99*(Pt 2), 167–189. doi:10.1348/000712607X206702

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, et al. (Eds.), *Practical aspects of memory: Current research and issues: Vol. 1. Memory in everyday life* (69–174). Oxford, UK: Wiley.

Bulf, H., & Turati, C. (2010). The role of rigid motion in newborns' face recognition. *Visual Cognition*, *18*(4), 504–512. doi:10.1080/13506280903272037

Burton, A. M., Wilson, S., Cowan, M., & Bruce, V. (1999). Face recognition in poor-quality video: Evidence from security surveillance. *Psychological Science*, *10*, 243–248. Retrieved from http://dx.doi.org.v-ezproxy.brunel.ac.uk:2048/10.1111/1467-9280.00144

Calder, A., Rhodes, G., Johnson, M., & Haxby, J. (2011). Oxford handbook of face perception. New York, USA: Oxford University Press.

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

Curio, C., Breidt, M., Kleiner, M., Vuong, Q. C., Giese, M. A., & Bülthoff, H. H. (2006). Semantic 3D motion retargeting for facial animation. In *Proceedings of the 3rd symposium on applied perception in graphics and visualization* (pp. 77–84). New York, NY: ACM Press.

Dobs, K., Bülthoff, I., Breidt, M., Vuong, Q. C., Curio, C., & Schultz, J. (2014). Quantifying human sensitivity to spatio-temporal information in dynamic faces. *Vision Research*, *100*, 78–87. doi:10.1016/j.visres.2014.04.009

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system: A technique for the measurement of facial movement*. Palo Alto, CA: Consulting Psychologists Press.

FaceShift Studio (Version 1.1) [Computer software]. Zurich, Switzerland. Retrieved from http://www.faceshift.com

Furl, N., van Rijsbergen, N. J., Kiebel, S. J., Friston, K. J., Treves, A., & Dolan, R. J. (2010). Modulation of perception and brain activity by predictable trajectories of facial expressions. *Cerebral Cortex*, *20*(3), 694–703. doi: 10.1093/cercor/bhp140

Girges, C., Wright, M. J., Spencer, J. V., & O'Brien, J. M. D. (2014). Event-related alpha suppression in response to facial motion. *PloS One*, *9*(2), e89382. doi:10.1371/journal.pone.0089382

Han, S., Jiang, Y., Humphreys, G. W., Zhou, T., & Cai, P. (2005). Distinct neural substrates for the perception of real and virtual visual worlds. *NeuroImage*, *24* (3), 928–935. doi:10.1016/j.neuroimage.2004.09.046

Hill, H., & Johnston, A. (2001). Categorizing sex and identity from the biological motion of faces. *Current Biology: CB*, *11*(11), 880–885. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11516651

Itier, R. J., & Taylor, M. J. (2002). Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: A repetition study using ERPs. *NeuroImage*, *15*(2), 353–372. doi:10.1006/nimg.2001.0982

Knappmeyer, B., Thornton, I. M., & Bülthoff, H. H. (2003). The use of facial motion and facial form during the processing of identity. *Vision Research*, *43*(18), 1921–1936. doi:10.1016/S0042-6989(03)00236-0

Knight, B., & Johnston, A. (1997). The role of movement in face recognition. *Visual Cognition*, *4*(3), 265–273. doi:10.1080/135062897395462

Krumhuber, E. G., Kappas, A., & Manstead, A. S. R. (2013). Effects of dynamic aspects of facial expressions: A review. *Emotion Review*, *5*(1), 41–46.

Lander, K., & Bruce, V. (2000). Recognizing famous faces: Exploring the benefits of facial motion. *Ecological Psychology*, *12*(4), 259–272. doi:10.1207/S15326969ECO1204_01

Lander, K., & Bruce, V. (2003). The role of motion in learning new faces. *Visual Cognition*, *10*(8), 897–912. doi:10.1080/13506280344000149

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. doi:10.1002/1099-0720(200101/02)15:1&lt;101::AID-ACP697>3.0.CO;2-7

Lander, K., Christie, F., & Bruce, V. (1999). The role of movement in the recognition of famous faces. *Memory and Cognition*, *27*, 974–985. Retrieved from http://dx.doi.org.v-ezproxy.brunel.ac.uk:2048/10.3758/BF03201228

Lander, K., & Chuang, L. (2005). Why are moving faces easier to recognize? *Visual Cognition*, *12*(3), 429–442. doi:10.1080/13506280444000382

Layton, D., & Rochat, P. (2007). Contribution of motion information to maternal face discrimination in infancy. *Infancy*, *12*, 257–271. doi:10.1080/15250000701518594

Lee, Y., Habak, C., & Wilson, H. R. (2010). Seeing an unfamiliar face in rotational motion does not aid identity discrimination across viewpoints. *Vision Research*, *50*(9), 854–859. doi:10.1016/j.visres.2010.02.013

Longmore, C. A., Liu, C. H., & Young, A. W. (2008). Learning faces from photographs. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(1), 77–100. doi:10.1037/0096-1523.34.1.77

Longmore, C. A., & Tree, J. J. (2013). Motion as a cue to face recognition: Evidence from congenital prosopagnosia. *Neuropsychologia*, *51*(5), 864–875. doi:10.1016/j.neuropsychologia.2013.01.022

Maguinness, C., & Newell, F. N. (2014). Motion facilitates face perception across changes in viewpoint and expression in older adults. *Journal of Experimental Psychology: Human Perception and Performance*. Advance online publication. http://dx.doi.org/10.1037/a0038124

Moser, E., Derntl, B., Robinson, S., Fink, B., Gur, R. C., & Grammer, K. (2007). Amygdala activation at 3T in response to human and avatar facial expressions of emotions. *Journal of Neuroscience Methods*, *161*(1), 126–133. doi:10.1016/j.jneumeth.2006.10.016

O'Brien, J., Spencer, J., Girges, C., Johnston, A., & Hill, H. (2014). Impaired perception of facial motion in autism spectrum disorder. *PloS One*, *9*(7), e102173. doi:10.1371/journal.pone.0102173

O'Toole, A., Roark, D., & Abdi, H. (2002). Recognition of moving faces: A psychological and neural synthesis. *Trends in Cognitive Science*, *6*, 261–266. doi:10.1016/S1364-6613(02)01908-3

Otsuka, Y., Konishi, Y., Kanazawa, S., Yamaguchi, M. K., Abdi, H., & O'Toole, A. J. (2009). Recognition of moving and static faces by young infants. *Child Development*, *80*(4), 1259–1271. doi:10.1111/j.1467-8624.2009.01330.x

Pike, G. E., Kemp, R. I., Towell, N. A., & Phillips, K. C. (1997). Recognising moving faces: The relative contribution of motion and perspective view information. *Visual Cognition*, *4*, 409–437. Retrieved from http://dx.doi.org.v-ezproxy.brunel.ac.uk:2048/10.1080/713756769

Pilz, K. S., Thornton, I. M., & Bülthoff, H. H. (2006). A search advantage for faces learned in motion. *Experimental Brain Research*, *171*(4), 436–447. doi:10.1007/s00221-005-0283-8

Pitcher, D., Walsh, V., Yovel, G., & Duchaine, B. (2007). TMS evidence for the involvement of the right occipital face area in early face processing. *Current Biology*, *17*(18), 1568–1573. doi:10.1016/j.cub.2007.07.063

Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience*, *18*(6). 2188. Retrieved from http://www.jneurosci.org/content/18/6/2188.full.pdf+html

Rhodes, G., Michie, P. T., Hughes, M. E., & Byatt, G. (2009). The fusiform face area and occipital face area show sensitivity to spatial relations in faces. *The European Journal of Neuroscience*, *30*(4), 721–733. doi:10.1111/j.1460-9568.2009.06861.x

Roosendaal, T. (1995). Blender (Version 2.66) [Computer software]. Amsterdam, Netherlands. Retrieved from http://www.blender.org

Schultz, J., Brockhaus, M., Bülthoff, H. H., & Pilz, K. S. (2013). What the human brain likes about facial motion. *Cerebral Cortex*, *23*(5), 1167–1178. doi:10.1093/cercor/bhs106

Schultz, J., & Pilz, K. S. (2009). Natural facial motion enhances cortical responses to faces. *Experimental Brain Research*, *194*, 465–475. doi:10.1007/s00221-009-1721-9

Solomon-Harris, L. M., Mullin, C. R., & Steeves, J. K. E. (2013). TMS to the "occipital face area" affects recognition but not categorization of faces. *Brain and Cognition*, *83*(3), 245–251. doi:10.1016/j.bandc.2013.08.007

Spencer, J., O'Brien, J., Johnston, A., & Hill, H. (2006). Infants' discrimination of faces by using biological motion cues. *Perception*, *35*(1), 79–89. doi:10.1068/p5379

Spencer, J. V., & O'Brien, J. M. D. (2006). Visual form-processing deficits in autism. *Perception*, *35*(8), 1047–1055. doi:10.1068/p5328

Stoesz, B. M., & Jakobson, L. S. (2013). A sex difference in interference between identity and expression judgments with static but not dynamic faces. *Journal of Vision*, *13*, 1–14. doi:10.1167/13.5.26.

Thornton, I., Mullins, E., & Banahan, K. (2011). Motion can amplify the face-inversion effect. *Psihologija*, *44*(1), 5–22. doi:10.2298/PSI1101005T

Thornton, I. M., Kourtzi, Z. (2002). A matching advantage for dynamic human faces. *Perception*, *31*, 113–132. Retrieved from http://dx.doi.org.v-ezproxy.brunel.ac.uk:2048/10.1068/p3300

Valentine, T. (1998). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology*, *79*, 471–491. Retrieved from http://search.proquest.com.v-ezproxy.brunel.ac.uk:2048/docview/1293583110?accountid=14494

Watson, T., Johnston, A., Hill, H., & Troje, N. (2005). Motion as a cue for viewpoint invariance. *Visual Cognition*, *12*(7), 1291–1308. doi:10.1080/13506280444000526

Weigelt, S., Koldewyn, K., & Kanwisher, N. (2013). Face recognition deficits in autism spectrum disorders are both domain specific and process specific. *PloS One*, *8*(9), e74541. doi:10.1371/journal.pone.0074541

Xiao, N. G., Perrotta, S., Quinn, P. C., Wang, Z., Sun, Y.-H. P., & Lee, K. (2014). On the facilitative effects of face motion on face recognition and its development. *Frontiers in Psychology*, *5*, 1–16. doi:10.3389/fpsyg.2014.00633

Xiao, N. G., Quinn, P. C., Ge, L., & Lee, K. (2012). Rigid facial motion influences featural, but not holistic, face processing. *Vision Research*, *57*, 26–34. doi:10.1016/j.visres.2012.01.015

Xiao, N. G., Quinn, P. C., Ge, L., & Lee, K. (2013). Elastic facial movement influences part-based but not holistic processing. *Journal of Experimental Psychology. Human Perception and Performance*. Advance online publication. doi:10.1037/a0031631